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

ED 018 975 EM 004 034

SEQUENCING OF INSTRUCTION IN RELATION TO HIERARCHIES OF COMPETENCE.

BY- BRIGGS, LESLIE J.
AMERICAN INST. FOR RESEARCH IN BEHAVIORAL SCIENCES

REPORT NUMBER AIR-G3-10/67FR PUB DATE OCT 67
CONTRACT OEC-1-7-071070-3771
EDRS PRICE MF-$0.50 HC-$4.32 106P.

DESCRIPTORS- *SEQUENTIAL LEARNING, *SEQUENTIAL PROGRAMS, *PROGRAMED INSTRUCTION, CURRICULUM DESIGN, *LEARNING, LEARNING PROCESSES, VERBAL LEARNING, PROBLEM SOLVING, COURSE ORGANIZATION, *EXPERIMENTAL PROGRAMS, EDUCATIONAL EXPERIMENTS, HIERARCHIES OF COMPETENCE, SCANDURA, GAGNE

RESEARCH ON SEQUENCING OF INSTRUCTION CAN BE DIVIDED INTO NINE TYPES ACCORDING TO DIFFERENT DIMENSIONS ON WHICH THE EXPERIMENTS VARY. SITUATIONS IN WHICH THE LEARNER CONTROLS THE OBJECTIVES ARE DIFFICULT TO EVALUATE EXPLICITLY BECAUSE OF DIFFERING CONTENT LEARNED. AMONG EXPERIMENTER-CONTROLLED LEARNING SITUATIONS, THE WORK OF GAGNE AND HIS ASSOCIATES YIELD THE MOST INFORMATION ON UNITS OF INSTRUCTION BECAUSE THE LEARNING STRUCTURE OF THE TASK IS CAREFULLY ANALYZED AND SEQUENCING IS BASED ON THE ANALYSIS. OTHER TYPES OF EXPERIMENTS INVESTIGATE SCRAMBLED AND LOGICAL ORDERS USING FRAMES RATHER THAN UNITS OF INSTRUCTION, BRANCHING VS. LINEAR PROGRAMS FOR AUTO-INSTRUCTION, AND EXPERIMENTER-PREPARED ADVANCE ORGANIZERS. A NINTH TYPE INCLUDES STUDIES IN WHICH SEQUENCING IS A STRUCTURAL PART RATHER THAN A DESIRED INDEPENDENT VARIABLE IN THE EXPERIMENT. FUTURE RESEARCH SHOULD TREAT LARGER SEGMENTS OF CURRICULUM, SHOULD EXTEND BEYOND THE MATH AND SCIENCE SUBJECTS CURRENTLY TREATED, SHOULD ANALYZE LEARNING OBJECTIVES FOR TYPE AND STRUCTURE OF LEARNING, AND SHOULD CONTROL EFFECTIVENESS OF PROGRAMING AND ENTERING COMPETENCY. A BIBLIOGRAPHY IS INCLUDED. (BB)

A

AIR-G3-10/67-FR

FINAL REPORT
Contract No. OEC-1-7-071070-3771

SEQUENCING OF INSTRUCTION IN RELATION TO HIERARCHIES OF COMPETENCE

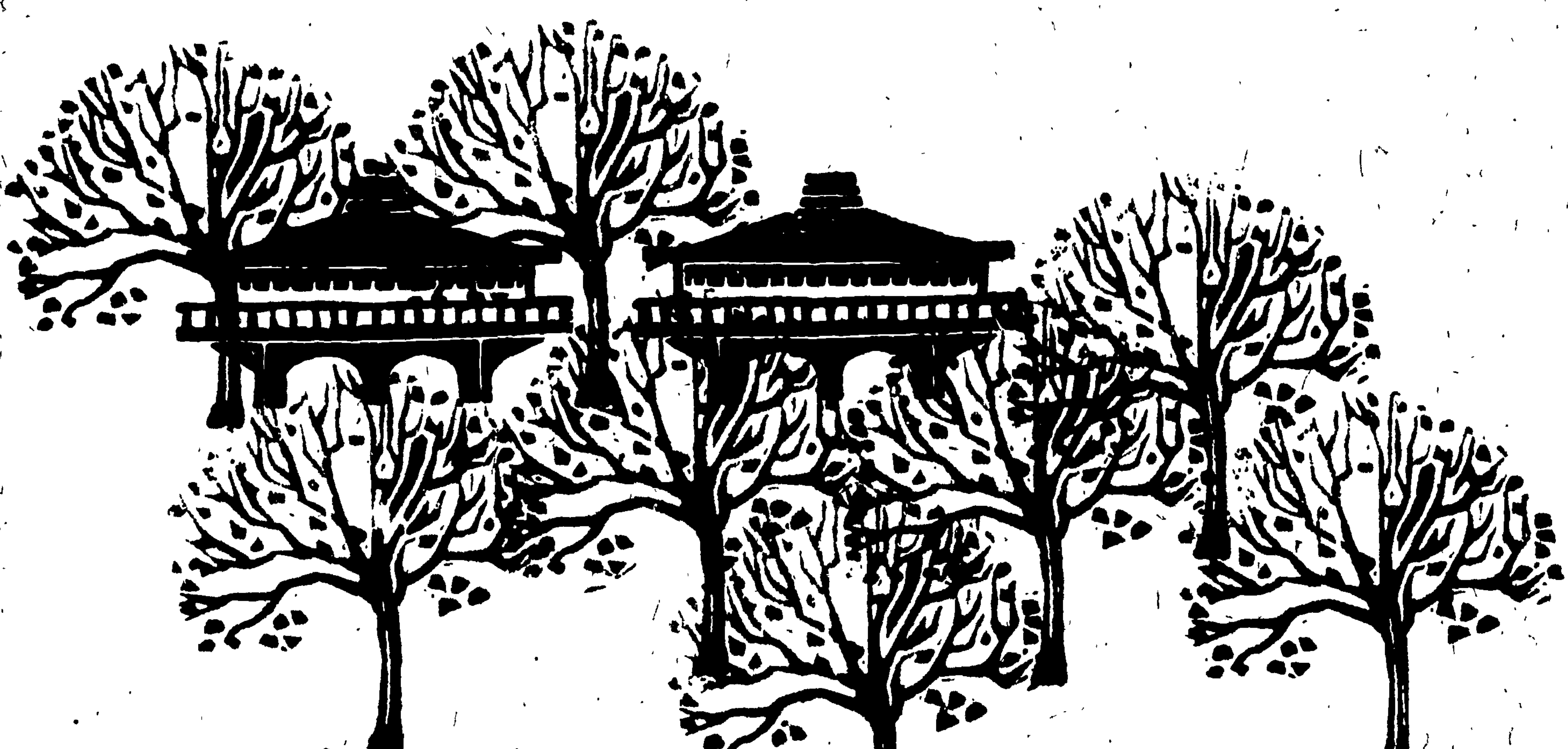
EM004034

Leslie J. Briggs

October 1967

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research



AMERICAN INSTITUTES FOR RESEARCH/PALO ALTO

AMERICAN INSTITUTES FOR RESEARCH

Center for Research and Evaluation
in the Applications of Technology in Education

SEQUENCING OF INSTRUCTION
IN RELATION TO
HIERARCHIES OF COMPETENCE

Contract No. OEC-1-7-071070-3771

Leslie J. Briggs

October 1967

AIR-G3-10/67FR

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

American Institutes for Research
Palo Alto, California

A•I•R ◇ CREATE □ P.O. BOX 1113, PALO ALTO, CALIFORNIA 94302 (LOCATION: 1791 ARASTRADERO ROAD) □ TEL. (415) 328-3550

Table of Contents

Foreword

Persons who were instrumental in identifying the need for a survey of the literature on the topic of sequencing of instruction include Mark A. May and Robert M. Gagne. The Office of Education arranged that such a survey would be made. This report is the result.

The writer expresses his appreciation for the consultative assistance of Robert M. Gagne, who provided many helpful comments and criticisms based on the draft of this report. However the writer is solely responsible for conclusions reached or for errors in fact or interpretation.

Appreciation is also expressed to Mrs. Barbara Rodabaugh, who so capably converted the usual forms of rough draft illegibilities into the present format of the report.

The author prepared this final report while employed by the American Institutes for Research. He is now employed by Dymedia, Incorporated, Palo Alto, California.

While this is the Final Report under this contract, it is considered as a review draft submitted to the Office of Education. Therefore only a few copies have been reproduced in the present form. The reviewer consequently is unable to furnish copies for general distribution. Later, wider distribution is planned by the Office of Education through its own resources, including ERIC. Requests for further information on availability should therefore be addressed to the Office of Education or its resource agencies.

Summary

Since the acquisition of knowledges and skills takes place in a cumulative fashion over periods of time, it is relevant to consider how different portions of the learner's time might best be spent. Experimental investigations directed to this end have dealt with the matter of varying the sequencing of the component parts of the instructional materials and noting the effects upon learning. While practically all investigators believe that some arrangements of material with respect to sequencing the instruction should be more effective than others, the theoretical rationales for this belief vary greatly. The purpose of this report is to review the research literature regarding sequencing of instruction in terms of the rationales employed, the experimental procedures followed, and the results and their apparent implications.

The experiments reviewed in this report are heterogeneous in four important respects. First, some experiments are based on the premise that it is the <u>learner</u> who should determine and control the sequence of events and materials utilized in study directed toward the mastery of the task, while other experiments are based on the premise that the <u>investigator</u> (or programmer of the instruction) is in the best position to exercise this control. Second, some experiments involve year-long or semester-long learning efforts, while others employ tasks requiring only minutes or a few hours of learning time. Third, the kind and degree of learner control or experimenter control varies from control over sequencing of rather gross units of learning effort to more precise, frame-by-frame units of short temporal duration. Fourth, many experiments were designed primarily to investigate sequencing of instruction while others were designed to investigate other independent variables which could be implemented operationally only by introducing variations in the sequencing of the different portions of the learning materials.

Because of this great heterogeneity in these and other details of the experiments reviewed, it was found necessary to attempt to identify categories into which experiments could be classified, so that the results could be discussed in terms of groups of experiments which were somewhat alike. Nine such categories were identified, and this report is organized in accordance with them for the purpose of reaching conclusions on the many aspects of sequencing which the research has involved.

Interestingly enough, some of the most clear-cut findings are reported for experiments which employed learning periods of intermediate length (e.g., several hours of learning rather than less time or a whole semester) and experimenter control at an intermediate level (e.g., the sequencing effects being evaluated were concerned with units of instruction corresponding to sub-tasks on an entire task, rather than with sequencing of single frames in a program or sequences involving weeks of instruction).

The significance of one type of experiment (by Gagne and others) is that when a task can be analyzed into a hierarchical structure,

sequencing the instruction accordingly takes best advantage of transfer of training and so results in the most effective learning. Another type of positive result pertains to what the reviewer would describe as "verbal substance learning," called by Ausubel "reception learning." In these experiments, use of advance organizers, presenting ideas in their most general or abstract form, facilitated learning of the details of the material, presumably by providing ideational scaffolding on basis of which details can be organized or subsumed.

Experiments which yielded "no significant differences," on the other hand, were either those exercising the highest degree of control over sequencing in the sense of control over the smallest elements of instruction ("frames" in programs) or those which permitted the learner much control over sequencing of rather gross portions of instructional materials. However this result of the review of the literature should not lead one to conclude that learning programs cannot be improved by close attention to the minute details of the instruction, on one hand, or to the overall strategy of design of large blocks of instruction on the other. On the contrary, the experiments comparing random vs. logical frame sequence employed such short program segments that the results cannot be generalized to longer learning periods, and the experiments with learner control failed to induce some of the variations in study sequences which were expected in order to test the hypothesis concerning learner control vs. experimenter control.

One difficulty encountered in attempting to draw conclusions from the experiments reviewed is the lack of a taxonomy of learning tasks which is acceptable to all experimenters. While nearly all authors of the research reports reviewed would concede that the importance of sequencing and the specific factors involved in optimal sequencing would vary with the type of task (type of learning), there is not anything approaching universal agreement on how many kinds of learning there are, and furthermore many reports do not classify the specific task in terms of the type of learning the author thinks it represents. While it is true that there is some disagreement on the basic hypothesis that there are several kinds of learning, even those subscribing to the hypothesis have not reached consensus on the number and identity of such learning types. Therefore the descriptive names given to tasks in research reports (e.g., concept learning, problem solving, etc.) do not mean the same thing to different people. It is perhaps partly for this reason that the present reviewer resorted to the kind of categories adopted for the purpose of this review. At the same time, however, the matter of sequencing is so ubiquitous a factor in the arrangement of conditions for a learning task that it reasonably could be expected that difficulty would be encountered in reviewing the data on sequencing. In other words, sequencing is a characteristic of all learning efforts, and it thus cuts across whatever "learning types" might eventually become widely recognized, accepted, and used in classifying tasks.

In this report, suggestions are offered for extensions and improvements of research on the sequencing of instruction. Such research

should eventually result in better understanding of how instructional segments need to be sequenced and how best to apply relevant conditions of learning for each segment. At the present time, while each investigator has some rationale for why sequencing should make a difference for the kind of task involved, the rationales, the specific aspects of sequencing variations employed, and the results are too scattered to offer confident recommendations for operational application. In a general sense, however, an analysis ofthe behavioral objectives for any course of instruction should be the starting point. A study of the extent of apparent interdependence or independence among such objectives should suggest the latitude one might have in blocking out the molar units of instruction as to their place in the course, and an effort to determine the internal structure for each objective in terms of structures which appear to be vertical, flat, or hierarchical in nature should provide a starting point for considering the more molecular structure of the course.

Chapter I: The Problem

Sequencing in Relation to Transfer and Course Structure

The purpose of this report is to summarize the research literature concerning the structure of knowledge as it relates to the optimal sequencing of units of instruction. The educational importance of the problem is as follows: if all the elements of skill and knowledge a student is to acquire during a given course of instruction are independent one from another (that is, if the learning of one element does not facilitate the learning of another), then the different elements could just as well be taught in any arbitrary or random sequence; but if the elements are dependent one upon another (that is, if the learning of one element transfers, thus facilitating the learning of another, then a careful sequencing of elements in terms of the direction of such transfer should be more effective than a random sequence.)

At the outset it may be anticipated that some school learning involves learning of independent elements (for which sequencing of instruction may be done arbitrarily or at random), while other samples of school learning involve dependent elements (for which one sequence would be much better than another). For example, if a student is to memorize the Spanish equivalent of 100 different English words, it may not matter in what order the word pairs are first presented to the student. Whatever the sequence of presentation, it will probably take several trials of practice, using either the entire list or parts of it at a time, with feedback provided, for the student to master the task. In this case, it is not order of presentation which is important; rather, learning will depend upon the amount of practice, the distribution of practice over the whole list or parts of the list, and repeated rehearsal with feedback.

On the other hand, if a student is to learn to solve linear equations, the whole picture is different, both with respect to what is presented and the order in which instruction is sequenced. In the vocabulary example, the entire content of the task to be learned is presented directly to the student, i.e., the entire list of word-pairs is presented in some sequence to him. This kind of learning is called "reception learning" by Ausubel (5) and "reproductive learning" by Gagne and Paradise (36). In the memory example, then, the entire task is presented or "given" to the student, the particular sequencing of elements not being very important. But to learn to solve linear equations, the student is not just presented a number of worked out equations to be mastered in the hope that he will thus learn to solve them. The student is not to memorize these particular equations and their solutions. Rather, he is to first master all the subordinate competencies it takes to be able to solve any equation of this type. What are these required subordinate competencies?

Gagne and Paradise (36) have presented their analysis of the task of learning to solve equations. Their analysis forms a pyramid or hierarchy of sub-skills or subordinate competencies. At the bottom of the diagram are basic aptitudes which students bring to the learning

task, such as symbol recognition. Next to these are the simple, general components of knowledge or competency, such as recognizing that $1 x = x$. Higher up in the pyramid are skills like performing multiplication of numbers in sequence. Continuing progressively upward in the hierarchy are combining fractions with the denominators, simplifying an equation by adding and subtracting numbers to both sides, etc. These subordinate competencies, unlike the example for memorizing, must be taught in a particular sequence (with some options within "layers" of the pyramid) rather than in a random sequence, and they are taught not necessarily by direct presentation of parts of specific equations, but by supplying certain instructional events, materials and exercises which lead to mastery of the subordinate skill. When all these competencies have been mastered through a proper sequence of well designed instruction, all the stimulus needed thereafter for the student to demonstrate mastery of the entire task is the instruction: "Solve the following equation: ____________." This kind of learning is called by Gagne and Paradise (36) "productive" learning to distinguish it from reproductive learning like the example of memorization.

How does one know when a segment of a course is "unstructured," e.g., is composed of independent elements which may be presented in any order during instruction, and when a segment has a hierarchical structure like solving equations, and hence the instruction should be sequenced in accordance with the structure of the task? Are there other kinds of structures in school learning other than "unstructured" and "hierarchical"?

Answers to both the above questions appear to require that somebody performs tasks analyses on a representative variety of school learning tasks to see first what kinds of structures result from the analyses. Unfortunately, as will be seen later, most such reported task analysis efforts have been restricted to educational objectives in mathematics and science. Nobody has yet, to the reviewer's knowledge, deliberately undertaken to analyze tasks from a wide variety of school subject-matter areas, in order to answer the question, "How many types of structure are there to the realms of knowledge of interest to public education?"

As to how the optimal sequencing of instruction is indicated by the analysis of the task, there is a need for experiments which involve teaching one group by a sequence based on the task analysis, another group taught by a sequence which violates all the implications of the analysis (such as by "inverting" the pyramid), and a third group taught by a random sequencing. A few relevant experiments have been reported, but they usually have not involved all three groups recommended above.

There is, then, a need to perform experiments in many subject-matter areas, for each of which one would first analyze the objectives of the course to derive the inferred structure of the learning required. One would next sequence the units of instruction in accordance with the inferred structure. One would next confirm or revise the inferred structure by testing whether or not students taught by the "optimal" sequencing learn better (easier, faster, or with better retention and transfer) than students taught by a random sequence or by a sequence

deliberately designed to run counter to the sequence implied by the inferred structure. In addition to yielding new knowledge on the above two questions, such experiments could suggest that empirically-based revision of learning programs could be further improved by such a procedure.

In following such an experimental procedure to "validate" the "optimum" sequence of instruction for either experimental or program-development purposes, one would have to interpret negative results cautiously. Negative results could be obtained for any of the following reasons:

1. Too large or too small a task was analyzed to yield a structure of practical usefulness for the design of sequences of instruction. Said differently, the resulting subordinate competencies were too large to be meaningful for arranging the learning units, or they were too small (e.g., a single S-R association may be "too analytical" a unit to be a component for sequencing purposes, even though it is a single "instructional event" within a <u>subsequence</u> of instruction designed to establish a competency appearing in the structural analysis).

2. The task analysis and the resulting inferred structure may be faulty. If so, a sequence based on it would be faulty.

3. The sequencing developed as an effort to translate the task structure into a learning program may be faulty. A different sequence derived from the same hierarchy may be better, due to unforeseen "options" in sequencing instruction within a given "layer" or "level" in the pyramid (e.g., there may be unforeseen horizontal transfer effects within a level as well as the vertical transfer expected among levels).

4. The skill in preparing sub-sequences may have been faulty, e.g., a unit in the sequence was inadequately "programed" (did not contain appropriate kinds of instructional events, for the type of learning involved (Gagne (31)). Or, the wrong <u>medium</u> of instruction may have been used. For example, if a good instructional event would be to supply an example for the student, perhaps a word was used when a picture or a real object would present a better stimulus. It may be said parenthetically that multi-media programs may often be optimal (Briggs <u>et al</u>. (15)).

5. The evaluation may have been defective. It would be desirable, in such an experiment, to test for the attainment of each competency in the sequence before the next unit in the instruction is presented. It also is necessary, of course, to test for competency in the total task.

Short of doing such three-group experiments as outlined above to validate both the inferred structure and the sequence derived on basis of the inferred structure, some alternate empirical procedures are possible, some of which were employed in the experiments reviewed in this report. One is to test untrained persons (not formally trained persons) for the subordinate competencies, and then have them try to do the whole task. If a person fails some of the competencies but succeeds on the task, the inferred structure is suspect (but not necessarily disproven,

since some lower level skills "disappear" when consolidated into higher order skills as tested later; see Gagne and Bassler (34)). If a person possesses all the competencies but fails the task, an effort should be made to identify a possible missing element in the analysis, e.g., a layer missing between the highest subordinate layer and the whole task.

