R E P O R T   R E S U M E S

ED 017 167   56   EM 005 975

PROGRAMED INSTRUCTIONS AND THE TEACHING OF MATHEMATICS.

BY- STOLUROW, LAWRENCE M.

ILLINOIS UNIV., URBANA, BUR. OF EDUC. RESEARCH

REPORT NUMBER PN-711151-01   PUB DATE 29 DEC 63

REPORT NUMBER 7A-806

EDRS PRICE MF-$0.25 HC-$0.88 20P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, *MATHEMATICS INSTRUCTION, DISCOVERY LEARNING, LEARNING PROCESSES, LEARNING THEORIES, *CURRICULUM RESEARCH, *METHODS RESEARCH,

THIS PAPER REVIEWS VARIABLES AND TRENDS IN RESEARCH ON PROGRAMED INSTRUCTION IN MATHEMATICS. IT WAS PRESENTED AT THE AAAS SYMPOSIUM ON LEARNING RESEARCH PERTINENT TO EDUCATIONAL IMPROVEMENT, CLEVELAND, OHIO, DEC. 29. 1963. (LH)

EM 005 975 BR. Title VII -A 806

PA-56 A

ED017167

TRL

TRAINING RESEARCH LABORATORY

Department of Psychology Bureau of Educational Research

University of Illinois 8 Lincoln Hall Urbana, Illinois

PROGRAMED INSTRUCTIONS

and the

TEACHING of MATHEMATICS

Lawrence M. Stolurow

Paper presented at the AAAS
Symposium on Learning Research
Pertinent to Educational Improvement,
Cleveland, Ohio, December 29, 1963.

U. S. Office of Education,
Title VII of the National
Defense Education Act,
Educational Media Branch
Project No. 711151.01

Programed Instruction and the Teaching of Mathematics[1]

Lawrence M. Stolurow
University of Illinois

This paper presents a summary of research on the teaching and learning of mathematics by programed instructional procedures. The research and findings are considered from a particular point of view with respect to their relationship to the developing technology of education.

Studies of Mathematics Teaching and Learning

It is not surprising to find that research on the teaching of mathematics with programed self-instructional materials is relatively extensive compared with other areas when it is realized that the proportion of self-instructional programs in mathematics is the largest of all subject matters (see Hanson, 1963). Most of the programs and the research relate to secondary school level mathematics; however, both the programs and the research actually range from elementary school through college. Topics covered include arithmetic, algebra (including Boolean), geometry, sets relations and functions, number theory, trigonometry, vectors, not to mention such areas of applied mathematics as conventional statistics (e.g., Hickey, Autor & Robinson, 1962) and linear programing for management decision making (e.g., Glaser and Reynolds, 1962).

---

[1]Paper presented at the AAAS Symposium on Learning Research Pertinent to Educational Improvement, Cleveland, Ohio, December 29, 1963. The University of Illinois work reported herein and the preparation of this report was made possible by a grant from the U. S. Office of Education, Title VII of the National Defense Education Act, Educational Media Branch Project No. 711151.01 titled: Comparative Studies of Principles for Programing Mathematics in Automated Instruction. Co-Investigators are Professor Max Beberman and Lawrence M. Stolurow.

Two Trends

Two trends, each with different objectives, are dominant in the research on mathematics teaching using programed self-instructional materials. One consists in the uses of mathematics as a convenient subject matter vehicle with which to study basic problems relating to the technology of self-instruction. The other consists in the study of mathematics as a conceptual, intellectual and behavioral domain. In some research, both of these objectives are involved for it is efficient to pursue the latter objective while also considering a technological problem.

The technological and substantive trends are by no means equivalent in their development nor is the pattern which they reveal the most logical one from all points of view. For example, it could be argued that the technological research which deals with problems of synthesis should follow the substantive research which deals with problems of analysis of behavior. From this point of view, the argument would be that we need to know what behavior we are to synthesize before we work on the techniques for accomplishing the syntheses. The fact that behavioral organization, or shaping, is indirectly accomplished by means of content control and the fact that this control can be exercised at different levels makes it possible to work on problems of synthesis while those of analysis are just beginning at a more molecular level. Any survey of research needs to keep the different levels of analysis clearly in mind for the research is to be related to the particular level of behavioral analysis that was used.

Problem of behavioral units. The use of the terms analysis and synthesis of behavior immediately suggest the need for some specification of the units used. This is an important unsolved problem that must be considered on intuitive grounds at the moment. There is little doubt, however, that there are molecular and molar elements of behavior. Neurophysiological analysis of behavior is clearly more molecular than an analysis of behavior in terms of observable gross movements of limbs and torso. Similarly, these movements are molecular in relation to the most molar aspects of behavior contained in a description of learning sets (e.g., Gagne' and Paradise, 1961) contributing to the solving of equations, e.g., "simplifying an equation by adding and subtracting terms to both sides" (Ibid, p. 6). Important in the analysis of behavior repertoires is the unit of analysis employed. Most current psychological theories of learning deal with units much more molecular than those of concern to the educator (see Stolurow, 1964). This difference in units used to describe behavior probably accounts for some of the failure in communication between the educator and psychologist, and the problem of behavioral units is critical to the technology of education. Needless to say, this problem arises in the research on the teaching of mathematics. Unfortunately, however, it is not one of the problems on which there is active research. The purpose in raising it here is that it is basic to the present treatment and examination of the research on mathematics teaching.

S R language. One way in which this problem of units enters rather obviously is in the application of the language of stimulus and response to behaviors more molar than those to which these terms are traditionally

applied. There are many different theories of learning that use S R language and not all of them use the terms stimulus and response to refer to environmental events and behavior at the same level of description. Guthrie, for example, uses S and R to refer to more molecular events than those to which Skinner applied these same labels (see Hilgard, 1956). Consequently, the application of S R language to the analysis of educationally relevant behaviors does not also imply the application of an S R theory of learning. Rather, its use is for objectivity in communication and description so as to minimize surplus meaning and to permit operational descriptions of material and procedures.

Studies Relating to the Technology of Programed Instruction

It seems useful to distinguish two types of studies relating to the technology of programed instruction. One is concerned with analysis and has implications for the psychological architecture of cognitive structures designed for educational purposes. The other is concerned with synthesis and the problems of construction of the cognitive and strategic structures which is the business of education. For purpose of this paper, the former will be referred to as architectural studies and the latter as engineering studies.

Studies With Engineering Implications

The research on programed self-instruction has concerned itself with the technology of teaching. This same emphasis exists in the research using mathematics materials. The implementation, or engineering, problems predominate and comprise the bulk of the research if not its more exciting developments.

Response Form

The form of the response to be used in a learning situation is settled upon as a result of a variety of considerations, one important one of which is the effect upon learning, retention or transfer. Thus, it is relevant to the extension of learning theory into educational engineering for us to examine the implications of various forms of response in relation to these three processes. In doing this, however, it is important to consider the form of response in relation to the student's repertoire. A response that is in the student's repertoire in the exact form required by the new learning experience is in a different class from a response that is not. The relevant factor in the design of a behavioral structure is the form of the behavior that is to be used. If it is already formed as required, then the engineering problem is one of putting it under stimulus control where the stimuli are at the proper level for the desired performance. However, if it is not already formed, the engineering problem is to assemble or shape that behavior which is available.

Once the psychology problems have been considered, then engineering decisions depend upon factors not directly related to the psychological outcomes or objectives. For example, the visibility of the response may be a factor, particularly in the early stages of development of a program, or in a new use of an established program as with a younger group of students than those on which the program was developed and is known to work.

Overt vs. covert response. The research on overt and covert response in programed instruction has indicated that the use of covert responses results in equivalent achievement in less time (e.g., Lambert, Miller & Wiley, 1962;

Stolurow and Walker, 1962). Consequently, the visibility requirements should prevail once it is established that the behavior is already in the students repertoire.

The finding that overt response does not add to learning in some mathematics programs (e.g., Lambert, et al, 1962; Stolurow and Walker, 1962) is of interest, since it might be assumed that responses required in learning parts of mathematics would not be in the student's repertoire. Since the number of studies and areas of mathematics used have been so small, it would be hazardous to generalize the present findings to all of mathematics. Certainly, the requirements to make the response visible are sufficient to warrant continued use of overt behavior in a mathematics course.

Constructed vs. multiple choice response. The psychological issues here are comparable to those associated with the overt-covert studies and relate to response availability. Consequently, the use of constructed or multiple choice response becomes a question of the probable existence of the desired response in the student's repertoire. If the response does exist, then the use of multiple choice permits the student to make his responses visible without also introducing delays that would occur if they were constructed. Data from mathematics are meager; however, Price (1962) compared these two response modes using mentally retarded students. The results can be interpreted in terms of response availability, for he found that the multiple choice mode resulted in superior performance when the students learned subtraction but not when they learned addition.