A variation of the above procedure could be to test on each competency, then to supply instruction on those failed, then test over the whole task. (This is much like branching programed instruction, in which units passed on basis of prior learning are omitted from instruction but failed units presented.)

A somewhat more laborious procedure (and frustrating for the student) would consist of presenting the whole task repeatedly, and after each failure, present instruction for one competency at a time until the student passes the whole task. This procedure might show that some inferred competencies are not actually required, thus removing "deadwood" from the instruction.

Some of the experiments reviewed in this report involved procedures like some of those discussed above. Other procedures have also been employed, as will be seen later, such as correlational analysis and use of a discrepancy index, noting departures from theoretical predictions.

Distinguishing Among Several Meanings of Structure

Before going on it appears desirable to differentiate among the following, somewhat independent, issues:

1. The manner in which existing knowledge in a subject-matter area or "discipline" was first discovered or acquired need not necessarily have a bearing on the present issue of "structure of a course." It often takes much research, experimentation, trial and error, or fortuitous everyday experience to first discover a new fact, principle, or other element of knowledge. Once discovered and communicated to others, it may be instantly learned, and be believed to be almost self-evident by the receiver of the communication, even though he "never thought of it before, himself." Thus even though much of the present knowledge in a culture was originally perceived only by invention, discovery, or agonizing search, the teaching of that knowledge to the student of today may require only a verbal statement like "warm air rises." This is not to say that a student should never be required to "rediscover" knowledge his teacher already possesses, but it is to say that the student need not invariably acquire the knowledge in the same way it was originally acquired by somebody else.

2. How total knowledge to date is divided up into disciplines is essentially irrelevant to how elements of total knowledge should be taught. The distinction between geology and chemistry is perfectly useful for labelling purposes, and for identifying a professional's field of competence. From the point of view of sequencing of

instruction, segment A of geology might much better follow segment B of chemistry than a traditionally taught segment C of chemistry, from the point of view of transfer of training. Discipline boundaries are hence arbitrary, used better perhaps to classify what a student is studying than to provide effective learning sequences.

3. How the knowledge in a discipline is organized as an outline of the field may be entirely different from the structure of the knowledge for learning purposes. Thus a "logical" outline or "structure" may be entirely useful for the professional to communicate with another about matters which both understand, but useless in guiding a student. It has been reported that the sequence of chapters in a textbook is often not the sequence which would be used in optimum design of instruction.

4. For learning purposes, then, "structure" means none of the above things. It means rather the description of the dependent and independent relationships among component competencies, arranged so as to imply when sequencing can be random or optional and when sequencing must be carefully planned, on the basis that transfer will be optimal in order to build up from simple skills to more complex ones.

A Speculation Regarding Types of Structure

As shown earlier, examples can readily be given for two kinds of course structure: "unstructured" and "hierarchical." For sake of terminology, and to speculate concerning gaps in the evidence from the literature the writer has surveyed, the following classifications are offered. (In later sections of this report distinctions are made among the various sizes of units which may be considered relevant to design of instruction. In practice, one might first make a structure for all the objectives in an entire course, and then make more detailed structures for single objectives or groups of objectives. This matter is obviously dependent on how coarse or fine the objectives are as initially stated. These matters pertain to the planning of overall strategy for the instructional course, and also to the details of learning of the individual competencies.)

1. A flat structure. This phrase is employed to describe the organization or structure of the course when it does not appear to matter in what sequence the instruction for the various major objectives or subobjectives is arranged. For a course with a flat structure one could either conduct the instruction in a random sequence or in any arbitrarily chosen order which one prefers. The significance of such a structure would be that the competencies gained in reaching each objective are independent from the competencies gained in reaching all other objectives. It is conjectured that if a course is well analyzed for the purpose of deciding upon the sequencing of the instruction, a truly flat structure will seldom be encountered. When one does encounter an apparently flat structure of a course, one can question the value of the course, or perhaps the objectives should be restudied to make sure that they are in fact behavioral objectives rather than content objectives (see Briggs

et al. (15)). An example might be a foreign language course involving written vocabulary, but no sentence translation or speaking. Another might be a history course limited to learning dates on which events occurred.

2. A vertical structure. This term is applied when there is one fixed best sequence, in which Objective A should first be taught, then Objective B, and then C, and so on. A course having such a vertical structure thus contains no "lateral transfer" among competencies which would otherwise appear at the same level in a hierarchical structure. It is believed that this type of structure would also be encountered infrequently for a course as a whole, although it may be encountered sometimes in planning the instruction within a single course objective, or a sub-objective. This type of structure then has only one competency per level.

3. The hierarchical structure. This kind of structuring is represented by a pyramid-shaped arrangement of the objectives of the course in which the objective at the top of the pyramid is a global, total course objective, and the subordinate objectives are arranged in layers. A hierarchical structure implies that all of the competencies within a layer should be taught before instruction for the next layer is begun (because vertical transfer is expected), although there may be options in the sequencing of the instruction within a layer (if lateral transfer is not expected). A hierarchical structure is a frequently-reported structure for carefully analyzed learning objectives or tasks (see Chapter IV).

4. Mixed structure. This would be illustrated by a course in which two or more major parts can be taught in random order, but where hierarchies may exist within the parts. This is called "parallel learning" by Ausubel and Youssef (8). A special case might exist as for foreign language, where vocabulary and sentence structure are both involved. Just how vocabulary and sentence structure are time phased would permit at least these options, desirable or not: (a) teach all the vocabulary, then sentence structure; (b) teach some vocabulary and some sentence structure, using familiar words only in the latter; (c) same as b, but introduce new words in the course of instruction on sentence structure. There may be instances of the reverse situation, where sequence is important among major objectives but not in the learning of the individual objective.

5. The special case of a flat structure requiring spiral sequencing of instruction. This type of structure is sometimes encountered when the major objective of a course is that the student learn to solve complex problems by analyzing each of several major components in a problem. An example of this has been encountered in discussions with subject-matter experts on learning to analyze foreign policy problems. Subject-matter experts in this area indicate that in making each foreign policy decision encountered in practice, it is necessary to analyze each of a dozen important factors going into each foreign policy decision in such a way as to appraise each factor individually and also to arrive at the best

trade-off decision as relating to each of the 12 factors. In arranging instruction in this case one might proceed first by teaching some individual concepts and principles needed later in analyzing each of the 12 factors to be considered in making a foreign policy decision; this constitutes the first "spiral." These 12 introductory sequences can be taught in any order. Next, a problem would be presented to the student in which the analysis is given to him for 11 factors, and he solves the problem by analyzing the 12th factor and arriving at a decision based on use of the analyses provided for the other 11 (second spiral). In the next problem given, two factors would essentially be left blank, requiring the student to consider the supplied information on 10 factors in terms of his own analysis for the other two factors. By spiraling the instruction the programmer or experimenter continues to supply an increasingly small part of the total solution, while the learner develops the competency to solve an increasingly large part of the problem. After a dozen such cycles or spirals he has acquired the capability to analyze a new problem by analyzing all 12 of its components, and making the decision for action. While there might conceivably be a better way to approach such instruction, the spiral sequencing at least represents one logical approach for this type of problem. This type of learning structure has been discussed by Glaser (40) and by Bruner (17), who refers to this sequencing procedure as "revisiting," or learning which "turns back on itself." In the "RULEG" system of programing discussed by Glaser, a wider variety of examples or finer discriminations may be accomplished by the later spirals.

More research is needed on learning structures, and how they relate to effective programing styles to establish the identified competencies.

The Pervasiveness of the Sequencing Problem

The problem of sequencing of instruction can be considered with respect to an almost infinite variety of units of instructional time -- decades, years, semesters, hours, minutes, seconds, or microseconds. It is necessary first to show why this is so, and then to delimit the scope of the problem with respect to this report.

Starting with the largest unit of time, the lifetime of the learner, one may recall current discussions concerning the need for lifelong learning in order that the person be equipped to adjust to a changing society in respect to work, leisure time, and new social problems. Much effort would be required to define the objectives for such a long-term learning period before one would be ready to think of the sequencing of the learning directed toward such goals.

Next one could consider sequencing in terms of long blocks of progress in formal education, such as grammar school, high school, or college. While one could conceive of attempting to define reasonable sequences of courses of instruction for different persons for such long periods of time, normally only certain courses are designated as prerequisites for later ones. While curriculum committees do prepare

general statements of "scope and sequence" of instruction covering periods of several years of study, it is easy to understand why formal experiments cannot be cited, as is the case for much briefer learning periods.

In professional training requiring college and graduate school instruction, there have been derived, largely on the basis of tradition and subjective judgments, required sequences of academic courses scheduled in terms of practical experience to follow such as internships, practicum courses, etc. While these arrangements also involve gross blocks of work to be sequenced rather than minute ones, they all imply the common belief that learning for practical occupational purposes involves much cumulative knowledge and skill. These must be sequenced in terms of the transfer of early learning to the performance of greater responsibilities involving increasingly complex judgments to be exercised in the early period of professional work.

While it is conceivable, again, that massive experiments could be conducted to evaluate and revise such gross sequencing practices in education and training, the normal course of events is to revise curriculum requirements based on rather informal judgments of the strengths and weaknesses in the performance of the graduates of such training programs.

Going now to the next smaller unit of instructional time, the conventional "course," say a one-semester one, we encounter the largest unit of time for which experimental results will be reviewed in this report. There exist both experiments and judgmental rationales in respect to the sequencing of instruction of this size of unit. Consider, for example, the textbook for a one-semester course. The author of the text may, in his foreword to teachers, indicate indirectly how he thinks the course is structured. He may ask that the teacher keep the instruction based on the order in which the chapters appear in his book. He may not give an explicit reason for this, and if he does, it may be based more on the "logic of the subject matter" than upon any inferred skill hierarchies. On the other hand the author may make rather well thought out statements of how the student is to progress from simple skills to more complex ones by studying the chapters in the order in which they are presented in the book.

It is, then, in units of a semester or less of learning time that research has been focused most clearly upon problems of course structure and sequencing. The results of the research reviewed here are derived largely from mathematics and science, in which competencies of rather well-defined sorts for a long time have been regarded as the goals of the instruction. This clear focus upon competencies (detailed behavioral objectives) is lacking in other subject-matter areas, notably social science. Apparently not enough efforts have been made to describe structures for the social sciences except in terms of subject-matter content. Since researchers have also avoided this area, there seems to exist no adequate basis for saying what the structure of these areas might be. At least in the writer's search of the literature, only one attempt was found to analyze a social science objective. This was offered by Gagne (32) for the behavior structure involved in voting behavior.

Because of these considerations, generalizations of findings regarding structure and sequence, as reviewed here, strictly speaking are limited mainly to science and mathematics. However the writer hazards the guess that hierarchical structures could be derived for history, if someone would clearly state the objectives for teaching history in behavioral, rather than content, terms. Then perhaps some of the existing history textbooks (content) could be utilized, along with appropriate instructional events, to make history instruction achieve its goals. This awaits a future effort.

In review, we have here glimpsed briefly the problem of sequencing from a life-long learning period to a one-semester course period, arriving at the level of a competency, some limited objective within a course such as may be taught (at the appropriate point in the overall sequence) in a rather small amount of time, say five minutes to one hour. It is, then, brief units of instruction which are to be arranged in a sequence to match the inferred structure of the course or of one objective in the course. It is this size of unit, then, which much of the research literature to be reviewed has dealt with.

However, one group of research studies has pertained to a still smaller unit, with reference to sequencing of instruction. These studies have dealt with frames as the units for the experimental study of various ways of arranging sequences. Frames is a term from linear, programed instruction, referring usually to no more than two sentences of text, with a blank word to be filled in by the student. Some such frames may require only a few seconds of time.

Finally, some researchers have investigated behavioral events within a single frame, such as reading time vs. response or decision-making time, measured as response latency in fractions of seconds, or delay of feedback in seconds or fractions of seconds.

This later kind of study is more minute than the behavioral units meant here by the word competency, but of course these are legitimate research areas in their own right. Thus while the reviewer sought primarily to report results of studies of hierarchies and sequences for tasks taking only a few hours of learning, other investigations at the frame level or the one-semester course level are also reported.

Distinguishing Sequencing of Competencies from Conditions of Learning

Briefly, this issue is the distinction between what is taught and in what order and how it is taught. Conceptually, what, what order, and how taught are obviously very distinct. As made evident by now, what refers to what competencies, not to subject-matter content or particular instructional media, materials, or stimuli. What order refers to the sequence of the units of instruction, each corresponding to a competency. It is the order as implied in the nature of the structural analysis which is the topic of this report.

However, in practice, it is often less easy to distinguish between what is taught and how it is taught. But roughly speaking, the research reviewed here at the competency unit level of detail represents what is taught and in what order. The literature based on frame order, however, gets into the issue of how the competency is taught. So in this sense, the present review covers both kinds of literature, but without the intent to review the dozens of theoretical, empirical and procedural ramifications of how to teach the competencies. To do that would involve review of issues reviewed already elsewhere, covering such topics as prompting vs. discovery, response mode, variety of examples, amount and kind of feedback, direct or indirect guidance, vanishing of prompts, vicarious reinforcement, etc. In some sections of this report, however, interactions of frame sequence with response mode, for example, are encountered.

At the outset it was said that the relative importance of sequencing will be found to be a function of the type of learning task. This was illustrated by the two examples of memorizing word pairs vs. solving equations. There is a need eventually, then, to analyze tasks representing all kinds of learning as well as representing many subject-matter areas, in order to learn about the relationships among course structure, how to teach, and how to sequence. The matter of sequencing of kinds of instructional events and their stimuli defines the areas of how to teach for each kind of learning. Thus the sequencing of stimuli within a competency-establishing unit of instruction becomes a matter in which sequencing is involved in how the competency is taught. Otherwise, we seek to keep sequencing of the units as a matter separate from the internal instructional characteristics of the learning program. The reader familiar with the implications of the several kinds of learning and their associated conditions of learning (Gagne (31)), and with the distinctions among conditions of learning, instructional events, instructional stimuli, media schedules, and media of instruction (Briggs, et al. (15)), will readily understand why this report is limited generally to sequencing of the competencies identified with units of instruction, to the exclusion of the matter of how the unit establishes the competency.

The implication of the above, and the complications avoided by the boundaries set for this review, may be illustrated by an example. Suppose a task has been analyzed, and the order of teaching the competencies decided upon. Suppose each of five competencies in the order to be followed involves a different kind of learning as identified by Gagne (31). The different kinds of learning make no necessary difference in the validity of the sequence to be followed, and it is just as easy to decide the sequence as if all five competencies required the same kind of learning. But if we were actually programing the details of instruction within each unit, we would have to follow five different guidelines as to the conditions of learning to be provided. Since this latter problem has been addressed already by Gagne (31) and by Briggs et al. (15), it is excluded from this report.

However, in some instances, one consideration in deciding upon sequencing of units would be the consideration of the types of learning

making up the hierarchy. Awareness of the full implication of the types of learning would lead one to be aware that, in general, the "higher forms of learning" (principles and problem solving) will tend to be found in the upper layers of the hierarchy, while "lower" forms of learning (associations and discriminations) will fall lower in level in the hierarchy. But exceptions to this trend may be noted if "cycles" of type of learning are involved. For example, to acquire a competency in Level V (near the bottom) in a hierarchy, a student may first have to form associations and multiple discriminations in order to acquire a concept needed for the competency. Then for a Level III competency (higher up in the hierarchy) the student may have to form discriminations among concepts on the way to grasping a principle. When the first draft of a structure suggests options in sequence within a level in the hierarchy, consideration of types of learning involved as well as competency at doing what may lead one to make a revision in the hierarchy to obtain a better sequencing of instruction.

This general tendency for the "higher" forms of learning to be found at the top of the hierarchy means that the more complex skills are built upon a history of acquisition of simpler skills acquired by "lower" forms of learning. It is of passing interest to also note that this going from simple to complex in the history of developing particular complex skills parallels the history of the child's development from conditioned responses, on through discriminations, chains, concepts, principles, etc. The history of establishing a complex skill in an older student, then, appears to add one more component to the well-known maxim "ontogeny recapitulates phylogeny"; the maxim then becomes "skill recapitulates ontogeny which recapitulates phylogeny."

It should be recognized at this point that not all learning theorists recognize the existence of several distinct kinds of learning, as discussed above. It is likely that those who agree that there are several kinds of learning would include Tolman (108), Gagne (31), Hilgard (48), Ausubel (5), and Pressey (83). It is likely that those who would disagree, at least to the extent of the implications as discussed above, would include Thorndike (107), Hull (51), Skinner (101), and Guthrie (43).

Since the major emphasis in this report is upon sequencing of series of competencies, rather than how to teach each competency, this divergence in views as to number of types of learning (and what the types are called) will be left at this point.

The next section of this chapter takes up an issue more directly germane to this report, namely alternate views as to how the importance of sequencing is regarded by various investigators.

Alternate Ways of Conceptualizing the Role of Sequencing

Up to this point in this introductory chapter, the reviewer intentionally has stated the problem of sequencing of instruction in the

light of a particular conception of learning which is compatible with the concepts of "course structure," reviewed earlier, and with the placing of heavy emphasis upon transfer of training as the underlying reason for being concerned with sequencing. This approach has been adopted up to this point in order to make the exposition of the problem as consistent and clear as possible (in terms of this way of stating the problem). A second reason for this approach to the introduction admittedly is that the reviewer himself finds it possible to conceptualize the problem in these terms. A third reason is that many of the experiments to be reviewed were designed by investigators who hold a similar view of the problem, possibly indicating that most researchers are able to think in these same terms about the problem. Therefore it is hoped that introducing the problem within this conceptual framework will enable still other researchers to utilize this report in planning further investigations of the problem.

Having said this, it now appears appropriate to recognize that there are alternate conceptions of the role of sequence of instruction. These differences arise in part because different investigators have different views as to who should control the sequencing, the learner or the investigator. But these differences also arise because learning theories differ, and because many researchers find learning theory not entirely adequate for grappling with problems in instructional design, such as sequencing.

While the diverse ways of conceptualizing the sequence problem will next be reviewed here briefly, it is important to keep in sight the fact that <u>almost all theoretical positions contain some reason for placing importance upon the sequencing of instruction</u>.

First, of all, four researchers have found it necessary to look outside the bounds of learning theory to find adequate conceptualizations for some problems in instructional design. These four views will next be presented, followed by other views not intended here to be labeled as either theoretical or atheoretical; these other views are given in an attempt to sample the range of investigators' thinking about problems related to instructional sequencing.

<u>Gagne</u>

The description of the problem of instructional sequence as presented in this introductory chapter up to this point follows most closely the viewpoints of Gagne. He has indicated that if he were faced with the problem of improving training he would "not look for much help from the well-known learning principles like reinforcement, distribution of practice, response familiarity" and so on, but rather to "the technique of task analysis, and at the principles of component task achievement, intra-task transfer, and the sequencing of learning to find those ideas of greatest usefulness in the design of effective training..." (33, p.181).

Gagne (32) has written on the topic of curriculum research, handling the topic in such a way as to integrate his previous work on the

conditions of learning (31) and his treatment of the topic of hierarchies of competence (29). Gagne stresses the importance that a theory of instruction focus upon the capabilities of the student, not the behavior of the teacher or the words in a book. He indicates that the purpose of stating objectives is not to pick content but to develop the design of learning in terms of the teacher-pupil communication or the characteristics of a learning program which later are evaluated. The capability to add fractions is an educational objective in the fulfillment of which one may select textbooks, motion pictures, laboratory equipment, and even teachers, but one does not select content. It is derived from the objectives. Gagne further defines a unit of content in the curriculum as a single student capability which is acquired under a single set of learning conditions, assuming that the prerequisite capabilities have been learned.

Gagne defines curriculum as a sequence of units arranged in such a way that learning of each unit may be accomplished as a single act, providing that the capabilities described by specific prior units have been developed. A description of a curriculum consists of (a) a statement of the terminal objectives, (b) the sequence of prerequisite capabilities, and (c)the new capabilities to be built upon prior capabilities. His method for deriving subordinate capabilities is discussed in a later chapter of this report.

Bruner

Bruner has emphasized the need for a theory of instruction which would be a better guide to educational practice than is learning theory. He further states that, as faulty as learning theory is as a guide to practice, it is misapplication and overgeneralization of theory which may bring even more serious consequences to education. Bruner says (17, p.524):[1]

> "...When I say a theory of instruction is prescriptive, I mean it is <u>before the fact</u>. It is before learning has taken place and not while and after learning has taken place. Let me give you an example of the kind of difficulty you get into when you assume that you can use the slender reed of learning theory to lean on. Take, for example, the case of programed instruction.
>
> "There is in the current doctrine (I will call it) of programed instruction the idea that somehow you should take small steps, that each increment should be a small step. Now, this idea is derived willy-nilly from a theory of learning which states that learning is incremental and goes in small steps. Nowhere in the evidence upon which such a theory is based--and it is only partial evidence--nowhere is there anything that says that simply because learning takes place in small steps, the <u>environment</u> should be arranged in small steps.

---

[1] Quoted with the permission of Dr. Bruner.

And so we set up a curriculum that also has small steps. In doing so we fail to take sight of the fact that, indeed, organisms from vertebrate on up through the highest primate, man, operate by taking large packets of information and breaking these down into their own bite size and that unless they have the opportunity to do that, learning may become stereotyped. At least it is a worthy hypothesis about instruction.

"A theory of instruction must concern itself with the relationship between how things are presented and how they are learned. Though I myself have worked hard and long in the vineyard of learning theory, I can do no better than to start by warning the reader away from it. Learning theory is not a theory of instruction. It describes what happened. A theory of instruction is a guide to what to do in order to achieve certain objectives. Unfortunately, we shall have to start pretty nearly at the beginning, for there is very little literature to guide us in this subtle enterprise."

In the same article Bruner goes on to outline four aspects which he believes are essential to the needed theory of instruction. These are: (a) factors which predispose a child to learn effectively, (b) developing an optimum structure of knowledge, (c) developing optimum instructional sequences, and (d) clarifying the nature and placing of rewards and punishments.

Concerning optimum sequencing of instruction, Bruner indicates, in agreement with other writers, that one must be clear about the type of lea-ning before developing the sequence. He lists six things to be provided for in designing sequences of instruction. These are: (a) arrange it that the student grasps the structure by induction from particular instances; (b) give practice in transfer when transfer is expected as a result of learning; (c) use contrast in the sequence;(d) avoid premature symbolization; provide for images first; (e) give practice in both leaping and plodding; small steps are sometimes necessary, but without great leaps involving guessing a child is deprived of his rights as a mind; (f) provide for revisiting--through use of spiral programs so as to not expect that the full value of a matter being studied is grasped always in a single block like a linear sequence.

Scandura

Scandura (94) points out that all teaching based on learning theory makes these assumptions: (a) that the principles useful for explaining laboratory learning are equally critical in the classroom; (b) that to make theory more adequate for the classroom requires discovery of more laws like the ones already discovered in the laboratory; e.g., the theorist may admit need for embellishments of existing theory, but he won't give it up; (c) what is needed is a new technology to apply present theory more effectively in the classroom.

Scandura suggests further that the needed theory of instruction

would involve the following assumptions: (a) some principles of learning theory are of only incidental or secondary importance in real-life learning; (b) the entering competencies a child has may be more important than such principles as reinforcement; (c) much of the learning tasks in real life have no counterparts in theory research; (d) task analyses, assessments of knowledge, and sequencing matters are more like the variables in a needed theory of instruction.

Scandura has also proposed a new language to replace stimulus-response language as a way of communicating about learning. In his set theory language, principles, not S-R connections, are the basic units of learning. The experimental uses and derivations of this language are reviewed in Chapter V.

Pressey

Pressey has long expressed dissatisfaction with learning theory for the purpose of planning instruction. Most recently his expressions of this viewpoint have arisen in the context of programed instruction, which he has much criticized (82, 83). His own preference is for methods he derived years ago (79, 80, 81), now known as adjunct autoinstruction, widely recognized historically as an early form of autoinstruction (62).

While Pressey has not characterized adjunct autoinstruction in terms of learning theory, his elucidation of the method draws upon general concepts not dissimilar to some concepts in gestalt psychology and some in cognitive theory. His method places considerable responsibility upon the learner to range about over the instructional material to create cognitive structure on his own, and to correct his misconceptions and strengthen and broaden his correct perceptions through the provision of feedback following responses to "practice test" items. By implication Pressey thinks that the student can find his own way, sequentially speaking, among the materials in a textbook. Further details on his views and methods are found in Chapter III.

Skinner

The views of Skinner concerning programed instruction have become so well known even by laymen that it seems unnecessary to review them here, because of the great amount of discussion in common news media. Suffice it for the present purpose to say that he, also, places great importance upon the matter of sequencing of instruction (102). It may even appear that he emphasizes sequencing more heavily than do other investigators because it becomes a matter for attention at the "frame" level of the program as well as in the overall behavioral analysis which sets the strategy for the sequencing of the frames in the program. Perhaps the greatest distinction between his specific rationale and other rationales presented here is that he does not rely heavily upon the concept of transfer, but rather upon reinforcement of sequences of responses to certain sequences of defined stimuli. Some reviewers, in fact, have referred to his programing concepts as "response oriented" and to alternate views as "stimulus oriented." Another distinction is his tendency

to apply widely the concept of shaping of behavior where others would speak of varying sets of conditions for different kinds of learning. One section of Chapter IV deals with research in sequencing at the frame level in programs.

Ausubel

This writer is generally thought of as a cognitive theorist, in contrast to a behavioral theorist. He has turned his attention particularly to learning from prose material or lectures, a form of learning he calls "reception learning" (5). Rather than speaking of analyzing a course into hierarchies of competence, as Gagne does, he speaks of achieving stable cognitive organization of ideas by use of advance organizers, general statements in the most abstract form of important ideas, introduced before the details of the instruction. His work is discussed further in Chapter IV.

Mager

Mager has investigated the learner's ability to sequence the instruction for himself, once an objective has been set. In one experiment the learner not only sequenced the instructional content, but he selected the content by the nature of the questions he asked (63). In a less extreme form of learner control over selection and sequencing of instruction, more guidance concerning objectives and learning resources was given the learner, but his sequence of actions was still quite independent (64), in contrast to the little control given by the learner in pre-arranged sequences, such as programed instruction. Mager has continued to participate in instructional programs which attempt to motivate the learner, to provide objectives, and to attempt to establish a desire for more learning on the part of the learner by use of the responsibility he takes for the learning.

Campbell

Campbell (20) conducted an investigation of learner-controlled sequences through the learner's freedom to select from those materials made available. He had previously found that by-passing of material on basis of a response to a test question did not appear to be either better or worse than permitting the learner to judge when to by-pass (19); nor did either basis for by-passing differ in effectiveness from a linear program. One difficulty in both experiments was that students tended to take and use the material as they found it, thus not exercising much the by-pass or self-direction freedom permitted. Still, since both conditions were as effective as a linear program, Campbell tends to favor the more learner-determined procedures, partly because of economy and partly in the belief that some objectives unanticipated by the experimenter might be achieved by some learners having greater freedom for self direction.

Summary of the Chapter

The purpose of this report is to review research in sequencing of instruction in terms of assumed course structures.

Views of learning theorists and experimenters vary widely as to the most important factors in promoting learning. However, for the kind of learning encountered most often in school, almost all agree that sequencing of the instructional units is an important matter in the design of instruction.

However, views differ widely as to why sequencing is important, and also as to how to provide effective sequences of instruction.

In order to relate sequencing systematically to other matters brought up in experiments and in design of instruction, the reviewer has introduced the problem from a particular point of view which is meaningful to him. This was done to make it possible to discuss related matters, including structure of learning tasks requiring various amounts of time for learning, kinds of structure which might be found by task analyses and their implications for sequencing, and learning conditions which may be affected when sequencing is varied systematically.

A brief review was attempted of theoretical positions which co-vary with rationales about sequencing. It was indicated that the experiments to be reviewed differ as to (a) who controls the sequencing, (b) the length of learning periods in the experiments, (c) degree and type of control or variation of sequencing, and (d) interactions of sequence with other characteristics of the learning programs.

Chapter II: Classification of Studies to be Reviewed

Descriptive Dimensions on Which Experiments Differ

It should be apparent from Chapter I that the researchers who have investigated the sequencing of instruction to provide competencies in the order needed to build up more general skills are studying how instruction may be pre-planned and pre-programed in ways designed to meet the educational objectives. In such studies it is assumed that it is the experimenter, acting as a course designer, and not the learner, who does the pre-planning. Just how much pre-planned control a course designer should exert over instruction is a matter on which opinions differ, as shown in the closing section of Chapter I. One basic difference, then, among the experiments reviewed, is the distinction between instruction pre-planned by the experimenter vs. learner-directed procedures.

This distinction, however, is not a simple all-or-none distinction. That is to say, there are various aspects of planning and control over learning sequences and procedures which can be assigned either to the experimenter or to the learner, thus making it necessary to form a profile or pattern of aspects of control over learning in order adequately to describe the specific control involved in a given experiment. Under the basic distinction between learner-directed instruction and pre-planned control by the experimenter, we may list now the specific matters to be controlled by one person or the other.

1. Experiments reviewed involve different amounts of learning time, from less than an hour to a full school semester.

2. Experiments reviewed vary as to who chooses the objectives of the learning, and how much specificity is involved in the identification of the objective. This ranges from "learn something about electronics" to specific responses in frames of programed instruction or specific subordinate competencies in a single task. Sometimes objectives are only stated in terms of items on an end-of-course examination.

3. Experiments reviewed vary as to how well the learning structure of the experimental task is described. Experiments by Gagne and his associates contain the most explicit and unambiguous identification both of the subordinate skills involved in learning the total task, and how these skills relate to each other in a learning hierarchy, thus showing why the adopted sequence is an appropriate test of the inferred structure described in the reports of the experiments. It is because of this, and because of the experimental designs employed, that the research by Gagne and others represents the most significant experimental evaluations available of the relationship between task structure and effective sequencing of instruction. Practically no other experiments reviewed defined the task structure well enough to demonstrate that the learning programs used in the experiments were designed appropriately as a way to validate the inferred structure. This is not to say that other experiments did not succeed in evaluating the effectiveness of the sequencing

in the learning programs in terms of the criteria. But it is to say that these other experimental reports do not show as explicitly as does Gagne the causal relationship between the structure of the task as analyzed and presented in the reports and the effectiveness of the instructional program sequence. Thus many experiments reported here demonstrate various degrees of effective sequencing, but not many show that the sequencing is based directly upon an analysis of the task. This is not a criticism of experiments which demonstrate a degree of effectiveness in sequencing of instruction in terms of the underlying rationale offered as a guide to the determination of sequencing. But this is a statement that since the reviewer sought to find experiments which relate the rationale on which sequence was planned to the problem of the nature of the structure of the learning task, Gagne's research stands practically alone in quality of research design, given the objective the reviewer sought to reach. Needless to say, experiments other than Gagne's have the compensating advantage of providing data on closely related issues which Gagne could not touch upon, given his objective, like that of the reviewer, to study the relation of task structure to sequencing of the learning program. It is this wide scope of the relevant experiments which obliged the reviewer to devise the classification of experiments reported in this chapter.

4. Experiments reviewed vary in the detail in which the learning is sequenced. For example, chapters in a book appear in a specific sequence, but this is no guarantee that the student reads the material in the order in which it is printed. In the experiment by Campbell (20), materials were prepared in "clusters" consisting of test items, maps, reading material, objectives, and film strips. The learner could use these materials, or only part of them, in any sequence he wished. In contrast, a linear program, especially if presented by a teaching machine, forces the learner, frame by frame and sentence by sentence, to read and respond in the sequence the programmer planned.

5. Finally, in the experiments reviewed, the instruction offered on the competencies which were the detailed objectives of the experimental lessons was sometimes programed in the sense meant by linear programmers, frame by frame, and other times was presented as blocks of reading material to be ranged over in whatever sequence the learner's study habits dictated.

Because of the above diversity in experimental procedures in the research reviewed in this report, it was determined that categories should be described into which experiments could be sorted, in order that the discussion of results could be based on somewhat similar experiments. In some cases, only one or two experiments are cited for a category, and in other cases many are cited. In one case, Category "Type I," no experiments are cited, but the general type of learning procedure is described for the sake of rounding out this account of the ubiquitous matter of the sequencing of instruction.

A brief identification of each type of experimental procedure follows.

The Types of Experiments Reviewed

Type I: maximum learner control. Under this category fall various educational arrangements under which the learner is given much responsibility and control both over what is to be learned and how it is to be learned. One example is earning a degree under an honors program in which the learning content is determined by the student, and the content does not even coincide with formal course outlines of objectives. The learner is also free to make visits, work in the library or laboratory, or employ any other learning resources to meet the objectives he has agreed upon with the instructor who serves as his advisor. Other educational programs may carry the label "independent study program," "adult education," "independent research," or "accelerate seminars." Class attendance is not normally required, and the student proceeds entirely on his individual initiative in selecting objectives, materials, study methods, and sequences of activity.

Type II: learner-controlled content and sequencing. Under this category fall many non-educational procedures such as adult self-improvement programs or industrial training programs of an informal nature. An instructor is present but not to sequence the instruction. Rather the instructor responds to requests from the learner, thus becoming an additional source of information for the learner. Either the instructor or the learner may choose the content, but the sequencing is learner-controlled. (Example: experiments by Mager.)

Type III: learner selection of materials and procedures. Under this category the experimenter or instructor normally sets the objectives for the learning and provides learning materials and resources. The learning materials are in well-defined separate packages representing resources the learner may use. The learner is not encouraged to search for his own material, but to utilize the material made available in any sequence he wishes. (Example: experiment by Campbell.)

Type IV: adjunct autoinstruction; mixed experimenter and learner control. The course designer, the instructor, and the textbook writer determine the materials available to the learner. The experimenter provides multiple-choice questions to be used during study by the learner. Although the printed instructional information may be in book or other form, the student studies assigned chapters or segments in the sequence determined by his study habits. The multiple-choice self-test items are regarded not as examinations or criteria but as ways to induce the student to respond, to receive feedback, and to undertake remedial study on objectives in which he is weak. (Example: experiments by Pressey.)

Type V: experimenter-determined sequencing of instruction in accordance with hierarchies of competence. The experimenter first analyzes the task and presents a structure showing how the subordinate competencies relate to each other. The experimenter then selects or designs a learning program following a sequence implied by the task structure. The experimenter then evaluates criterion data in terms of the hypothesis underlying the analysis of the task and the sequencing of

instruction. (Examples: experiments by Gagne and his associates.)

Type VI: experimenter-determined sequencing of frames in programed instruction. Linear program frames are presented to the learner, either in the sequence thought best as determined by the programmer or in some other sequence such as random sequence determined by the experimenter. Interaction of sequence with other characteristics of the program may or may not be involved in the experiment. (Example: experiment by Hamilton.)

Type VII: learner-determined branching in autoinstruction. The individual learner follows either the original linear programed sequence or departs from it to receive remedial instruction when he is doing poorly or to by-pass some of the regular program when he demonstrates prior mastery of relevant competencies. The branching may be determined either by the experimenter when a student passes or fails a test item during the sequence, or by the learner's subjective evaluation of how well he has understood a sequence. A recently employed alternate basis for branching is computer-determined branching in which departures from the linear program may be based on a whole series of responses rather than a response to a single test item. (Example: experiments by Coulson and Silberman.)

Type VIII: experimenter-prepared advance organizers. Pre-planned written statements are prepared by the experimenter which consist of statements of important concepts in the learning program in their most general or abstract form. These statements are administered to the student prior to the detailed learning program. The theory is that these statements, called "advance organizers," help the student develop a cognitive structure under which the detailed material later to be read becomes better learned and retained, because the organizer creates a cognitive framework providing the capability by the learner to establish subsumption or ideational anchorage. (Example: experiments by Ausubel.)

Type IX: sequences pre-planned by the experimenter to test hypotheses about effective characteristics of learning programs. In these experiments the experimenter's main interest was in analyzing the effectiveness of some characteristic of the learning program such as the requirement for overt responding, the use of a wide or narrow variety of examples of a concept, practice in a wide variety of problem-solving instances, etc. In order to study the desired independent variable in the program, the experimenter rearranges sequence to set up the desired experimental conditions. His interest in sequence may thus be secondary to his interest in the associated characteristics of the frames as they are sequenced. (Examples: all experiments in Chapter V.)

While the above classification cannot be defended on any single theoretical basis, it did seem necessary to devise these types in order both to show the wide range of experiments in which sequencing is a factor and to discuss experiments which are somewhat alike. The reader may wish, for each experiment described, to refer to the five numbered points listed in the preceding section of this chapter in order to note

for himself how the experiment may be classified on these five descriptive dimensions on which experiments differ. At the same time, the reader may wish to continuously evaluate the data presented in terms of the basic dimension on which the experiments vary, namely learner-determined sequencing vs. experimenter-determined sequencing.

It may be noted further that experiments described in the first seven types listed above for the most part involve use of actual instructional material suitable either for an educational course or an industrial training program. Many experiments involve instruction in only a limited portion of such a formal course, while other experiments conducted in the classroom consisted of the materials for the entire course. In some instances the experimental procedure was implemented in the classroom and in some instances in a laboratory situation.

The remainder of this report is organized in the following manner: experimental findings for Types I through IV are discussed in Chapter III, which thus reports experiments which gave much control of the learning to the learner. In Chapter V experiments are summarized in which the experimenter controlled the learning sequence by pre-planning efforts consisting either of an autoinstructional program or some other material sequenced in greater or lesser detail as the experimenter desired. In Chapter V, by contrast, experiments are described which were conducted primarily to test some variable in learning other than sequence, but in which data were produced which are relevant to the purpose of this review.

Chapter III: Experiments Involving Learner-directed Instruction

Type I: Maximum Learner Control

Independent study plans of various kinds have been used in education to permit the student maximum choice of course content and maximum self-determination of how he acquires his competencies. Much leeway is given, under the guidance of an advisor-instructor, in the objectives a student may choose to pursue, the materials and resources he will use, and the independent study methods to be adopted.

In degree programs, a student may either be preparing for an "honors examination" in some broadly defined area of liberal arts or other study, or he may be simply studying independently as a means of preparing for examinations in particular courses, while attending normal classes to meet other course requirements as a part of the total degree requirement.

At Ohio State University following World War II, many different educational procedures were used to permit veterans to complete their degrees by the most expeditious legitimate means possible. These provisions included course credit by examination, independent study credit, "accelerate seminars," four quarter per year attendance, and other variations. These are summarized by S. L. Pressey, Educational Acceleration: Appraisals and Basic Problems, The Ohio State University, Bureau of Educational Research, 1949.

The adult learner, working on his own program of self-education, represents the widest latitude in self-determination. The "honors program" represents a wide latitude of self-determination in degree programs.

While almost no responsible person advocates a similar degree of self-determination for the education of young children, yet the "core" requirements in some ungraded plans and other individualized study plans do permit much more than the usual variation in the amount of time a student spends on each core requirement. The nature of the "requirements" however may be as externally-determined as in conventional schools, the student's freedom being more in the nature of how fast he progresses and in what enrichment materials he studies and how he learns what he does learn.

Usually, for the younger learner pursuing individualized study, a definite series of competencies must be mapped out, and the student's progress appraised frequently in terms of defined units of instruction. If the materials employed are conventional rather than "programed," the details of how the student learns from the material may be quite a personal matter, never recorded. But in general, a scope and sequence is defined for the student, and evaluations of his progress may actually be more formalized in accordance with competency units than are appraisals in conventional group instruction. Such "individualized" programs may provide greater latitude for "enrichment" exercises--in fact one reason for assessing progress on "required" skills is to enable the

student to complete them rapidly and spend more time in study of matters which correspond with his personal goals and interests.

It is evident that there is by no means great homogeneity among various "independent study" or "individualized" programs, even at the same age level, with respect to just what aspects of the program are under the student's control to a greater extent than with conventional group instruction.