Stimulus Encoding

There are several problems relating to the presentation of mathematical concepts for efficient teaching. For example, the use of "boxes" or simple geometric forms such as squares and circles to represent variables instead of letters of the alphabet is a case in point. Apparently, empty boxes that could contain a variety of different numerals is a superior form of encoding to the use of letters, particularly for students at the lower ages. Another problem concerns the choice between algebraic and geometric presentations of a problem. In some unpublished studies[2], for example, we have required students to learn a formula that applies to some, but not all features of a display. It was found that a few students gave geometric solutions, whereas most of them gave algebraic solutions To check this finding, some groups were deliberately given a geometric solution principle and others an algebraic solution principle. The latter performed better than the former, even though the two solution principles were potentially equivalent in effectiveness. Some ambiguous data relating to this problem come from a study by Hickey, et al, (1962). These investigators failed to find differences which they expected to in favor of graphics as contrasted with symbols in teaching Boolean algebra. Unfortunately, the data are meager on encoding problems; they suggest that symbolism can make a difference in the rate of learning. Since data are almost non-existent for retention and transfer, their implications for these objectives are unknown.

[2]Stolurow, L. M. and McHale, T. (1962-63) under USOE Contract 2-20-003, Title: "Psychological and Educational Factors in Transfer of Training."

Related is the problem of stimulus support in the presentation of mathematics materials for learning. Rigney and Budnoff (1962) used both pure prompting, pure confirmation procedures and combinations of them in teaching Boolean algebra. They found that pure confirmation, the condition with least stimulus support, led to lower error scores in learning than did the mixture of prompting followed by confirmation, the vanishing condition. This was true for both upper and lower intelligence groups. However, the reverse was true for the middle intelligence group. It is generally assumed that prompting which maximizes stimulus support is a desirable initial learning procedure for it raises the probability of the correct response. Once the behavior reaches a level high enough to withstand the withdrawal of stimulus support, then confirmation could be used to minimize stimulus support. It is not clear why the middle group would not respond in the same way as the extremes. These data suggest that some other factor (as yet unknown) was operating to produce the unexpected results from the middle group.

Angell and Lumsdaine (1962) also studied vanishing and found that it resulted in equivalent performance scores to those of a group for whom stimulus support was not withdrawn. However, two weeks later their 5th and 6th graders who were trained with the vanishing procedure achieved higher retention scores. Their results are, therefore, consistent with the theory described above.

One of the significant uninvestigated problems is discrimination. Training with mathematical symbols would suggest the need to differentiate them would arise with many students.

Feedback Characteristics

The requirements for optimum feedback in complex learning situations are poorly understood. The particular events which follow response seem to have several potential dimensions of effect upon the student. The most salient of these is the reinforcement effect, but it is typically confounded with reward, information and motivation effects. If it is assumed that any event following a response can have implications in one or more of these four dimensions, then each is potentially variable independently and may have separate effects on behavior. Teachers and programers differ greatly in the language they use to inform the student of the correctness of his responses; consequently, they could differ in their use of language relating to the reward, information or motivation effects of feedback. Presumably, an effective program would make selective use of feedback to provide each of the four aspects of it as appropriate and important for optimum effects.

Eigen and King (1962) used a program that taught numerals and the concepts of "oneness" to "nineness" to five and six year olds. With some students, they added trinkets to verbal knowledge of results; however, the trinkets produced no differences in student performance. This suggests that the concrete aspect of reward is not too critical even at this early age.

Studies With Architectural Implications

The development of a psychological architecture for educational engineering has two aspects. One is the delineation of associative structures, or cognitive organizations, that pertain to knowledge of the subject matter. The other is the delineation of strategies. Gagne', et al (1962) provide a key to the analysis necessary for the identification of the hierarchical structure of "learning sets". Their key is the question "What would the individual have to know how to do in order to perform this task, after being given only instructions?" By asking this question, each learning set is specified at a level, and by repeating the question, every subordinate level is described down to the simple, the most general and the lowest learning sets. A "hierarchy of knowledge" becomes explicit by this process. However, this is not a sufficient procedure for generating an instructional program, since other objectives also are to be accomplished than those pertaining to knowledge. For instance, "cognitive styles" or strategies of search and selection also are sought. To secure comparable information on the structure of strategies, a different question is asked. It is concerned with the procedures, methods and techniques to be used. Consequently, the question is "What must the learner do in order to perform this task?" We can think of the answers to this question as a set of operations performed on the knowledge requirements identified by the first question.