Nevertheless it is recognized that programs do differ greatly, at a given age level, in the degree to which the student chooses his objectives, the degree to which he has the sequence of competencies spelled out for him, and the degree to which he studies "unprogramed" (conventional) materials in contrast to programed instruction, teaching machines, computer-assisted instruction, etc.

While "independent study programs" thus differ among themselves as to what the student decides and what the planner-teacher-advisor decides, taken collectively such programs may be assumed generally to allow the student to make more of the instructional decisions than is the case for most of the other types of experiments reported in this chapter.

With special reference to sequencing of units of instruction, some of these programs may be even more pre-planned by people other than the learner, while others may indeed involve more permissiveness in allowing the student to choose his goals and his path to them. If programed instruction is used for some or all of the competencies, it is the rate of progress that varies, and otherwise it is like a Type VI program which is pre-planned down to the frame level.

This discussion of Type I is included for rounding out the discussion; no experiments are to be reported nor are evaluations of independent study programs to be reviewed here, except for the reference made above to the monograph by Pressey. While there is, indeed, a research literature reporting evaluations of such programs, the reviewer decided not to attempt to summarize those reports here. Such evaluations often necessarily are based on percentages of students successfully completing such programs, and other statistical data not based on criterion examinations. Such criteria, as relevant as they are for the purpose, are unlike criteria employed in the experiments reviewed in this report.

Type II: Learner-controlled Content and Sequencing

Mager (63) has reported a study in which the sequencing of information asked for by learners was found to be different from the sequence of topics in standard textbooks and course outlines.

Six persons, who thought they "knew nothing about electronics" but who expressed an interest in learning something about the topic, were subjects for an experiment. Each subject participated in a series of private sessions with the instructor. In each session the learner had

student to complete them rapidly and spend more time in study of matters which correspond with his personal goals and interests.

It is evident that there is by no means great homogeneity among various "independent study" or "individualized" programs, even at the same age level, with respect to just what aspects of the program are under the student's control to a greater extent than with conventional group instruction.

Nevertheless it is recognized that programs do differ greatly, at a given age level, in the degree to which the student chooses his objectives, the degree to which he has the sequence of competencies spelled out for him, and the degree to which he studies "unprogramed" (conventional) materials in contrast to programed instruction, teaching machines, computer-assisted instruction, etc.

While "independent study programs" thus differ among themselves as to <u>what</u> the student decides and what the planner-teacher-advisor decides, taken collectively such programs may be assumed generally to allow the student to make more of the instructional decisions than is the case for most of the other types of experiments reported in this chapter.

With special reference to sequencing of units of instruction, some of these programs may be even more pre-planned by people other than the learner, while others may indeed involve more permissiveness in allowing the student to choose his goals and his path to them. If programed instruction is used for some or all of the competencies, it is the rate of progress that varies, and otherwise it is like a Type VI program which is pre-planned down to the frame level.

This discussion of Type I is included for rounding out the discussion; no experiments are to be reported nor are evaluations of independent study programs to be reviewed here, except for the reference made above to the monograph by Pressey. While there is, indeed, a research literature reporting evaluations of such programs, the reviewer decided not to attempt to summarize those reports here. Such evaluations often necessarily are based on percentages of students successfully completing such programs, and other statistical data not based on criterion examinations. Such criteria, as relevant as they are for the purpose, are unlike criteria employed in the experiments reviewed in this report.

<u>Type II: Learner-controlled Content and Sequencing</u>

Mager (63) has reported a study in which the sequencing of information asked for by learners was found to be different from the sequence of topics in standard textbooks and course outlines.

Six persons, who thought they "knew nothing about electronics" but who expressed an interest in learning something about the topic, were subjects for an experiment. Each subject participated in a series of private sessions with the instructor. In each session the learner had

to ask questions of the instructor, as no lecturing or reading were involved. Tape recordings of the sessions revealed the following facts: (a) subjects differed in amount of knowledge already possessed on the topic, as revealed by their questions; (b) subjects wanted to know about concrete objects before they wanted to know about electronic theory, i.e., they wanted to progress from simple wholes to more complex wholes in understanding, not from abstract to concrete, and not from elements to wholes; (c) there was a considerable similarity among the six persons in the nature of the sequences of information called for, even though no specific objectives had been stated other than "desire to learn something about electronics."

Since this experiment was designed to determine whether the sequence of information the student asked for was different from sequences planned by course designers for formal courses, no criterion measures were administered. Thus it was not intended to decide between the relative effectiveness of learner-determined sequences and sequences determined by instructors or course planners.

However, Mager suggests that the usual tryout of draft instructional materials as a method of program improvement requires the student to help improve the details of sequencing matters already decided in more gross terms by the programmer. He suggested that learner-determined sequences might be more motivating to learners, since learners attempt to dovetail new material with what they already know. Thus finding out what typical learners already know, coupled with a sequencing of new information in the order the learner wants it, he argues, would constitute a good method for developing instructional materials.

In discussing his findings, Mager failed to mention that were specific performance objectives and standards set as the goals of the instruction, students might have asked for information in different sequences than they did in this experiment. Thus the procedures he used might be more effective for keeping the student interested in an optional learning program than for establishing specified competencies. Nevertheless the value of the experiment might lie in suggesting that if the learner had been given specific competencies as objectives, the sequence of information he asked for might still be different from that which a programmer might design. Since he did not wish to introduce "tests" into the experiment, however, the relevance of the finding for the necessarily test-oriented nature of education in general remains somewhat uncertain. One might say that one of the objectives of his experiment was to see how long a student would stay with the topic under the teaching procedure employed--a not unworthy objective for public education, but one needing to be harmonized with stated objectives and evaluation of their attainment.

In a follow-up study, Mager and Clark (63) did give adult trainees a detailed list of instructional objectives and some suggestions on learning resources, with freedom to follow their own schedule and procedures for learning. The learning time taken was reduced by an average of 65 percent from the time required for the formal course, but no

second control group was used to show how the formal course might have been redesigned for improvement in efficiency and effectiveness. Nevertheless the implication is that adult trainees, if given well-stated objectives and availability to sources of relevant information, might well achieve the objectives more efficiently than by use of existing types of conventional classroom training.

Both these experiments have suggestive value for public education, but they would need to be repeated with added elements to the experimental design in order to have more definitive value as guidelines for educational practice.

Type III: Learner Selection of Materials and Procedures

Campbell (20) reported a series of experiments similar to those described above by Mager and Mager and Clark, except that materials were packaged in such a way as to give students freedom in the specifics of what materials were used and the sequence of use. Briefly, Campbell intended to compare the effectiveness of programed instruction (a programmer-determined sequence of study) with student self-direction. The theoretical basis on which Campbell hypothesized student self-direction to be superior was also similar to Mager's discussion. Two factors were stressed: (a) meaningfulness of materials <u>to the learner</u>, and (b) motivation. Campbell believed that when problem-solving activities are needed for highly structured material, small-step, fixed-sequence programs might interrupt the student's line of thought, or detract from the motivating properties of discovery methods. Also, he thought, the student is his own best judge of when an idea has been grasped, and this judgment is more easily exercised under self-direction than under programed instruction. Third, the individual memory span was believed to be best employed in a self-determined route of study with materials provided by the experimenter, thus bringing a need for a flexible size of step.

Of several subject-matter areas sampled in a series of experiments, the only significant difference favoring self-direction over programed instruction was for mathematics, and that difference occurred only after coached practice in self-direction. The self-directed group was provided the following materials: (a) a short basic text (actually a short linear program with only infrequent questions), (b) supplementary examples and explanations, (c) self-testing questions, and (d) a two-page outline of the entire lesson. The programed group used a linear program, self-paced. Both groups were free to consult the teacher at any time. Without coaching, there were no significant differences between the two groups. Since students benefited from the coaching in the use of the self-directing materials, there is a possibility that more prolonged use of self-direction methods would be needed without coaching for the superiority of the method to appear. Those who used self-direction with most benefit tended to be the better achievers among the students.

One drawback to the study was that whereas the author identified

mathematics as a hierarchical task, it is not clear that the implications of this were capitalized upon fully in either method of instruction, e.g., it is not clear that either the programed instruction materials or the self-directed materials made it clear to the student what the individual subordinate competencies were, and how attainment of each was to be signified. Had the "hierarchical task" been fully analyzed, as discussed in Chapter I, <u>and a copy of the hierarchy given to the student</u>, both procedures might have been more effective, and it is at least consistent with Campbell's theorizing to suppose that the benefits of self-direction might have been enhanced the most!

Other experiments reported in the article by Campbell failed to demonstrate differences between programed instruction and self-direction in Far East geography, principles of global geography, Presidents' names (ordered specifics), or history. While these materials, except Presidents' names, also are subject to the same comment made above about the apparent failure to make the subordinate competencies explicit, Campbell explains the negative findings in terms of the difficulty of jogging students out of their passive set for just absorbing materials as they are given out. He also points out that since self-direction was not significantly inferior to programed instruction in any of the subject-matter areas, but the materials often could cost less and take less study time than programed instruction, the self-directing method is economically important.

Granting Campbell's point that students long accustomed to didactic methods and a set to "receive" rather than to discover or self-direct their learning may require much practice in self-direction to make it work best, the reviewer would emphasize again the possible importance of making the structure of the task clear to the learner in terms of the achievement of each identified subordinate competency, so that the learner can focus his self-direction efforts upon mastery of each competency in turn. Some early years of experience with such learning might indeed eventually enable students to develop greater freedom and responsibility for larger-sized competencies. In this sense, the "size of step for self-direction" would then be increased with age of the student. Perhaps future long-term experiments such as those reported by Campbell can show in more detail how this can be achieved.

The reader may have already noted considerable similarity among the views and experiments of Mager, Campbell, and Pressey. Each of these have something to say about motivation, interest, self-direction, and the importance that the student organize for himself the information he reads. Ausubel, also, might be cited as sharing some of these views. The reviewer has granted the cogency of all these views, but has suggested that providing the learner an outline of the competencies he is to achieve and an indication of the order in which he should achieve them might be a procedure which would capitalize upon all these views and also put into effect the transfer among competencies demonstrated by Gagne to take place when learning is so sequenced. It may also be noted that the tendency for many experimenters to use together in an experiment an analysis of <u>hierarchies</u> of competencies and <u>programed</u>

instruction is not a necessary togetherness, although a convenient one to maintain the control of the experiment in order to test the inferred structure of the learning. This suggests an experiment employing sequenced units in accordance with the task analysis, the units being learned by methods under the control of the learner.

Since the reviewer has indulged his opinion in the preceding paragraphs, it might introduce a balancing note to say that Pressey, while still interested in adjunct autoinstruction, most recently tends to stress much the holistic view of learning--the need to grasp things as a whole, as in a poem or a picture.[1] Whether bringing in a perceptual principle in characterizing learning is best may depend in part upon the type of learning involved; it may be especially relevant to problem solving.

Type IV: Adjunct Autoinstruction; Mixed Experimenter and Learner Control

Pressey first developed and demonstrated a method of instruction now known as adjunct autoinstruction or auto-elucidation. Over a period of years he and his students employed several mechanical aids in the application of this technique (Pressey, 79, 80, 81) in a wide variety of classroom situations in order to answer a variety of questions about the technique (Pressey, 82; Stolurow, 105). As indicated by the above heading, this procedure may be considered a mixture of experimenter-determined control and learner control.

The procedure of adjunct autoinstruction is essentially a simple application of the timely and consistent provision of feedback to the student. A simple mechanical device or chemically-treated answer card giving one signal for "right" and another for "wrong" is employed with multiple-choice questions to enable the student to determine how well he has learned from study of a chapter in a textbook, listening to a lecture, watching a film or demonstration, or any other means of initial presentation of instructional materials, information, or ideas to the student.

The procedure is as follows: first, the student reads the chapter in a textbook (to use textbook study as an example), then indicates his answer to each question in a "practice test" or "self-check test." Since the feedback device signals "right" or "wrong" to the student after each response made, this feedback confirms correct learning and indicates which questions require further (remedial) study. The student can repeat the test after further study until a perfect score is earned, or different questions over the same material can be presented on successive "practice" tests. Since a wrong response is seen at once to be wrong, the student can correct his error by rethinking the question and choosing another answer, or if need be, he can refer to the page of the text given opposite the question on the test sheet, and re-read the

[1] Personal communication, 12 October 1967.

passage before selecting another answer, which will also be followed by feedback. Thus the student stays with each answer until he has it correct (sometimes called the "retained situation" to distinguish it from the "vanishing situation" in which only one response is made before going on to the next question).

In this application of the method, the student is encouraged to study the chapter by noting headings and overall organization, glancing at pictures and charts in the text, and otherwise "ranging over" the material before going on to reading the materials in their printed sequence. The practice test questions are not necessarily arranged in the order of the text, and the student is free to skip around on the items if he wishes.

One feature of adjunct autoinstruction is that it can be used to enhance or support the learning begun by the student with <u>any media of instruction</u> (book, lecture, film, demonstration, field trip, etc.). Besides employing the demonstrated effectiveness of corrective feedback and confirmation, it makes use of the "retained situation," which has been shown by Briggs (11) to be superior to the "vanishing situation" and to performance without feedback, as on a conventional test.

Pressey (82) has reviewed much of the laboratory and classroom evidence, which he and his students have gathered over a period of years, showing the superiority of the use of adjunct autoinstruction over various methods of study and teaching which do not employ timely and systematic feedback. By timely is meant the providing of feedback soon after the initial study effort has been made, not days or weeks after. By systematic is meant the application of the feedback procedure consistently, day after day, in the course of study. Stolurow (105) has reviewed thoroughly many of the research questions studied and classroom and laboratory conditions employed in research with the adjunct technique, including research in laboratory, classroom and field experiments with a related device called the Subject-Matter Trainer (Briggs, 12).

Since earlier research had established the effectiveness of adjunct autoinstruction as compared to normal classroom teaching procedures and to students' regular study effort, Pressey later conducted studies to compare the effectiveness and efficiency of adjunct autoinstruction with programed instruction, since he saw theoretical and practical drawbacks to the latter. These comparisons of adjunct autoinstruction with programed instruction have not involved the same massive research efforts as did the earlier research; because of this Pressey has labeled the findings as suggestive and requiring confirmation through larger-scale experiments, yet all the reported comparisons show adjunct methods invariably equal to linear programs in learning effectiveness and much more efficient in terms of learning time (Pressey, 83; Pressey and Kinzer, 84).

Due to the fact of Pressey's tendency modestly to emphasize the informality of his recent experiments and the need for larger experiments to check the results, it seems appropriate to cite the results as seen by another reviewer, Anderson: (3, p.152)

"Pressey and Kinzer have presented a drai .tic demonstration that the small-step, linear program can be unnecessarily redundant and wasteful of students' time. A group which completed the first 84 frames of the Analysis of Behavior did no better on a posttest than a group which read a 643-word summary of the same material. However, it took students eight times as long to complete the program as it did to read the summary. A group which read the summary and then answered nine 'auto-elucidative' questions actually did significantly better on the posttest than the group that received the program. No one has ventured a replication of this study yet. And, it is true that the study entailed a short, easy section from a rather redundant program, that intact classroom groups were assigned to treatments, and that the posttest consisted of a 'carefully graded essay examination.' Still the results of the study are not easily dismissed."

Hershberger (47) employed adjunct questions with tersely written prose, and with discursive prose which contained underlining and other typographical cues, simple and complex, to distinguish core material from enrichment material. A control group received the materials without typographical cues. The adjunct questions appeared in small groups after each page or so of prose. The self-evaluational test items were found reliably to enhance learning and retention of core content without distracting from enrichment content. Complex typographical cueing failed to increase learning, but simple underlining of core content increased performance on a core-oriented criterion. The effects of simple cueing and the use of self-evaluation test items are independent and additive in improving achievement. Low-ability students were benefited by terseness only if it was accompanied by test items. High-ability students benefit from either terseness or test items, or both. Low-ability students receive most benefit from self-test items and a terse text.

Ausubel, reviewing data on review and retention (5, p.118) made the point that rereading as review is most beneficial about two weeks after study, while test items as review are best used shortly after the original learning.

Concerning use of terse prose, Anderson (3, p.152) states: "The reviewer believes that the terse summary statement should become a benchmark for verbal, self-instructional programs, and other forms of instruction, too, for that matter. If a program cannot outperform a summary, it has no place in the schools."

Feldman (Margaret E., "Learning by programed and text format at three levels of difficulty," Journal of Educational Psychology, 56, June 1965, 133-139) administered programs at three difficulty levels and texts created from the identical words, by filling in the response term as a part of the text. Low-ability learners on a transfer test did better from the texts than from the programs. She concluded that for those learners, responding to program blanks interferes with the

student's organization of the materials, and that good charts can overcome the effects of difficult wording. On a cloze test of retention, no differences between programs and texts were found.

One aspect of Pressey's views and practices has only recently been subjected to experimentation. Pressey believed that the students should read an entire chapter of text before responding to self-test items. The purpose of the test items, as he saw it, were to confirm learning which has taken place, to correct errors, to provide practice in discriminating right from wrong among the multiple-choice alternatives, and to receive guidance in portions needing further study and correction.

Many persons, including this reviewer, have questioned this practice of delaying the introducing of the set of test questions until after the student has first studied the entire chapter. While not questioning the value of the first fast overview, the suggestion was that the questions be attempted, one or two at a time, as the student first reads limited portions of the chapter. Such an "intermediate size of step," it was argued, would be more effective and efficient than reserving all the questions until the whole chapter had been read.

So far as the reviewer is aware, Rothkopf was the first person to conduct an experiment dealing with this specific issue, although Hershberger had obtained favorable results using test items after every page or so of text.

Rothkopf (89) reports a study designed to find out if adjunct, test-like questions have generally facilitating effects upon learning from written instructional material, and whether it matters where these questions are asked in the course of reading. He cited the findings of Hershberger, reviewed above, but Rothkopf believed it remained undetermined whether previously reported benefits of self-testing were a specific effect or a generally facilitating learning set. This question remains, Rothkopf believed, because the test items used by Hershberger during the reading were similar or identical to the test items employed in the criterion test.

Rothkopf inserted the adjunct test questions at different intervals in the reading materials for different groups of students, and within intervals, some students had questions only while others received both questions and answers. Thus some groups responded to the test questions before beginning reading of an entire chapter, while other groups had a few questions before or after the reading of each section of the material. Still other groups received the questions after reading was completed.

Rothkopf concluded that the adjunct questions shape effective inspection behavior and thus are highly useful in teaching the specific skills which were the objectives of the printed material. He further found that responding to questions <u>after</u> reading the materials provided both specific and general facilitative effects upon learning, whereas providing the identical questions before the reading yielded only

specific effects, which effects are greater with answers given than with answers not given. In general Rothkopf's data suggest that an optimum place to insert adjunct questions in the text is after approximately 1,000 words of reading. This may be contrasted with the ratio of blanks to words in linear programs, in which a response is often required after every 10 or 15 words of reading. Rothkopf's findings also appear to confirm other findings suggesting that adjunct methods are more economical in learning time than the prevailing type of linear programed instruction.

It is concluded, then, that adjunct methods can be employed in such a way that the learning achieved goes beyond the content of the specific questions employed, in agreement with earlier findings reported by Pressey (82) that the learning transfers to an end-of-course test in which all items are different from the practice-test items. One procedure which Pressey employed to insure such an effect in a course in educational psychology consisted of use of three sets of adjunct questions per chapter, the first designed to test for the content of the chapter, the second to require use of the content in making judgments, and the third giving the student questions requiring applications of the content to practical, lifelike situations. In use of each of the three tests, a student first tried the entire set, and then tried again only those questions missed the first time. Thus items passed the first time are "dropped out" on the second time through. A "card sort" device designed by Briggs (12) made such item dropout automatic from trial to trial, so that a student stopped work when all items had been dropped out (responded to correctly).

Skinner (102) has opposed the use of multiple-choice items for learning purposes on the grounds that the wrong alternatives presented, if chosen by the student, are learned and remain as errors in the student's later responses. Pressey has countered that the use of feedback in the "retained situation," in which a wrong response must eventually be followed by the correct response, insures that the student learns both to make the correct response and to recognize that the wrong response was wrong. Furthermore he also learns why a wrong answer was wrong, due to the feedback and the restudy effort which negative feedback requires. Pressey also argues that a practical goal in life is to discriminate right from wrong answers, and that this training is provided by the adjunct method of responding.

Karraker (55) tested the hypothesis that plausible wrong responses made to multiple-choice self-test items are recalled later as being correct, as Skinner contends. Seventy-two college freshmen were stratified into high and low ability levels, then randomly assigned to conditions of no knowledge of results or knowledge of results following responses to multiple-choice questions. All students first responded to a multiple-choice test in educational psychology and then took a recall constructed response criterion test. The score the student earned was based on the number of responses that were the same as the plausible wrong responses made in the multiple-choice test. Under the no knowledge of results condition, significantly more errors were made on the criterion test than under the knowledge of results condition. The

conclusion is that when knowledge of results is supplied after responses to multiple-choice questions, the student does not later recall plausible wrong responses as correct responses.

When an objective of the learning is to be able to discriminate right answers from wrong answers representing common misconceptions actually held by other people (and this is usually the basis for selecting the "mislead" alternatives), the multiple-choice question has much merit. This does not mean that the study materials or the way the questions are phrased should lead the student to make errors. Rather, the material read should enable the student to choose the correct answer and to discriminate it from wrong answers.

For moderately difficult conceptual material, there is some evidence that the question should be challenging enough to require some effort to respond correctly (Goldbeck and Campbell, 42; Goldbeck and Briggs, 41; and Feldman, cited above), and that under these conditions actually being required to make an overt response (in these studies a written, constructed response), is better than simply being shown the right answer.

On the other hand, for memorization tasks it is better that the student be shown directly what the answer is; that is, for a memory task direct prompting is more effective than confirmation (Briggs and Angell, 14; Briggs and Hamilton, 16). For memorization tasks and for some conceptual tasks, it may or may not matter, when the student is directly shown the answer, whether he overtly repeats the answer or not; except for learning to spell words by rote. Probably if it is an associative, conceptual response, he need not overtly make it, but if it is involved with response learning (such as spelling), the student should actually make the response after being shown the correct one (e.g., after being prompted). Kaess and Zeaman (54) presented students with learning material relating to definitions of psychological terms, in the form of multiple-choice questions with the right answers underlined. Students who merely observed the correct underlined answer did as well as students who observed the correct answer and then overtly punched that answer on a punchboard (a mechanical form for recording multiple-choice answer selections).

While the nature of the task determines whether multiple-choice responses or constructed (open end) responses would be best in learning under conditions of feedback after each response, and while the nature of the task also determines whether or not direct overprompting of responses is desirable as compared to less highly prompted, challenging responses, made by context cues of minimal strength to lead to a correct response, it can be said without equivocation that the making of initial wrong responses is almost never desirable pedagogically. This is not, however, the same as correctly discriminating right from wrong on a multiple-choice test. But to write materials and test questions in such a way as to provide the student with just enough help (but no more) to enable him to respond correctly is an impossible programing task without very complex equipment, except by resorting to copy frames. In short,

while there is evidence that for much conceptual learning one should prompt close to the threshold of the right response, rather than far above it, this cannot be judged precisely by the programmer, so it is often (not always) better to provide feedback after each response, thus enabling the student to correct the wrong responses he does make due to miscalculation on the part of the programmer, than to make a too-redundant program. Adjunct methods provide these conditions, and when responses to multiple-choice questions are suitable, as opposed to the alternatives of spoken responses or other constructed responses, adjunct methods are more economical in time than are constructed responses as required in most forms of programed instruction.

Due to the fact that the study materials are separate from the test questions and the feedback mechanism (punchboard, chemoscore card, etc.), they provide an admirable set of conditions for remedial instruction, often called "refresher training" by the military. This is because the student can be asked to attempt each question in turn, and to read the available materials only when he makes an error. (The page number to be read is simply noted next to the test question.) "Remedial" instruction in education may take place soon after the original learning, or years later. "Refresher" training normally is given months or years after the original training, the interval being used for work on a job either related or unrelated to the training. Meyers (72) has reported adjunct methods to be effective, economical, and well received under these circumstances.

When multiple-choice items are used for adjunct autoinstruction, how many alternatives should be offered in addition to the right answer: 1, 2, 3, 4, or more? How many alternatives are best for conventional testing or evaluative purposes? The issues are different in the two cases.

Taking the instructional use of multiple-choice questions first, and recalling that it is desired for the student to not only learn the right answer but to discriminate it from commonly made wrong answers, one issue then becomes, "How many commonly made wrong answers are there?" If there is only one confusable, often-made wrong response, clearly a two-choice item is called for, as it is useless to introduce irrelevant or implausible wrong answers. (The number of common wrong answers can be empirically determined by giving the test questions in open-end form to a group of students like those to be taught by the adjunct methods, after they have read the materials.)

But if there are many often-chosen wrong answers, then what? One approach is to rewrite the materials to reduce the frequency of choice of some of the wrongs. Another approach is to elicit the right answer once in one item, and if the matter at hand is important, give practice in discriminating the right from many wrongs in successive items.

If there are a limited number of likely wrong answers, then in general the smaller the number of misleads the better. Kaess and Zeaman (54) found that the fewer the number of distractors presented initially, the

more quickly the students mastered the materials in subsequent trials. Wrong responses made initially did tend to be repeated on later trials, even when feedback was given consistently. Students who were just given the correct answers initially made fewest errors on later multiple-choice trials. This finding is in agreement with those of Briggs (11), that on four-alternative items both rights and wrongs tended to persist even though wrongs may be overcome after repeated trials. Although errors did become fewer in subsequent trials, it is best to avoid initial errors as it takes time to correct them. It is not known, however, whether such eventually-correct responses are retained over long periods less well or better than initially correct responses. Again, arranging it to strike a balance between a meaningful, thoughtfully made right response on an important issue and the avoidance of initial error takes much care. But the data on rote learning in favor of over-prompting as a way to avoid error nevertheless should not carelessly be generalized to conceptual, problem-solving situations. The desirable means for avoiding initial error are different for problem solving than for memorization.

With reference to multiple-choice items for conventional evaluation, in general more alternatives are desirable than is the case for learning, within limits. Although five-choice items were widely used a few years ago in standardized tests, the present trend is to four-choice items, for good reasons. Again, there often are not many functional (frequently chosen) wrong answers. Even when there are many possible wrongs, it has been shown mathematically (Tversky, 110) that the optimal number of alternatives is three, with reference to the amount of discrimination information each added alternative yields to the measurement. But at this point we need to relate the discussion of adjunct methods back to the emphasis in Chapter I upon subordinate competencies, and also to other issues raised in the present section of this chapter.

Why did Rothkopf find, in support of the views of Pressey, that adjunct questions aid learning most when coming after a moderately lengthy reading passage (1,000 words)? Why would they not function best when interspersed among paragraphs in the text? In the reviewer's opinion, the reason is that materials in texts are not optimally sequenced in accordance with subordinate competencies. Thus the more loosely organized in this respect the text is, the greater the need for the student to range about in the text to organize his thinking. For this reason, meaningful understanding of the entire learning task covered by the chapter may be developed by the student independent from the sequence of paragraphs and independent from the order of the questions on the test sheet. If this is so, it helps explain Campbell's findings, cited in an earlier section, that student self-direction among four parts of the total materials was sometimes better than following the sequence in the programed instruction booklet.

If materials could be sequenced in the order that the competencies are needed, the reviewer suggests, then posing a few adjunct questions after each portion of the text would be advisable, and in this case, the order of responding to the questions should be fixed to correspond to the sequence of paragraphs in the text. But if texts remain faulty in

this sequencing sense, then the questions need not be presented and answered in a particular order, since the student will re-order both text and questions in his study effort. It would still help, however, to cross index questions with page numbers, to make the student's remedial study more convenient.

The task of discovering how to sequence the text in a way meaningful to the student, as Campbell suggests indirectly, may be so difficult that it is better to use texts as they are with adjunct methods than to rewrite the texts or rewrite the programed instruction. At least with present methods, the rewrite attempt may be no more suitable for the students than the original version, since writers and programmers sequence materials differently than the student would, as suggested in Mager's study, cited earlier.

Now is the sequencing problem really this formidable in texts and self-test questions, as well as in other types of programs relevant here, or haven't we yet developed an effective approach to instructional materials design? Would a greater benefit result through a broader and more intensive use of empirical procedures to keep the designer more in close touch with the student through such expansions over the usual empirical methods as those used by Mager (63) and by Markle (66)? Mager used empirical procedures to find out what of relevance the learner knows at the outset, and how he would prefer to learn. Markle employed empirical procedures to define the learning problem, to determine how long to use films before switching to printed materials or manual practice, and to bring achievement of the experimental group taught by an empirically designed multi-media course up so much that the distribution of scores for the experimental group did not overlap the distribution of scores for the control group. No statistical significance tests are required to determine the relative effectiveness of such a successful instructional design, made possible more by the empirical procedures employed than by use of theory. Would use of both theory and empirical procedures be even better? This is the basis for the procedure for the design of multi-media courses reported by Briggs *et al.* (15).

Chapter IV: Experiments Involving Pre-planned Instruction

In this chapter are reviewed those studies in which the experimenter rather than the learner has primary control over the sequencing of instruction. As shown in Chapter II, the specific ways by which sequencing may be controlled by either the learner or the experimenter vary considerably among the experiments reported. In Chapter III, several types and degrees of learner control over sequence were described and the results reviewed. In this Chapter IV, several forms of experimenter control are reviewed. Also, as implied in the title of this chapter, experimenter control usually involves pre-planned instruction, arranged in advance in form of a linear program, a branching program, or printed material. It may be noted that with pre-planned instruction, the experimenter may be primarily interested in getting the units of instruction in the optimum order or he may be interested in the order of program frames within a unit. In one group of experiments, it is the use of introductory material which is given primary attention.

Type V: Experimenter-determined Sequencing of Instruction in Accordance with Hierarchies of Competence

The reader may note that the above heading coincides closer than do other headings with the title of this report. The title of the report, in turn, reflects the title of the original proposal to conduct a review and the title of the contract authorizing the review. In short, the project was planned most directly at the outset to report the results of the experiments here labeled "Type V." In this sense, the present section is more central to the original purpose of the project than are others.

However, as work on the project progressed the pervasiveness of the problem of sequencing became increasingly apparent, and the scope of the experiments reviewed kept expanding. In this sense, the report "outgrew" the original contract plans, but for the sake of administrative consistency the title remained unchanged.

Said differently, the original plan was to review experiments in which sequencing was arranged to implement an explicitly-stated experimental task structure representing a hierarchy of task components, the structure being represented in such a way as to indicate the presumed direction of transfer from one subordinate task component to another. However since many other kinds of experiments became of interest, it appeared that sequencing was the most common element among all the experiments reviewed.

Viewing, then, all experiments reviewed in this report, it can be said that:

1. All experiments reported involve the sequencing of instruction.

2. Many experimental reports identify the structure of the task, in some degree at least.

3. Some experimental reports identify the task as hierarchical in structure.

4. A few experimental reports (by Gagne and others) show the hierarchical structure, and the expected direction of transfer, and furthermore present the results of the sequenced learning in terms of verifying the assumed task structure.

In this section of the report, then, are reviewed those experiments with experimenter control over sequencing in which the authors said that their experimental tasks were hierarchical in nature. As will be seen, not all authors described their tasks fully enough that the learning data can be accepted as evidence for or against the presumption of hierarchical structure.

Experiments by Gagne and his Associates

This series of experiments is distinguished from the Other Hierarchy Studies (next sub-section) because it is only the experiments by Gagne and others which were conducted and reported so as to include all the following elements: (a) an explicit rationale as to how productive learning takes place; (b) a schematic representation of the task components among which transfer is predicted to occur in a given direction; (c) a representation of entering competencies in the form of both aptitudes and relevant skill components; (d) a means for testing for the actual occurrence or non-occurrence of the predicted transfer in terms of specific sub-skills, rather than a more general inference as to the validity of task structure based on overall effectiveness of the sequence designed to be appropriate for the inferred structure.

In an early report in the series of experiments reported here, Gagne (29) differentiated between the acquisition of knowledge or capability on classes of tasks or stimuli from the acquisition of particular responses, thus distinguishing productive learning from reproductive learning. Using the task, "Finding formulas for sum of n terms in a number series," he first presented a type of task analysis which shows the inferred hierarchy of capabilities supporting task performance (29, p.359). The capabilities in the hierarchy were identified by starting with the criterion task and asking, "What would the individual have to be able to do in order that he can attain successful performance on this task, provided he is given only instructions?" This question was then repeated successively for the subordinate classes of tasks identified by the answer. Gagne theorized that the competencies so identified serve as mediators of positive transfer from lower-level competencies to higher level ones. Starting with the acquisition of lowest level competencies, attaining each higher one depends on positive transfer, which is dependent in turn upon recall of lower level competencies and the effects of instruction. (Note carefully that this recall is from one unit of instruction to the next unit, during the initial learning period. The cruciality of this point will be made clear later.) At the very bottom of such hierarchies may be found either the entering relevant competencies

brought to the course from prior learning, or in the case of the task being illustrated by Gagne (above), basic abilities were identified at the bottom (symbol recognition, recognition of patterns, and number abilities).

Gagne administered test items over each competency to ninth grade boys. Each boy was tested first for the final task; if he failed, the test for the next highest competency was given and so on down the hierarchy until the boy passed on a competency. At that point the testing stopped temporarily, and a learning program designed to teach the next higher competency (previously failed) was administered. Then the test items were given over the remaining levels. The results showed, first of all, that the boys exhibited wide individual differences in patterns on the tests, illustrating different "entering competencies." There were no instances in which a person able to perform a "higher level" competency failed on a lower one. After completing the instructional portion for previously failed competencies, the boy was again tested, yielding scores on competencies both before and after administration of the learning program. Though there was no control of or correction for the effects of the added test trial, as the author acknowledged, comparison of "before" and "after" scores on each competency supported the idea of the knowledge hierarchy.

In reference to the abilities at the bottom of such hierarchies, Gagne theorized that if learning programs were of perfect effectiveness, everyone would pass all the component tests in the hierarchy, the variance would be zero, and all correlations of tests on the various competencies with basic abilities would also be zero. But if learning programs are not perfectly effective, the probability that a person will acquire each competency will be increased to the extent of his score on a test of the basic ability. Further hypotheses about the ability measures were investigated in another study, reported next.

Gagne and Paradise (36) went on to deal with sources of individual differences in <u>rate</u> of learning from a learning program. They suggest that observed differences among learners in rate of learning are not due to variations on some underlying general ability to learn fast, but rather upon (a) the number and kinds of learning sets (competencies, knowledge) the learner brings to the situation and secondarily (b) upon his standing in respect to certain basic abilities relevant to the competencies to be acquired as they are identified in the theoretical hierarchy for the task, and (c) his level of general intelligence.

Reasons for failure to learn to perform a task after taking a learning program could rest upon the following: (a) some subordinate knowledge may have been left out of the learning program, (b) insufficient practice or some other program characteristic may have resulted in poor recall of a subordinate competency, or (c) the program may have been defective in guiding thinking required to induce the necessary integration of subordinate competencies.

Gagne and Paradise hypothesized that a substantial amount of the

variability during learning is attributable to the attainment or non-attainment of learning sets relevant to the final task the program is designed to teach. "Accordingly, as the learner progresses upwards in the hierarchy, his rate of learning should depend increasingly upon relevant abilities" (36, p.3). In contrast, the correlation of general ability with rate of learning should remain constant as the learning progresses up the hierarchy.

As contrasted to rate of achievement, the fact of achievement rather than nonachievement, in a moderately ineffective learning program, might depend increasingly upon relevant abilities as one goes up the hierarchy, e.g., those of low ability "drop out" due to failure on some competencies, while those of more ability continue to master competencies in spite of defects in the program.

The task of solving linear algebraic equations was selected to test some of the above hypotheses. A learning hierarchy was designed by the logical analysis procedure previously described. The resulting hierarchy (36, p.3) showed six levels of relevant competencies and, at the bottom, three specific relevant abilities, such as defined by factor analysis studies. The authors recognized that other persons, especially proponents of "modern mathematics" might derive quite different hierarchies. It is not a matter of there being only one "right" analysis; it is a matter of finding empirical "validation" for the method at hand, in terms of the hypotheses to be tested.

The procedure employed in the study by Gagne and Paradise was as follows: First, basic ability tests were administered. Then the learning program was given, using eight booklets over eight class days. Then a performance test of solving equations was given, followed at once by a transfer test involving unfamiliar forms and unfamiliar symbols. On the next day tests were given on each learning set in the hierarchy. Learning rate was recorded by having students draw a line on their program answer sheets upon a signal given every three minutes.

As to overall achievement on the performance test and transfer tests, the results indicated that the program was only moderately effective.

As to transfer from learning of subordinate sets to higher ones in the hierarchy, students were scored as "pass" or "fail" on test items over each of the component capabilities. Patterns of pass-fail from lower skills to higher ones were then analyzed for each learner. A pattern of Higher +, Lower +, indicates learning was achieved on two successive levels of competence in accordance with the theory. A -- indicates failure on such a set of two competencies in accord with the theory. A + -, a pass on a higher level but a failure on a lower one, is contrary to the theoretical prediction. A - + is not in opposition to the theory, for a less than perfect program. (Each + or - was based on a single test item, due to time limitations. Two items for each competency would have yielded more reliable measurement, the authors point out.)

A tabulation of the frequency of the above four conditions on pairs of test items for adjacent levels in the hierarchy confirmed theoretical predictions on transfer from one competency to the next in proportions of instances ranging from .91 to 1.00.

On the matter of the correlation of basic "relevant" abilities with total performance (performance test, transfer test, competency tests, and learning time), the values were relatively high, as predicted, as compared to correlations of judged "irrelevant abilities" with performance.

Correlations among the various performance measures were fairly high, with persons doing well on the total task being those who passed the greatest number of tests over subordinate competencies. Also the best performers were the fast learners and the high scorers on the transfer test.

On the relations between relevant abilities and learning sets, the results generally confirmed the prediction that rate of learning depends decreasingly upon relevant abilities, whereas the achievement of the learning set depends increasingly on such abilities as one goes up the hierarchy. There was also evidence for the corrolary assumption that since rate of learning comes to depend less on relevant basic abilities, as learning progresses up the hierarchy, it does increasingly depend upon achievement of successive subordinate competencies.

An implication of the findings is that as learning progresses, relevant abilities will correlate increasingly with achievement only for programs which are somewhat defective in mediating transfer from one level to the next.

Other implications of the findings are that (a) the sequences of units in a program should correspond to the hierarchy so that instruction on each subordinate competency will be given when it is most needed, and (b) that frames within a unit of instruction for a given competency should provide recallability of previous learning and guidance for thinking enabling the learner to integrate prior learnings in the performance of a new task. On the characteristics of frames which provide such functions, more will be said later in this report.

Finally, if programs are designed in accordance with learning hierarchies, and tests made over each subordinate skill, these tests can be useful for diagnosing needs for remedial instruction and for revising the program.

Further validation of the general theory of learning hierarchies is contained in a report by Gagne *et al.* (35). In that study a learning program on addition of integers was prepared in four different parallel forms. The content was analyzed into a hierarchy, as in the studies reported above. Instances of confirmation of the theory that acquisition of higher competencies in the hierarchy depended on prior mastery of competencies lower in the hierarchy ranged from 97 to 100%. The

four parallel programs were designed to determine whether variations in amount of guidance and variation in amount of repetition would have marked effects upon achievement. The results for those two independent variables in the programs were mostly negative. The relative effects of those two programing variables were small in comparison to the consequences of adding or omitting subordinate competencies. Stated differently, if moderately good instruction is provided in the proper sequence, as compared to instruction not so ordered, the effects of this may overshadow other qualitative features in how the material is programed. This may account for the frequency of "no significant difference" findings in research designed to isolate "style" aspects of programing.

Still another task in mathematics was analyzed into a hierarchy by Gagne and staff, University of Maryland Mathematics Project (37), and self-instructional programs again were used to study effects of still other programing variables. The task in elementary geometry consisted of "specifying sets, intersections of sets, and separations of sets, using points, lines, and curves." In this study, the importance of sequencing of topic order was again confirmed in terms of the number of instances confirmed of higher competency acquisition being dependent upon acquisition of those lower in the hierarchy. However, the variables of variety of examples during learning and passage of time between stages of learning had no evident effect upon the learning effectiveness of the program.