Associative Structures

The hierarchical structures of knowledge identified by Gagne' and his colleagues (Gagne', 1962; Gagne'& Brown, 1961; Gagne' & Dick, 1961; Gagne' & Dick, 1962; Gagne', Mayor, Garstens & Paradise, 1962; Gagne' & Paradise, 1961) represent associative structures which depend upon positive transfer for their efficient formation. The units of which these structures are built are more molar than those typically studied in the learning laboratory. For example, at level V (Gagne' & Paradise, 1961) symbol recognition is a class of behaviors, not a single stimulus response connection. It is a learning set in that there is a common principle involved in the student's responses to each of the exemplars of the class of stimuli. Gagne' has suggested that the basic level be identified as one that is specified by pure factor tests. These, then, are alternative ways of specifying the elements of molar associative structures.

Fundamental to the performance of a learning set is the more molecular learning involving individual stimuli and responses as in the learning to recognize an individual symbol such as a summation size or an integral, etc. Once the set of symbols relating to an area of mathematics has been learned, then the student is at the basic level in the Gagne' and Paradise hierarchy--level V--"symbol recognition".Implied, but not specified in their analysis, is the prior learning of the molecular structures that make up the class of things labeled a learning set.

Among the many topics studied in research on programing, some provide information relating to conceptual structures. For example, relationships between aptitudes and achievement, size of step, and organization.

Aptitude. The aptitudes of the student are the structures used in building higher levels of knowledge and skill. In fact, any hierarchy can be considered as a strcuture built upon a base identified by aptitude tests. The general ability level of the student as measured by MA or IQ test is typically used in education for selection purposes, but seldom for differential instruction. Aptitude tests measure more specific types of performance relating to the content of a program. From this point of view, it seems reasonable to assume that efficient instruction can compensate for some individual differences in general or specific abilities. Smith (1962) and Cartwright (1962), for example, report data to confirm this.

The organization of a program seems to determine the abilities used by the student in achieving a particular level of performance. For example, with a program teaching the concept of a fraction, general ability accounted for more of the variance when the step sequence was mixed than it did when the same steps were systematically organized. With another program, Dick (1963) found that different abilities accounted for the majority of the variance in achievement test scores, depending upon whether the students worked alone or in pairs. Verbal aptitude was more important than quantitative when students worked in pairs, but the reverse was true when they worked individually. Eigen (1962) found a low order interaction effect between IQ and method of presentation. With a horizontal format and with machine presentation, there was a significant correlation between IQ and achievement, and also between moding level and achievement. However, this was not true for the vertical format.

The theory presented by Gagne' and Paradise (1961) relates to the correlations between relevant abilities and rates of attainment of learning sets. They predicted that this relationship would decrease with progression upwards in the hierarchical set. They found support for this position and also found an increasing correlation of relevant abilities with achievement of learning sets, and for a low, but constant correlation between irrelevant abilities and achievement of learning sets.

If we assume that reading ability is a very basic learning set, then the fact that Feldhusen and Eigen (1963) failed to find it related to achievement on a sets, relations and functions program also fits Gagne's prediction.

Size of step. Step size is ambiguous as a general concept but somewhat more meaningful when considered in relation to a single program consisting of versions with different numbers of steps. Evans, Glaser and Homme (1960) found that a smaller step program was more efficient than a larger step program for teaching the conversion to number bases other than 10. They found the smaller step program resulted in significantly fewer errors and better performance on a delayed retention test. Shay (1961) related IQ and amount learned by fourth graders to step size and found no significant relationship.

Organization. The way in which materials are organized for presentation to the student theoretically could make a significant difference in his ability to learn, retain and transfer the knowledge taught. The data reported to date indicate that a variety of different sequences can

produce equivalent achievement scores (e.g., Cartwright, 1962). There is a question about the comparability of the achievement, however, in terms of the specific kinds of information learned by the students under the different conditions. It is possible to achieve the same scores but on different items.

While several different sequences may produce equivalent mean scores, they may have different implications for the abilities required to do this (e.g., Cartwright, 1962; Smith, 1962). Furthermore, different sequences of frames can have effects that are revealed by retention and transfer scores. Cartwright (1962) found that one sequence was better for retention but that another was better for transfer.

The most penetrating study could come from the use of a program that had been shown to produce positive transfer between learning sets. With such a program, the order of the sets could be reversed to see if it would alter performance on the achievement test as revealed by the measure of transfer used by Gagne' & Paradise (1961).

Strategies

The various strategies used by students are sets of operations performed on classes of stimuli. A single operation such as attending to a stimulus is a basic strategy. Other strategies are based upon simple attention strategies or combine several into a larger behavioral unit.