Despite the above negative findings in regard to programing style variations (e.g., variety of examples) and their effects upon task acquisition, it was thought desirable to measure retention of knowledge for the same students. In a follow-up study, Gagne and Bassler (34) measured the retention of these students both on the total task and on each component subskill nine weeks after the learning. The retention for the entire task was very high except for one group which had previously received a narrow variety of examples in the learning program. The level of retention, overall, ranged from 108% to 128% (this was not labeled as evidence of reminiscence, due to the presumed facilitating effects of repeated testing). In contrast, the level of retention for subordinate competencies ranged from 60% to 88%, showing the individual skills much more susceptible to forgetting than the performance on the task as a whole. This difference in retention of the part-skills vs. the total skill cannot be considered evidence that the part-skills need not have been learned in the process of learning the whole skill, because the contrary was shown to be the case in the original acquisition data.

From the practical point of view of maintaining ability to perform this task, the forgetting of the subskills which originally helped master this task is of no importance, as these learners retained (or even gained) competency on the task as a whole. But if some of these same subskills are needed for new tasks to be learned later, this loss in retention of subskills is important, and deserving of efforts to prevent in future programs for other students; some remedial work for these experimental students could improve learning of related tasks later, if the forgotten competencies are involved in the later tasks. The implication is that

this program should be revised to exercise more the subordinate skills for students who will take the program later. For students in this experiment, relearning would be necessary in order to master new tasks which are dependent on the same subskills. One difficulty for insuring retention of subordinate skills is that the data show that whereas there is high continuity (within the hierarchy) in skills learned from the program, the patterns of skills lost over the nine-week interval are highly random, with the pattern varying among persons. Thus learning goes "in accordance with the hierarchy," e.g., a person learns all the competencies up to a point and then starts failing consistently or he keeps learning and masters the whole task. Forgetting, on the other hand, is random rather than ordered along the lines of the hierarchy. This finding appears to make difficult both the prevention of forgetting of subskills (as more repetition would be needed in all the parts of the learning program), and the correction of such forgetting, since for a group of students, somebody would need remedial work on about all parts of the program.

It is possible that this finding in regard to loss of subskills has an important bearing upon the issue of linear vs. branching programs. Even though the bypassing of instruction on a competency the student already possesses seems reasonable in terms of learning economy, the added exercise such a student would receive on a linear program might be worthwhile in retarding forgetting. Just when added repetition, or review, should be scheduled is however a complex issue in itself. At any rate, the importance of studying retention of linear and branching programs is clearly a relevant part in evaluating the relative merits of these two types of program.

Discussion of experiments by Gagne and others. Anderson (3), in a recent review, states that the notion of hierarchies as dealt with by Gagne and his associates cannot yet be said to be definitely tested. He cited two reasons for his statement: (a) that the correlational type of analysis employed by Gagne is not sufficient evidence of the hierarchy notion, and (b) that an experiment by Merrill (70) had resulted in findings contrary to Gagne's hypotheses concerning hierarchies.

In this sub-section, Anderson's first objection is discussed, e.g., that based on correlational analysis as evidence. His second comment, concerning the experiment by Merrill, will be taken up after that experiment is described in the next section of this chapter.

The first thing that might be noted about Anderson's statement regarding correlational analysis as evidence is that there is one study which is not solely correlational in its evidence on hierarchies of competence, namely the first one cited in the series by Gagne (29). Here, "before" and "after" comparisons were made. The point about this study is that it showed, first, by correlational evidence, that the task was hierarchically arranged. Then, as a next step, it deliberately set out to change the individual learners, and achieved considerable success at this. The latter method, despite its drawbacks, is an experimental procedure. The other experiments in the series also demonstrate by

correlational methods that the task is hierarchically arranged. What they did not do is take the additional step of experimental manipulation to change the learners.

It is noteworthy that Gagne's original study (29) showed first that learners could not progress in learning task components beyond a certain point (different for different people) and then that these same individuals could so progress when specific learning was undertaken to insure mastery of successive components. Thus this method showed first that the task was hierarchically arranged, and then that the "progressive mastery" hypothesis was tenable. This type of "before" and "after" evaluation of performance has the advantage of highlighting the specific things the person could not do before training, then showing how his subsequent performance can be improved by training on these (for some Ss) limited parts of the task. (The implications for efficient remedial education should be apparent also.)

This "before" and "after" procedure, as conducted by Gagne, has some disadvantages, from the point of view of experimental design, which he pointed out, as indicated in the preceding review of his experiment. From the practical point of view of how large a difficulty a learner can be faced with in experimental situations, his procedure is perhaps better than the three-group experiment recommended by the reviewer early in this report. In that procedure, Ss faced with training by randomized sequences (or worse yet by "inverted" sequences contrary to the implications of the inferred hierarchy) might be so frustrated (if the task chosen is hierarchical) that it might be difficult for him to persevere in his efforts until the desired test data were completed. Perhaps, however, such experiments should be attempted; the use of measures of perseverance on the task and frustration measures might yield supplementary data.

It would appear, then, that Gagne's original experimental procedure is about as good a test of the hierarchy as can conveniently be designed and implemented. And, as has been stressed here, the data are not restricted to correlations in performance among adjacent levels of subordinate competencies.

Other Hierarchy Studies

Merrill (70) employed a corrective procedure of instruction intended to make sure that the learner did master each subtask before he was given instruction on the next higher subtask in the hierarchy. His experimental instruction consisted of programed instruction in a hypothetical science in which principles and other information could be hierarchically ordered. Criterion performance was ability to make satisfactory predictions and interpretations of the workings of the hypothetical science. A detailed linear program was employed; it was sequenced in accordance with the hierarchy. The administration of the program was followed, however, by several conditions of practice and correction before the criterion test. Merrill states that the group which performed best received

only a summary of each lesson (or level in the hierarchy) prior to the correction review exercise which preceded the retention tests. Since both the linear program and the summary were reported to have been sequenced in accordance with the hierarchy, the findings appear relevant not to the validity or nonvalidity of sequencing in accordance with the hierarchy but rather to the comparative effectiveness of the program and the summary. Merrill himself makes the statement, "Apparently the summary statements and review technique used in this study were very effective instructional techniques." It might be a corrolary that the program was not such. It may be recalled that terse textual summaries have previously been found superior to programs in effectiveness (Anderson, 3). Viewed this way, Merrill's findings are not regarded by the reviewer as evidence either for or against the hierarchy notion.

An alternate interpretation is that Merrill did not have a learning task whose elements formed a hierarchy (i.e., depended upon each other for learning facilitation). His diagram might logically appear to support the belief that the task was a hierarchical one, but the empirical results do not bear out that expectation. The finding that for Merrill's task it made no difference in learning outcome whether prior topics were mastered or not is directly contrary to Gagne's results (29), suggesting that either the two tasks are not of the same nature or that sequencing was not done well as an implementation of the task structure. It seems to be going beyond the data to conclude, as Merrill does, that for a hierarchical task, mastery of an early sub-task does not facilitate learning of a later sub-task higher up in the structure.

In a follow-up study, Merrill and Stolurow (71) turned attention to the summary procedure, found to be an effective (control) condition, used in the previous study. The purpose of their study was to separate out the effects of three characteristics of this control condition in the earlier study. Accordingly, a hierarchically organized summary was presented in one group as the initial learning condition before S attempted to solve problems. In other groups the same summary was used as a remedial technique following errors on practice problems and was compared in effectiveness as a review technique with two other conditions: (a) presentation of only those summary statements which had a bearing on the problem at hand regardless of its hierarchical position, and (b) specific review consisting of step-by-step presentation of the solution of the problem.

The results of this second study, as interpreted by the reviewer, are as follows: The terse summary of system information and principles, sequenced in accordance with the hierarchy which was analyzed for the task, was the best initial condition of instruction. However, once a student began making errors on trying to apply the principles to the solutions of practice problems, the best remedial condition was the specific summary consisting of step-by-step presentation of problem solution. While the reviewer's interpretation differs from the language in which the report is described by the authors, it is believed that the most significant finding is that when problem solving is the task, a terse, well-written statement of information and principles presented

in a sequence in accordance with the hierarchy as the task is analyzed represents an effective initial instruction procedure.

It is risky to have the S begin trying to solve problems before he has a complete preview such as the hierarchical summary because providing correct answers to problems as the only instructional information is not effective. Once an error is made it is best to review the correct solution, then to use the summary as a remedial technique or to present again those selected principles which should aid the S in arriving at a solution. Other experimenters have found errors once made difficult to correct. Campbell (19) found that detailed branching instruction after detection of a misconception often failed to lead to its correction, thus confirming the importance of devoting much care in developing the basic program as compared to dealing later with errors arising out of failure of the program.

In summary of the experiment by Merrill and Stolurow, the reviewer feels that issues other than the validity of the notion of hierarchies were being tested. These issues seem to include: terse vs. detailed instructions; how to correct errors on problem-solving tasks; how to balance initial and remedial instruction; when to introduce problem solving practice with reference to prior exposure of relevant principles. Viewed somewhat differently, having discovered a terse summary which was better than a linear program, the experimenters then sought to appraise the summary when used (a) for initial training; (b) for remedial training; or (c) for review, as compared to using summaries only on problems failed or trying to conduct remedial review by giving the answers to the problems.

Buckle (18) also reported an experiment designed to test Gagne's theory concerning vertical transfer among subordinate competency levels in a hierarchical structure for a task. In his introduction Buckle cited two experiments which he interprets to provide evidence contrary to Gagne's hypothesis concerning hierarchies. The first experiment he cites is that by Merrill (70) reviewed above.

Next, however, Buckle cites the experiment by Gagne and Bassler (34) as additional evidence against the theory. He says, "It was found that on the retention tests higher level learning sets were achieved while subordinate sets had been forgotten." Apparently Buckle overlooks the significance that these retention data presented by Gagne and Bassler were based on tests given nine weeks after completion of learning of the task. In citing this as evidence against the learning hypotheses regarding a hierarchical task, Buckle apparently also misconstrues Gagne's hypothesis concerning hierarchies. Gagne's hypotheses do not include the prediction that what is found by a learning experiment concerning component to component transfer within hierarchies would also be found on measures of retention. In fact, as discussed earlier in this report, the interesting feature of the experiment by Gagne and Bassler was that whereas subordinate competencies can be shown to be needed in the initial acquisition of a task capability, the task itself may later be performed satisfactorily even though some of the subordinate knowledges have been

forgotten. Hence citing the findings from Gagne and Bassler is not relevant to the theory of hierarchies, as made explicit in various of Gagne's reports. The principle distinction missed by Buckle is the distinction between the learner's ability to recall an earlier competency while attempting to master a new one during the learning period, and the learner's ability to recall specific components of the task nine weeks later.

In the course of his report, Buckle states an hypothesis based on the work of Ausubel (5) who distinguishes between derivative learning and correlative learning. In derivative learning information presented late in an instructional period might be apprehended perfectly well by the learner, as he can derive its import due to the cognitive structure of the moment rather than by specific recall of earlier supporting material. In correlative learning, on the other hand, the learner would not be able to grasp the meaning of a new item of information if he had forgotten a related item presented earlier. Buckle cites derivative learning instances as contrary to Gagne's hypothesis, and correlative learning instances as supportive of Gagne's hypothesis. More specifically, Buckle advances this hypothesis, "Symbolic learning will undergo correlative subsumption, while verbal learning will undergo derivative subsumption for both overt and covert response modes during learning." In essence, Buckle appears to mean that Gagne's hierarchy ideas would be relevant for symbol learning but not for verbal learning. Learning is the form in which Buckle's hypothesis is stated, but delayed retention is the condition under which he tests the hypothesis and relates the results to Gagne's hypotheses. While there are other matters which are confusing in Buckle's report, these will not be enumerated here, since results reported are based on delayed retention measures.

When good retention was shown on symbol test items on the delayed tests, Buckle cited this as evidence that the material was meaningful or derivative in nature in accordance with his hypothesis. Lower retention for the verbal responses he cites as an indication that the material was not meaningful.

Buckle did not present his analysis of the task to enable the reader to determine for himself whether, in his judgment, the task was hierarchical in nature. There appears to be danger of administering a linear program and labeling it as instruction for a hierarchical task without providing sufficient data as to the nature of the task.

Comment by the reviewer. Referring back now to the earlier cited criticism by Anderson of Gagne's hierarchy hypotheses, with reference to the experiment by Merrill, discussed above, our conclusion has been that Merrill went beyond his own data in his interpreting his task as hierarchical. It must be said that positive results, showing that mastery of an early skill component does transfer to learning a later one is a supporting argument to the notion of hierarchical structure when viewed in terms of all the kinds of evidence Gagne presents. On the other hand, negative findings, e.g., failure to find transfer from one sub-task to another, may simply mean that that task was not hierarchical. Also,

before such negative findings could be viewed as evidence against the hypothesized course of learning in hierarchical tasks, it would be necessary to be sure that the negative results are not subject to one of the five experimental errors listed in Chapter I.

It must be constantly kept in mind that there is no known hypothesis to the effect that all tasks are of the hierarchical sort. However if an experimenter classifies his task that way, the basis for this should be shown. Such a basis was not shown in Buckle's report. Neither Buckle's study, nor Merrill's, nor that of Merrill and Stolurow criticized their task structure on basis of the empirical findings reported.

Practical Factors in Applying Hierarchies

Schutz, Baker and Gerlach (97) evaluated a 276-frame constructed response program in capitalization in the following way: They arbitrarily set as the criterion of success a score of 33 out of 40 correct and administered the program to 4th graders and 7th graders. They took the precaution of determining that no subject in either group could meet this criterion on a pretest. From the results on the criterion test they concluded that the program works very well for some individuals but not for others, and they further analyzed criterion performance and concluded that Ss who failed to learn the capitalization rules during the program failed to show improvement from pretest to posttest even though they filled in responses to the program correctly. They concluded that learning the rules is an important prerequisite to capitalizing samples correctly, and they further recommended that in revising this program the student should be required to state the rule before filling in the capitalizing response to any problem posed in the sequence of frames.

Analysis of the total results also revealed that some students did not have the prerequisites assumed to be present at the start of the program. Thus a need was seen in practical applications for pretesting the students to be sure the assumed prerequisites are there, then assuring that the students can recall and state relevant rules before having them make constructed responses to capitalization instances in the frames of the program.

This appears to be a clear example of programs in which a student can successfully fill in the required responses without mastering the objectives the programmer intended. It is not known whether the responses required in this program were overprompted, copy responses, or whether the basic fault lies in analysis of the task or in faulty testing for mastery of the rule before presenting examples. In either instance it is doubtful that the blackout technique (Holland, 49) would give the answer for approach to program improvement because it seems designed more to identify extraneous material in the program than to identify needed missing content. It would not be a new observation to say that many programmers are too intent upon the specific responses to be elicited by the program and are thus distracted from employing a sequence which is really logical based upon an analysis of the task.

Schutz, Baker and Gerlach in another study (98) set out to determine whether a concerted effort to carefully analyze tasks, carefully analyze entering prerequisites, and carefully sequence programed materials could lead to a really outstanding improvement in the success of individual students in the classroom. While they did achieve improvements in programs by the empirical techniques they used for capitalization, fractions and punctuation, and while the sequences of sub-skills in their programs differ significantly from sequences implied in textbooks, their overall conclusions and discussions reflect the practical difficulties encountered in conducting such work in the classroom instead of in the laboratory.

They say, "We seriously underestimated the power and complexity of the cumulative prior educational experiences of pupils and teachers alike. We are convinced now that these determinants...precluded the possibility of any...available instructional technique...being absolutely successful in the typical classroom under typical conditions, irrespective of the adequacy of the behavioral analysis upon which the instruction is based or the technical sophistication of instructional sequencing."

So much for the practical difficulty these researchers encountered in achieving the classroom payoff from empirical and behavioral procedures which are based on good theoretical and experimental grounds and supported by solid data when the learning is conducted for shorter time periods under the controlled conditions of the laboratory.

These investigators also reported trouble in maintaining behavior until the criterion test, and felt that there is often too much cueing in programs when long-term retention is of interest. In regard to spiraling sequences of instruction, they feel that it does not insult the good learner as much as does sequential redundancy, and that this fact along with the distributed practice occasioned may recommend spiraling when the overall sequence takes account of the structure of the task. This is perhaps another way of saying that even though the task is highly structured and thus should be carefully sequenced, one should not forget to inject review exercises after the sequences have individually been completed.

In conclusion, these investigators observed that issues which have received much prior attention, such as size of step, form of response and the like, are insignificant in their effects as compared to the impact of failure to follow the instructions required to implement in classroom practice a careful behavioral design and sequence. In short, they call for a shift in focus from the details of program writing to the matter of how to control teachers in such a way as to implement the advantages that programs potentially do offer.

After the Fact Measures of Patterns of Learning and Retention

Two reports by Odom (76, 77) are considered of great importance in that they represent an empirical after-the-fact analysis of the sequence

and grade placement of skills in the elementary school as these skills are now taught. Using the Odom Diagnostic Test of Capitalization and Punctuation Skills, a total of 1818 school children were tested in one class each from the 4th, 5th, and 6th grades in each of 19 schools in a large California district. The results of the actual achievement of the children at each grade level on each of 49 sub-tests of punctuation skill and 37 capitalization skills was compared with the grade placement of instructional materials in these skills in elementary school textbooks approved for use in the state of California.

The results showed that 10 capitalization skills are currently mastered at the grade level adopted for presentation by the texts, and 18 capitalization skills are mastered only at a later grade level than indicated by the materials. The comparable figures for punctuation are one skill mastered earlier than implied in the text, three skills mastered at the designated level, nine skills at a later grade level, and 24 skills not mastered anywhere to the criterion required in the elementary school grades. The criterion for passing or failing in each instance was determined on basis that 50% of the students in the total grade level would pass two out of four of the test items for each skill, but less than 75% of the students would pass three of the four items.

The statistical summaries of results by grade level for each of the 86 basic skills in capitalization and punctuation present an empirical picture of what is now being accomplished in terms of the sequence and grade placement of skills in adopted materials. Odom reports that teachers in general were not aware of the specific skill levels of individual students, and therefore four recommendations were made: (a) a diagnostic instrument should be used in the public schools on a periodic basis, (b) individuals should be given remedial practice in weak skills, (c) individual rates of progress should be allowed, based on the skill profile of the individual, (d) an experimental study of teaching methods should be made to identify the most effective instructional procedures before recommending the sequence and grade level place which ought to be taken as objectives in public education.

Transfer of Method vs. Transfer of Specific Skills

Smith and Moore (104) have discussed learning how to learn in the sense of general transfer such as in Harlow's experiments (45), in contrast to the specific transfer from one sub-task in a hierarchy to another. Thus they define learning set as a primitive form of transfer which occurs when the learner works on a series of complex but similar problems. They identify this as a transfer of method rather than as transfer of content (specific capabilities), as may be the case within a single task. They point out that no previous study has been reported to determine the effects of such general sets when a student is exposed to a long series of different programed instructional materials. They arranged it that 7th graders in various groups took varying numbers of programs up to seven in a counterbalanced order, and concluded that the number of programs taken was significantly related to the achievement level in later programs. Thus the experience of taking the first three

programs regardless of the order of the three or the content areas, led the Ss to make higher achievement scores on later programs as compared to groups who took the later programs first. However, they discuss a ceiling effect of a motivational and attitudinal nature which indicates that there is a limit to the proportion of time which a student should spend in learning by programed instruction.

Affective Implications of Sequencing

Fowler (27) has discussed the importance of sequencing in the context of education of disadvantaged children. While in general he cautions against the assumption that their problems are necessarily different from children found in normal groups, he does recommend, for every individual child, keeping on the alert for the need for individual psychological diagnosis and developmental stimulation in some ordered coherent manner. He refers to the frequency among disadvantaged children of defects in symbolic skills which are prerequisite for much school learning. He believes sequencing is important in steering the child through the intricate elements of the reality structure to be encountered in real life. He thus sees a developmental implication of sequencing which has long-term implications in blocks of instruction, compared to short-term frame-by-frame or topic-by-topic sequences in specific instructional materials. He recommends as does Bruner (17) that the child, disadvantaged or otherwise, first be familiarized with concrete stimuli such as familiar actual objects as a necessary prerequisite for gaining imagery and motor-produced sensations, the formation of which should precede establishment of corresponding symbols. The young child particularly needs a learning task involving active physical manipulation with concrete objects, followed by the formation of verbal and abstract equivalences of such stimuli. Certainly in the formation of concepts, as both Gagne and Bruner have discussed, exposure to an appropriate variety of physical objects with simple yes/no feedback may in many instances be more effective than any additional verbal instruction.

McNeil (68) reports the highly important finding that whereas boys are typically inferior to girls in learning to read when taught by human teachers, they made equal or superior progress as compared to girls in learning to read by programed instruction. Furthermore, when these same boys were given the next lessons by conventional teacher-conducted instruction, they reverted back to their inferior accomplishment relative to girls. This finding supports the contention of many educators that the education of boys suffers in many group situations, particularly if the teacher rewards and punishes boys on the basis of their social behavior rather than on the basis of their academic achievement. The informal, impersonal, and impartial reinforcement provided by programed instruction appears to provide an equal opportunity with girls in mastery of the reading task by removing the instruction from the social environment. This reduction in group interaction, both with the teacher and with class members, favors promotion of achievement by boys who probably, under stimulation of their peers and the teacher in the group situation, display aggression and independence which diverts them from the reading task and which leads to social disapproval.

Fitzgerald and Ausubel (26) have commented upon cognitive vs. affective factors in the learning and retention of controversial materials. In learning which represents understanding of "other side" arguments, the conceptual schema outstanding in the cognitive dimension of attitude of the learner is usually devoid of relevant subsuming ideas to which the new information can be related. That is, a person who holds view A can more easily assimilate additional information in agreement with view A than he can learn contrary information supporting view B. This is because the learner has more advance organizers and cognitive anchorage under which he can subsume additional learning of the familiar view A, in contrast to learning of information supporting view B.

Type VI: Experimenter-determined Sequencing of Frames of Programed Instruction

Scrambled vs. Logical Sequence

Most developers of programed instruction material would agree that they believe the effectiveness of their programs rests upon the following factors: (a) learning materials are carefully sequenced in an order which corresponds to the author's ideas of how learning should best progress, beginning "where the learner is at the outset" and taking him step by step into more complex learning; (b) the incremental steps are small and well prompted so that the learner can successfully take each step; (c) the learner makes active overt responses to each frame; and (d) the learner receives reinforcement for each correct response and corrective feedback for the infrequent wrong responses he makes.

In point (a) above, is seen the programmer's belief that sequence of materials is very important. Not all programmers, however, analyze tasks in the form of hierarchies, as Gagne and others have done. Some begin by writing "criterion frames," which are unprompted frames designed to test the student's mastery and immediate recall of a small sequence of prompted, "learning frames." Thus the programmer may work as though he thinks of the task as having a "vertical" structure, as defined in Chapter I, since he produces or arranges his criterion frames in "linear" order. Since the programmer's goal is to get the criterion frames in the "right" order to be used in a linear program, he may never deliberately attempt to find out if the task being expressed in the form of criterion frames is hierarchical, vertical, flat, or mixed (see definitions in Chapter I). But at any rate the programmer assumes, explicitly or not, that the total task mastery is accomplished by developing subordinate "behaviors" (competencies) in a definite order; consequently the learning frames also must be presented in the logical order, not a random order.

Once the criterion frames are arranged in the presumed "right" order, then the programmer writes teaching frames designed to enable the learner to succeed on each criterion frame (competency test). Thus the programmer assumes that sequence is important both for the

competencies tested by each criterion frame, and for the series of teaching frames within each competency.[1]

It is therefore apparent to programmers that they could test their assumptions about the importance of sequence by comparing their "logical" sequence with a "random" sequence, through experiments in which either the competencies (sub-sequences defining units of instruction) are "scrambled" in order, or the frames within a competency are scrambled, or both. It is possible to think of these three kinds of "scrambling" experiments with programed instruction as ways to study the "gross" or the "fine" structure of the program, or both.

As will be seen, since the results of these experiments are mostly negative, the no-difference findings are not regarded as indications that sequence is of no importance in "full length" programs. This is because the programs used in these experiments were short enough that the learner was probably able to recall materials out of sequence, and thus overcome the effects of "scrambling" the sequence. In some experiments the students were not prevented from looking backward and forward at will. Under these conditions, successful completion of a scrambled program is probably best interpreted as a sign of high motivation.

Gavurin and Donahue (38) compared the original sequence with a scrambled sequence of the first 29 items of the Holland-Skinner program (50). Although the scrambled group required more trials to learn to criterion, there was no difference a month later between the groups in retention level. It is quite clear that results of this experiment for 29 items cannot be generalized to longer programs because of the ease with which students might recall the substance of the 29 items and mentally try to remanipulate them to provide their own structures and to utilize information from an earlier frame in searching for an answer to a later frame.

Payne, Krathwohl, and Gordon (78) indicated that no one doubts that if an entire course were scrambled learning would be retarded. They point out therefore that the size of unit in which sequence is destroyed should be a factor in the results. They also indicated that type of task should be another factor in the importance of sequence, in that for learning to spell a list of words sequence should make no difference, while for learning a principle it should. They also assumed that degree of prior knowledge and hence different degrees of familiarity with the topic might enable Ss to adjust with diffferent degrees of effectiveness to a scrambled order within a block of instruction. In addition, the ability of a frame to stand alone because of its verbal instruction and the type of repetition in frames would also be factors affecting the

[1] To be sure not all programmers would describe what they actually do in these terms, depending on their theoretical leanings, but the practice is pretty much the same among experienced programmers. Some would say they are "shaping behavior" by "reinforcing responses" in such a way as to "bring responses under control of the desired discriminative stimuli."

the results of logical vs. scrambled orders.

These investigators arranged three programs in order of the judged degree of inter-item dependency. Three sections of the Krumboltz and Wiseman program (58) in educational measurements were judged as follows: The segment on percentile norms would contain the least inter-item dependency, the segment on validity and correlation would contain the most inter-item dependency, and the segment on age and grade norms would fall intermediate between the two. Three judges agreed upon this rank ordering of the three segments. The programs were then rearranged so that one group received all segments in the original linear or logical form. Other groups received some of the three segments in the logical form, and some in the scrambled form. No significant differences were found in criterion measures of learning and retention among the groups, but error rates on frames during learning were higher for the scrambled form of the program.

Roe, Case and Roe (86) utilized a 71-item program on elementary probability to study the logical-vs.-ordered-sequence question. They apparently conducted this study because of their interest in previous studies of response mode in which no differences were found among the groups giving overt responses, covert responses, constructed, or multiple-choice responses to instruction presented by machine, programed text, or lecture. These negative findings concerning program variables deemed significant by Skinner (102) led them to conclude that the only thing remaining that matters must be sequence of frames. Their study on scrambled vs. ordered sequence also resulted in no significant difference, and they suggested that results with other programs would be a function of the length of the program, the nature of the task, and individual differences in ability and prior knowledge. They added the observation that the scrambled order may keep students more alert because of the additional effort it requires to avoid errors. For example, the typographical errors which were in the programs were spotted by some members of the scrambled group but not by members of the logical sequence group.

Levin and Baker (60) changed the original sequence among frames in only the third unit of a four-unit program on geometry. They chose the third unit because it had been found to be the most difficult of the four units on earlier tryouts. They found no difference between the scrambled and the original sequence in acquisition, retention or transfer. They observed, however, that this total program was not an effective one for the groups employed, as the overall criterion scores were low. They also speculated that the results for the scrambled sequence might be interpreted as being due to the distributed practice thus resulting. In the discussion they suggested that most existing programs are probably neither completely structured nor without structure.

Hamilton (44), in three separate replications involving 68 fifth- and sixth-grade students, employed four versions of an autoinstructional program in music fundamentals. Teaching frames were presented in logical or random order in combination with two covert response conditions: specific (reading and thinking up answers to blanks in frames, with no

confirmation given), and nonspecific (reading only of frames without blanks). The hypothesis was that greater learning should be found for groups receiving the logical sequence and specific responding, and that less time should be required with logical sequence and nonspecific responding. However the results showed the specific response mode superior in both learning gain and learning time. The sequence variable interacted with response mode, the random nonspecific version being the most successful treatment and the random specific the least.

Hamilton was interested in four separate theoretical matters which programmers and experimenters have investigated, discussed, and debated. The first of these was logical sequencing; the second was reading only response vs. reading and covert explicit responding; the third was overt vs. covert responding (writing or thinking answers); and the fourth was confirmation vs. no confirmation. She theorized that in other experiments on scrambling the fact that confirmation was provided at the completion of each frame enabled the learner to overcome the effects of randomized sequence. In fact she interprets the learner's need to actively reorganize the randomly sequenced frames as a very effective form of active responding, if all the information is there (as in reading or confirmation conditions), if the concepts to be acquired are few and simple (as in the program she employed), or if the randomization among <u>frames</u> leaves the <u>sequence of units</u> in the logical order.

Hamilton recommended further examination of the sequencing variable <u>in combination with</u> other variables, and with longer programs and different subject matters.

<u>Interpretation</u>. While there appears to be general agreement that experiments with longer programs are needed to resolve the matter of sequencing as it has been investigated with programed instruction, there is also a need for more experiments which distinguish clearly between sequencing of <u>frames</u> and sequencing of <u>competencies</u>, e.g., the distinction between sequence of <u>units</u> or parcels of frames and single frames.

There is also a need to identify the assumed structure of the task, as Gagne has done in his experiments, because the importance of sequencing should vary with the nature of the structure of the task.

Further, there is a need to examine the way the sequence of instructional events is altered when frames or groups of frames are altered in sequence.

Next, there is a need to determine how the effects of scrambling within sequences (which alters instructional events) differ for the different types of learning, such as the types distinguished by Gagne (31).

Hamilton has shown the importance of the need to study the effect which altered sequence may have upon the way other variables in learning function in mastery of the task under logical and random sequencing. For example, confirmation may not be needed for well cued frames arranged

in logical sequence, but confirmation may be needed if the sequence is scrambled.

If this series of experiments indicates anything, it is probably the remarkably adaptive and organizing abilities of the highly motivated human learner, a medium of instruction often neglected (Melton, 69). The possibility that "scrambling" on a highly selective basis may actually improve learning has been shown by Hamilton, by keeping the learner actively searching for order from a partly unordered program. Such a program may require more time, but measures of retention and transfer should indicate whether it is time well spent. There are probably limits to which the student can be pressed in such altered programs. But in a sense, any prose passage which requires the reader to infer some of the ideas dealt with is this type of program. And some well written prose, in the sense of sequencing of words and sentences, can be dull and lead the student to stop reading, or it can be highly effective.

It may well be that studies previously conducted on degree of prompting should be reinterpreted and expanded in a search for a measure of prose which is effective through challenge rather than effective by shaping of behavior (Goldbeck and Briggs, 41). It is when the program is designed to "communicate the present culture" (establish the competencies in the hierarchy) that sequencing is to be arranged to establish the task performance. Programs designed to stimulate creativity and enquiry (Gagne, 30) may require characteristics quite different from programs to transmit the culture. Whether sequencing would play a very different part in such programs is unknown.

But to return to the facts, in the experiments cited above which employed short programs, there is no case in which scrambling, as a main effect, was significantly inferior to logical sequence. This finding must be accepted at face value for these programs, even though possible explanations for the reasons behind the results are very interesting. So students can and do perform as well under scrambling of the types used for the programs used. How long students would continue to study with such materials is unknown. Nor is it known how well they liked the scrambled materials, nor how many students noticed anything unusual about the programs. No experimenter reported a refusal to complete a program.

The one significant effect reported was the interaction of scrambling with response mode in the experiment by Hamilton, whose report provides food for thought even though her results were largely contrary to theoretical predictions.

In this series of studies, measurements of delayed recall were used in some experiments but not in all. It is unfortunate that more learning experiments do not include measures of retention and transfer. In the course of preparing earlier reviews of the literature on various topics, the reviewer has often noticed how many times results of experiments change, even to the extent of reversing direction, from immediate retention tests to delayed retention tests. Considering that education

is such a long process for our children today, and has been getting longer all along, and considering the findings by Gagne and Bassler (34) on the differential loss of retention for tasks and for particular subordinate competencies often needed for learning of tasks to come later, it would be important to discover conditions which provide effective long-term retention. If this were done, the time taken to complete training in a field of knowledge might be shortened. It also might be well if more experiments utilized effective programs. An extremely common finding is that criterion scores were low, whatever the diffferences found among independent variables used.

At least one comforting thought about this group of five experiments is that the results are in agreement. This is the exception rather than the rule in studies of individual characteristics of programed instruction. In almost all other variables studied (response mode, size of step, amount of prompting, linear-branching) some experiments can be found in which a given characteristic (e.g., overt responding) was superior to other response modes, some which find the characteristic inferior, and many which find no difference. Even worse, several reviewers, summarizing results for the same sets of experiments, draw different conclusions, some insisting that overt responding has been found essential to learning, some that it has not, and the remaining safely concluding that it depends on the type of task, or that the programs were too short to draw conclusions. One investigator (Holland, 49) has even invented a technique to distinguish a "program" from a "non-program," in order to decide for himself which reported experiments really pertain to issues in "programed instruction."

See closing chapter on Recommendations for Research for methodological suggestions on an approach to clarifying some of the above issues.

Programing Variables Associated with Frame Sequence

This group of experiments deals in order with three programing variables, as independent variables; in each experiment, the sequencing of content among frames was altered in order to test the independent variables of interest. This group of studies thus deals as directly with variables in the conditions of learning as with sequencing per se, but the two matters are so intertwined that it was thought relevant to review these experiments here. The headings, following, identify the independent variables in each study.

Homogeneous vs. Heterogeneous Practice Problems. Traub (109) employed a programed instruction sequence for the task of making graphical addition of positive and negative integers. The instructional program was designed on basis of a hierarchy of sub-tasks as analyzed by the programmer. The basic main line program was in programed instruction form which provided information on the sub-tasks presumed to be required. This included instruction on how to locate positive integers on the number line, how to draw arrows in solving problems, how to draw arrows of

specified length from the number line location of a specific integer, etc.

Three groups of Ss, after completing the main line program, then received practice in applying these procedures, each group having a different set of problems. One group received a variety of heterogeneous problems, another group received homogeneous problems, and a third group received a review of the main line program without taking practice problems.

The most effective learning was accomplished by the group which received heterogeneous problems following programed instruction. Even though this group found the heterogeneous problems more difficult as compared to the homogeneous problems and even though more errors were consequently made in solving these practice problems, fewer stereotype responses resulted and the Ss apparently gained more information from these difficult problems as indicated by the criterion measure. Apparently after learning the principles involved in such a procedure as graphical addition by use of the number line, it is best that the practice exercises involve a range of arithmetic values, a balanced mixture of positive and negative numbers, and hence a range of context within which the procedures are applied. One might say that learning to solve problems may be viewed as a two-stage process: whereas the initial programed instruction on the principles of graphical addition should be sequenced in accordance with the hierarchy and highly prompted so that the student receives much response guidance, the best way to gain skill in the procedure is to receive practice in a wide range of problems which, although difficult, provide good transfer to the performance of other problems so long as feedback is provided.

Clustering Frames for Related Concepts. Newton and Hickey (75) report a study of sequence effects in programed learning of a verbal concept. The study involved variations in the sequential placement of sub-concepts relative to other sub-concepts in mastery of the main concept. The major concept dealt with by the program was that of gross national product as being a function of three kinds of spending. The sub-concepts are each identified with one of the three kinds of spending, namely consumption spending, investment spending, and spending by the government. The program consisted basically of four clusters of frames, one cluster dealing with each of the three sub-concepts and the fourth cluster dealing with how the three sub-concepts are combined in the definition of gross national product.

Three variables were studied with reference to how frames in these four clusters were sequenced. One variable pertained to the order in which the sub-concepts were introduced, a second variable related to the linear distance in the program between frames which belong together as representing the same concept, and the third variable was whether rule frames or example frames appeared first.

The overall conclusion was that the position of frames makes a difference in speed of learning and amount of transfer, but the fastest learning is not necessarily conducive to the most transfer. There was

significantly poorer performance on the criterion test when learning of sub-concepts was most remote in the program from the integration of the concepts to form the major concept. The poorest performance resulted from a sequence in which sub-concepts were placed together at the beginning of the program with maximum separation from the definition of gross national product and when, on the other hand, gross national product was defined at the beginning with much separation from frames dealing with the two of the three sub-concepts which a triple interaction indicated were most important in transfer.

Generally speaking, this study may be interpreted to mean that there should be close temporal contiguity in rule frames which relate major concepts to sub-concepts so that not too many examples or details of each concept are brought up before the major relationships are made clear. This finding appears consistent with the idea of spiral programs and with the use of advance organizers.

Inductive vs. Deductive Sequencing. There has been some speculation among programmers and educators whether some learners may learn best by an inductively sequenced program while other learners may learn best by a deductively sequenced program. Also, the same question has been raised as to whether in general (for most students) an inductive or deductive sequencing is superior.

Evans, Homme, and Glaser (25), in their discussion of the RULEG system for the construction of programed learning sequences, appear to believe that rules should be stated first, followed by completed or partially completed examples, after which the student completes his own unprompted examples, thus demonstrating understanding of the rule. These authors also imply a hierarchical structure of tasks in that they point out that a statement which is a rule in one frame may become an example in a later frame, thus representing an increase in complexity of the learning accomplished. The main programing question which these authors discuss is not whether or not the rule should precede the example, which they assume to be the case, but rather just how quickly the program should go from prompted or partly prompted examples to unprompted examples (the question of how rapidly to fade or vanish the prompts).

In general, a deductive approach in which rules are presented before examples might be identified with a didactic or prompting approach to instruction, whereas the reverse case of presenting examples before the rules may be identified as an inductive or discovery method of learning, much talked about and much favored by many educators in the absence of much empirical verification. It would appear reasonable to hypothesize that a deductive approach would be best for learning discriminations, concepts and principles, but an inductive approach would be best for learning to solve problems (see the review by Briggs and Hamilton, 16). Two studies by Belcastro have a bearing on this issue.

In the first study Belcastro (9) compared four techniques of programing algebra. The four experimental programs were non-verbal inductive, verbal inductive, non-verbal deductive, and verbal deductive. The

four experimental programs differed significantly from the control group in mastery of axioms, with the verbal deductive being superior. There were not found significant differences among the other three conditions. In the second study (Belcastro, 10) the verbal inductive program presented applications or examples of yet unstated definitions or principles followed by the discovery and statement of the definition or principles. Words were used in the program both to give directions to the student and to provide essential background material. In the non-verbal inductive program the sequence was the same as in the verbal inductive, but words were used only to give directions, not to provide background material. In the two deductive sequences statements of definitions and principles were encountered first, followed by the requirement to respond to examples and applications. In the verbal condition, again, words were used both to give directions and to provide background, and in the non-verbal condition only to give directions.

Again the deductive method was found superior to the inductive, and the verbal deductive condition was superior to the non-verbal deductive condition. This superiority of the deductive method was more clear-cut for immediate retention than it was for transfer in subsequent applications of the algebra learned. In spite of these findings, the author hypothesized that a long-term retention study would reveal that the inductive method could yield better recall even though the Ss are more familiar with the deductive method. There appears to be little evidence from the data presented to support this hypothesis, so its statement may simply reflect the intuitive attractiveness which the discovery method holds for the experimenter.

Type VII: Learner-determined Branching in Autoinstruction

A word of explanation is needed for the phrase "learner-determined" in the above heading. The decision to use a branching rather than a linear program is a decision by the experimenter. It is also the experimenter who sets the rules which determine when the learner will "branch." But the actual fact of "branching" from the basic sequence results either from a choice made by the learner or from his objective performance. Hence the phrase "learner-determined."

Crowder (24) introduced a variation of programed instruction first identified in terms of the early product, a "scrambled book," and later identified as "intrinsic" programing, as machines replaced the books. The scrambled book is not read in page sequence, as are other books. The learner is given a paragraph or two of information, then a multiple-choice question. If the learner chooses the correct alternative, he is directed to a page on which he is congratulated for his perspicacity, reminded of why he is right, and then directed to another page having new material and another test question. If the learner chooses the wrong answer, he is directed to a page which tells him that he is wrong, often scolding him for his carelessness or inattention. He is also either given an explanation of why he was wrong, or he is given a hint and directed back to try again. Alternatively, he may be given a remedial sequence consisting of more simply written material, or he may be sent

back to read again some information previously responded to correctly, along with the friendly admonition to use that information to think out the answer to the present problem.