Search. Attention habits relating to individual cues, or classes of cues, represent one type of search strategy. Search is a general set of operations that can be subdivided into sets. One set consists of those

in which the search is associated with the spatial arrangement of stimulus. The visual fixation habits in reading a printed page, a diagram, table or such mathematical materials as an algebra or geometry page are specific examples of this type of search. Another type is scanning and a third is tracking, etc. Displays that are not static require scanning and tracking skills. These skills serve as strategies for securing information and relate to speed of discovery in the contexts in which they are used.

Integration. Another strategy consists in relating successively presented information such as sounds or separately presented ideas. To cope with these, the student needs to develop integration strategies. One example at a basic level is sound blending. More complex levels of integration occur when isolated, but relatable items of knowledge are presented in prose or when the student is taught a mathematical principle such as associativity or commutativity and then must use each of these in some order to solve a problem or to develop a proof. The deliberate formation of many-to-one sets of associations is an example of a convergent associative structure involving an integration strategy.

Diversity. The opposite strategy is one generating diverse sets of associative connections. Here, the student learns to associate several responses with a single one. This results in the formation of divergent associative structures which seem to be related to originality (e.g., Maltzman, 1960).

Discovery. The term discovery has many different meanings of which three that are fairly general can be identified. One refers to a class of procedures used in teaching as when something is "taught by the discovery method". An example in mathematics is the set of procedures described by Polya (1962). In this use of the term, discovery refers to techniques for promoting discovery on the part of the student (e.g., Henderson, 1962; Hendrix, 1947).

A second usage refers to the experience of the student. In this usage the "a ha" experience is a discovery experience. The strategies used by students to discover a solution or principle are different from the strategies used by teachers to get students to discover.

A third usage refers to the product of a discovery experience. Of the three usages, the first two are of greatest interest.

In programed instruction, we can distinguish among three sets of operations relating to discovery. First, steps can be written so that the student is required to discover which stimulus is to be the occasion for a particular response. The student may know how to add before he learns to add when he sees a capital Sigma. Secondly, steps can be written so that the student must discover which response to make. Whereas the former can be referred to as cue discovery, the latter is response discovery. A third type is mediator discovery as when the student must discover a principle or formula to use in order to generate a correct response. Discovery can be said to occur whenever the student is given incomplete information and must fill in that which is missing.

The missing information defines a requirement for a strategy and it is assumed that different strategies are required to provide the different types of missing information.

If this analysis of discovery learning is related to discovery teaching, then it is apparent that the latter consists of, not one, but rather, a set of operationally distinguishable procedures, each of which is associated with an aspect of a learning set that is to be discovered.

Some preliminary unpublished data reveal that with the UICSM program, approximately the same level of achievement results from programs written to teach cue, response and mediator discovery. Furthermore, Wolfe (1963) found that students taught by a general discovery method the previous year learned equally well when subsequently taught by either the discovery or an expository method. Thus, there seems to be no reason to fear that a mixture of the two methods would interfere provided the students had discovery teaching initially.

Several studies comparing parts of Unit I of the UICSM course as taught by programed text and by a trained teacher resulted in no difference in average performance achieved by students at different ability levels. This suggests that the discovery approach can be programed to teach as effectively as teachers specifically trained to teach by the methods used in the UICSM program. However, it also is true that when the class means were compared, the range of means for those given the programed instruction materials were more homogeneous than the means of the teacher-taught groups (Brown, 1963).

Gagne' and Brown (1961) compared discovery, guided discovery and the Ruleg methods with ninth and tenth graders. They found that all groups learned significantly, but that the greatest amount of learning was produced by the guided discovery and the least amount by Ruleg.

Summary

Research on programed instruction in the teaching of pure and applied mathematics shows rather clearly that effective learning can be produced by this approach. Not only are there more programs in mathematics than in any other area, but also there are more studies using these programs than those from any other single topic. These studies fall into two groups: (a) those simply using a mathematics program to study a general problem of programed self-instruction, and (b) those using programed self-instruction to learn about mathematics teaching. The former were considered in terms of their engineering implications for mathematics teaching. The latter were considered in terms of their architectural implications in the design of associative structures of knowledge and of strategies through the preparation and use of self-instructional programs. The architectural process is conceived as a continuous development that builds upon the existing structures represented by aptitude scores. Procedures for developing new structures have two facets. The ones concerned with cognitive structures are those dealing with the size of step and the organization of content elements. The ones concerned with strategic structures have been less well delineated and studied but include search, integration, diversity, and discovery. The last is a complex strategy of considerable interest to educators. Research on programed instruction indicates that it can be used to teach: (1) as effectively by discovery as trained teachers while also producing more homogeneous achievement levels; and (2) most effectively by "guided discovery" rather than by either less controlled discovery or by the use of rules and examples that eliminate discovery.