This "intrinsic" program thus was partly different in content for each learner, depending on the particular pages the learner was directed to as a function of his particular wrong choices, and the particular remedial instruction thought needed by the program in the case of a given kind of error. The particular sequence of pages used, then, also was a function of the learner's performance. Just as learners were "branched" to new remedial material or sent back to re-read old material when they failed, they could also "bypass" some material if they could answer a test question correctly first.

For some unknown reason, most of the programmers who write scrambled books have also adopted or tried to adopt Crowder's almost inimitable style of humoring the student along, with "gasps of horror" when he is careless (never stupid), and verbal smiles of congratulation when he shows keen insight. Mager, one of the more successful humorists in the business, sometimes capitalizes upon the known curiosity of the learner in "peeking" at the "wrong answer" page even when he knows the right answer, just to see what he would be told if he made that error. Therefore Mager sometimes deliberately asks a question nobody could miss except intentionally just to catch the learner up for "peeking" when he "cheats." It is a joy to read programs by these two master craftsmen, but it is not known whether their programs are good because of or in spite of their humor; nor is it entirely clear whether students learn more by branching programs or by linear ones. But at least many such programs avoid the dead seriousness of dead easy "copy frames."

Crowder has also developed electro-mechanical devices for presenting intrinsic programs (see review of devices by Briggs, 13). These do the "page turning" for the learner who needs only to push the coded buttons to be told what to do next. One such device was able to present short motion picture sequences as well as projecting pages of print. Later on it was found by others that branching programs lend themselves to computerization. Some of the earlier demonstration programs with computers did little different from scrambled books except to turn pages automatically, but this was an improvement as page-turning is a chore with scrambled books, and once one loses his place, he may never find it again.

Branching programs strike most people as eminently sensible because of the following considerations. Learners do exhibit a variety of entering relevant competencies, even on a fairly simple learning task, as Gagne and Paradise have shown. Learners do respond with differing patterns of error to a learning program. After the program has been taken, learners do show great variability in patterns of forgetting or remembering of subordinate competencies, as Gagne and Bassler have shown. Learners do differ in what they can infer from the same information in a program. Learners do miss the same test item for different reasons. For all these reasons, as well as others not stated, one could

reasonably expect that branching programs should be found to be superior to linear programs because they appear to provide the information the learner needs when he needs it, without his being bothered by information he doesn't need.

With all this in mind, it is difficult to understand why there is not more evidence showing branching superior to linear programs. But many experiments have failed to find such superiority for branching programs. In one report Coulson and Silberman (21), summarizing a study conducted in 1959, said that branching saved time but did not increase level of performance. They also described a study conducted in 1960 using better equipment, but still no significant differences were found in performance level. But in a subsequent experiment, described in the same report, an "optional" type of branching determined by the learner was superior to a fixed sequence.

Roe (85) also found a simple branching program not superior to a linear one. Coulson and others (22), employing groups matched on mental ability, had one group take a fixed sequence program in logic, consisting of 233 frames; each member of the other group received different numbers and sequences of items depending upon his performance. Branching was based both upon errors made and students' subjective evaluation of understanding. Scores for the branching group were superior (at the .05 level). The time difference was not significant. In this study some use was made of the computer capability to branch on basis of a series of relevant responses rather than upon a response to a single item.

It may be that some of these obtained results are the effects of making errors in the decision to branch ahead or to back-track, occasioned by three practices often employed in branching. First, the learner may be given a test question at the outset to see if he might skip a sequence of instruction. In the effort to pass the item, the student may implicitly recall some information wrongly, or make wrong inferences, and so make an error. The error response as well as the erroneous thinking leading up to the error may persist in spite of subsequent presentation of new instructional material designed to overcome the error. Second, a learner may be given an initial, very brief sequence of instruction, followed by a test question, with the plan to supply more information if the student fails. Again, the effects of being wrong may not thereby be corrected. Third is the case of the "false positive" in which a student responds correctly to a test item by chance or for the wrong reasons; on basis of the correct response he is permitted to bypass some of the regular material. It is discovered later that the bypassing was a mistake, but subsequent remedial instruction is often not then very effective (Campbell, 19).

In recent years, equipment for the student station for computer-assisted instruction has been greatly improved. Early stations were limited to typewriter input and output. With the use of the cathode ray tube and light gun, providing both printed and pictorial display, and with the use of film strips, movies, and tape recorders in conjunction with the station, a greater range of instructional stimuli can be

employed. If also accumulated historical data and series of responses made earlier in the program relevant to the next branching decision to be made can become the replacements for branching on basis of a single test item, the results for branching programs might surpass those for linear programs.

Suppes (106) expresses the view that while careful attention to sequencing of curriculum materials must be important, this importance nevertheless is far overshadowed by the extreme individual differences in capability for rate of progress. He illustrates his meaning by citing individual differences in learning rates for specific tasks, for drill material paced in rate by the computer. In his discussion he introduces the interesting idea that overt correction responses after feedback are needed for kindergarteners and first graders, but not for adults, to be sure that the feedback is effective.

In reporting a study consisting of Russian/English word pairs Suppes reported the rather unusual finding that the optimum size of unit for practice is typically found to be the largest of several units employed. Thus when lists of Russian/English words were presented in different sizes of units as an application of the whole/part question, college students and junior high school students did better practicing a list of 108 items than a list of 6, when time for mastery of the total list is the criterion. The results of course might not be the same if English/Russian were employed. But the general rule he applies is that when learning is faster than forgetting, the experimenter should show all 108 pairs on the first trial. But when forgetting is faster than learning, only one item should be presented; all exposure of a given pair should then be contiguous. The mathematics for in-between learning rates is given by Suppes. At any rate, Suppes related the results of his studies to the capabilities of computers to provide highly individualized adaptive programs which can be adjusted rapidly from moment to moment in accordance with the error rate, rate of progress, response latency, and size of unit with which the individual is successful in dealing. Surely, it would seem, a unit size based on the capabilities of the individual student, as can be managed by the computer, should be much more effective than any set unit size for an entire group of subjects at any age level.

It would appear, then, that in the future <u>branching in rate of presentation</u> should become a new variable in such research. At any rate, basing branching on something other than the results of a single test item appears desirable, due to (a) the unreliability of a single item, (b) to the danger of failing to assess a competency in a relevant way in view of the next steps to be taken, and (c) to the possible adverse effects of making errors. At the present time, then, it is not entirely clear whether the somewhat disappointing results from comparison of linear with branching programs is due to inadequacy in evaluating the student's competence as a basis for branching decisions, to adverse effects of making errors, to removal of needed redundancy from branching programs, to inadequate basic programing, or to upsetting certain sequencing effects which need to remain unchanged, as in the basic linear program.

It may be easier to achieve greater gains with computers for drill than for abstract, conceptual material. At least it remains for future research to demonstrate great superiority for branching in conceptual learning. It is possible that large-scale programs in multi-media course design, using computer-assisted instruction as one medium, such as is now planned by the Office of Education and the U. S. Naval Academy, will answer this question.

Type VIII: Experimenter-prepared Advance Organizers

In this section is discussed the learning of meaningful prose material, such as learned from reading textual material or listening to lectures. In the experiments reported here, it is the experimenter who requires the learner to read a certain kind of introductory material, called organizers, before reading the instructional materials. The experimenter prepares the reading material of both kinds, and he controls the amount of reading and study time devoted by the learner to each kind. In this sense, the experimenter controls sequence of study, e.g., study of the organizer before study of the instructional material. Just how the learner reads or "studies" each part for the allotted number of minutes is controlled by the learner, e.g., there is nothing to prevent his skipping around in the reading materials.

Ausubel calls learning from reading or lecturing "reception" learning (5). By reception learning he means what some mean by didactic instruction, in which the entire content of what is to be learned is presented to the learner in final form. Reception learning thus differs from discovery learning in which the learner generates some of the materials or ideas. The student's role in reception learning is as a receiver and organizer of information and ideas. The message transmitted can consist of <u>either rote material or meaningful material</u>. Rote material is material in which the content arrangement is arbitrary--that is, the learner can't figure out the reason for it--he just accepts it, like a name or telephone number. Reception learning of meaningful material is reception of non-arbitrary relationships which can be rationalized by or for the student. Reception learning thus must not be confused with rote learning, because it can represent highly organized ideas of a meaningful nature, which the learner does receive, but he also organizes in terms of his own cognitive structure. It is because the essential ideational content, even though meaningful, is <u>given</u> to the learner in the materials he reads that this is classified as one form of reception learning.

In an early experiment, Ausubel (4) found support for his hypothesis concerning the role of "advance organizers" in the learning of new material. He theorized that in learning of unfamiliar but meaningful material, if the learner reads a short introductory statement first, which contains the most abstract or general statements possible about the topic, this organizer gives facilitation in advance to learning of the detailed lesson. This is because the cognitive structure of the learner is hierarchically organized, so that grasping first an inclusive concept facilitates the subsumption of the less inclusive concepts in

the learning material. The new matter becomes incorporated into the learner's cognitive structure if it is subsumable under existing general concepts already there. The older general idea provides anchorage of the new specific ideas as well as subsumption. Thus an "advance organizer" prepares the way for learning of detailed new material.

In such an experiment, it is important to show that the organizer does not really constitute additional learning time for materials measured on the criterion test covering the lesson itself. Otherwise, just the added study time for the organizer might account for the results. Ausubel therefore took two precautions: (a) the organizer was studied by an independent group which did not receive the lesson itself; these people scored little above chance on the lesson criterion; (b) a control group studied an introductory passage of equal length to the organizer, but having content less relevant to the lesson than does the organizer; this permits assurance that it is the <u>relevance</u> of the organizer, not its mere presence for study before the lesson, which facilitates learning.

In another experiment Ausubel and Fitzgerald (6) employed organizers to study the case of two successive lessons which are logically related, so that learning of the first should facilitate learning of the second. This is a typical educational situation, in which new learning builds upon old. This is called sequential learning by these investigators. Under these conditions, the learner's ability to discriminate between the two sets of concepts in the two lessons has important implications for learning and retention. This discriminability is partly a function of the clarity and stability of the previously learned concepts. When such discriminability is low due to inadequate prior learning, new learning and retention can be enhanced by use of "comparative organizers," to help the learner detect the similarities and differences between the old learning and the new. Results of marginal statistical significance were obtained in support of this hypothesis.

In a later study, Ausubel and Fitzgerald (7) tested the hypothesis that "expository organizers" enhance learning and retention of new materials for learners of low verbal ability. They used an organizer to be studied for six minutes, followed by two successive lessons on the physiology of pubescence. The results indicated that the organizer facilitated learning and retention by low ability learners of the first lesson at nearly the .01 level of confidence. Also, learning of Part I facilitated learning of Part II.

In total these experiments confirmed the general theory concerning organizers, but the differences cited often were significant only at the .05 or .07 level of confidence, and must be considered marginal.

Scandura and Wells (96) have discussed advance organizers as a concept somewhat similar to enrichment. They indicate that in high school texts in mathematics, historical material is often used in introductions as motivating materials. An alternate kind of introduction consists of examples of how principles work before teaching of the principles themselves is undertaken. In college texts, strangely enough, the

historical and illustrative materials come last, as though students at this level do not need the motivational effects such materials are presumed to bring about.

Scandura and Wells employed two kinds of introductory materials for two tasks. Group G material (approximately 1000 words) was a brief introduction to sets, the definition of binary operation, and the abstract definition of a mathematical group in terms of the five group axioms. Group T materials (also 1000 words) were on a brief introduction to topology, the definitions of a network, of a closed network, of traversing a network, and Euler's four rules for traversing a network in a single journey.

The two types of introduction were historical (H) and organizers (O). The historical introductions on the two topics were also about 1000 words, and were used as controls. The organizers used in the experimental groups were mathematical games.

Results indicated that the organizers were superior to the historical introductions, and this difference was greater for topology than for G materials. The reviewer believes that a better evaluation of the effect of the organizer would have been possible if another group had received only the materials, without either form of introduction, but with study time equated with groups which did receive introductions.

While the terminology is different, the reader may note some similarities between Ausubel's concepts and those of Gagne concerning the hierarchical nature of learning. While Ausubel speaks of development of a cognitive structure while Gagne speaks of competency acquisition, the implications for sequencing are not unlike. It is interesting to compare the two approaches in relation to two different ways of stating the learning objectives.

Gagne's research in mathematics deals with a subject-matter area widely believed to be "highly structured." In Gagne's work, this structure is a behavioral one, as stated in form of competency on classes of task performance. To do this kind of research, objectives must be stated in behavioral terms.

Topics such as history normally are not stated in behavioral terms. Such educational areas possibly lend themselves to Ausubel's approach.

If one accepts the fact, fortunate or unfortunate, that much of school learning does consist of reading or hearing prose material in order to pass tests of comprehension and recall of the principal substance of the material, then one can work with Ausubel's theory. It might be better for education, however, if course designers were required to state behavioral objectives; then analyses could be done and checked by sequencing experiments to determine whether those objectives were structured hierarchically or not.

In closing this comment on the similarities and differences in the

theoretical concepts, approaches, and technical terms as employed by Ausubel and Gagne, the following discussion by Ausubel may sound quite familiar to one who has read Gagne's writings. Ausubel says (5, p.86):

> "In sequential school learning, knowledge of earlier-appearing material in the sequence plays much the same role as an organizer in relation to later-appearing materials in the sequence.
>
> "Hence new material in the sequence should never be introduced until all previous steps are thoroughly mastered. Perhaps the chief pedagogic advantage of the teaching machine lies in the ability to control this crucial variable in sequential learning.
>
> "Most complex tasks, particularly those which are sequential in nature, can be analyzed into a hierarchy of component learning sets or units. The rate of learning these units and the extent to which they can be recalled are more highly related to final achievement on the learning task than are general intellectual ability or more specific cognitive aptitudes. Serious breakdowns in learning can often be attributed to the inadvertent omission of a logically essential component unit from the total task or to its inadequate integration with other components.
>
> "Heterogeneous presentation of stimulus material that does not provide sufficient repetition to allow for mastery is not only less effective than homogeneous presentation in learning a principle, but also does not facilitate the learning of a reversal principle during the transfer period. Reversal learning in rats is similarly facilitated when the first of two discrimination problems is overlearned."

Chapter V: Sequences Pre-planned by the Experimenter to Test Hypotheses about Effective Characteristics of Learning Programs

In this chapter there are reviewed experiments which for the most part employed artificial, contrived, "laboratory tasks," in which sequence was varied primarily as a way to implement some other independent variable whose effect upon learning the experimenter wished to evaluate. In these experiments sequencing *per se* thus is of only indirect or secondary interest as compared to interest in the independent variable.

The experiments reviewed are classified in accordance with the type of learning involved in the experimental task, as identified either by the author of the experimental report or by this reviewer. This creates some difficulties due to the lack of standardization in the definitions of "problem solving," "principles," and "concepts." Some writers simply group these as "meaningful learning," but this is not very satisfactory either. Other writers, by the way they identify the type of learning represented by the experimental tasks they have employed, show wide divergence in the usage of these three terms. Thus the organization of this chapter admittedly is unsatisfactory as to distinctions made in classifying experimental tasks. The added label of Type IX, of course, is an arbitrary one used for the purpose of this report, as are the designations of the previous eight "types" of sequencing studies.

Problem Solving

Scandura's Research in Problem Solving

Scandura, with various associates, has conducted a series of studies opening up what he refers to as a new discipline of "psycho-mathematics," which he sees as a parallel to psycho-linguistics.

Scandura's views on a theory of teaching were discussed in Chapter I. Basic to his views is the conception that the principle, rather than the association, is the basic unit in meaningful learning (Scandura, 91). Also, he introduces a new language of communication about learning which he calls a scientific set-function language (SFL). This language deals with (a) objectives, (b) information presentation in learning, (c) what is learned, and (d) input variables (prior knowledge).

Applying SFL to problem-solving as encountered in mathematics and in tasks in the learning laboratory, he deals with four components in problem-solving activities. The language for these components, he feels, had to be developed because he views mediational and operant theories as too cumbersome, and cognitive theory as not precise enough.

In one experiment (Scandura, 95) during learning *S* encounters four pair of stimuli to be associated and used in a transfer task:

1.  - Black $S_1 - R_1$

2.  - Large $S_2 - R_2$

3.  - White $S_3 - R_3$

4.  - Small $S_4 - R_4$

Scandura asks, when these four stimuli pairs are learned, what has the learner learned? He says it is two principles, not four associations, and the principles are: "If triangle, then color; if circle, then size."

These principles, if learned, usually are employed in responding to the first transfer task, if at all. Thus SFL denies the primacy of the association.

The response given on the first transfer task enables a prediction to be made of the response on a second one.

Transfer Task 1 was a small black triangle and a large black circle. Test 2 was a large white triangle and a small white circle. Fifteen Ss were trained to an overlearning criterion on the original task. All but three gave the correct responses to both transfer tasks.

This result, Scandura says, cannot be interpreted as stimulus generalization, so he believes the Ss discovered the two underlying principles while learning the four original S-R pairs, and they later applied these principles to the transfer task.

Since an association refers to a one-to-one relationship, and a concept is a many-to-one relationship (the response red to all red stimuli), the latter relationship is the one pertaining to meaningful learning. "If triangle, then color" is also a relationship not based on a single S-R association, and it is what was learned--a principle.

Scandura assumes that four elements, I, D, O, R, are needed to specify a principle. "The stimulus properties in the set 'I' tell when the rule (D, O, R) is to be used; those in 'D' tell which properties determine the response; and the combining operation 'O' tells how the response properties 'R' may be derived from those in 'D'. In the preceding task, 'I' includes colored triangle, 'D' color, 'O' color naming, and 'R' the color names" (95, p.276).

Scandura employed the above language and theory regarding the nature of learning in tasks which could be labeled as either learning of principles or problem solving. He employed a task called "Abstract Card Tasks" in a series of laboratory-type experiments (90, 91, 92, 93). The results suggested that unless S can show in his performance that he can make use of terms or statements presented as a part of the problem,

he cannot understand descriptions of a superordinate topic. Also, it cannot be assumed that providing S with definitions of prerequisite verbalizations is sufficient; practice in using the referents in the prerequisite material improves learning on the superordinate criterion.

This series of studies may generally be interpreted as supporting some of the points in Gagne's rationale (29) of the significance of hierarchies for the sequencing of various units of instruction.

The series of studies by Scandura and others is not reviewed here in the detail it may deserve because the reviewer feels that, taking the series as a whole, the contributions made in defining a language, a methodology, and a prototype task for the study of problem solving are to date greater contributions than are the results so far reported. However, as this work continues, and becomes more heavily oriented to use of classroom materials, as begun in one study already reported (93), the results in the near future may clarify the educational significance of the overall results.

Other Studies of Problem Solving

Morrisett and Hovland (73) report a comparison of three varieties of training in problem solving. During the experimental trials subjects operated keys in a multiple-choice response mechanism to record their interpretations of the series of stimuli. The materials required Ss to discover that keys 1 and 2 were to be pressed representing an up/down arrangement of stimuli, keys 3 and 4 were to be pressed to represent left/right pairs of stimuli, keys 1 and 3 were to be identified with pairs of circles, and keys 2 and 4 with pairs of triangles. The task was to discover these relationships between stimulus characteristics and correct key responses and then acquire proficiency in responding to repeated trials of the stimulus elements. This kind of experimental situation required S to discover what rule the experimenter was using in determining how reinforcement was provided, e.g., the signal "right" or "wrong."

One group of Ss received 48 presentations of the same 4 items as the practice condition, a second group received 24 presentations of two sets of 4 items, and a third group received 16 presentations of three groups of 4 items. This arrangement meant that all groups responded to the same number of stimuli, and hence is somewhat comparable to the study by Traub (109) in that it represents a different range of stimuli during practice in preparation for a transfer test. The results of this study on the transfer test indicated that group 3 performed the best. The conclusion is that practice on multiple problems involving similar principles yields better transfer than practice in equal amounts on a smaller range of problems.

Short and Haughey (100) compared the effectiveness of a multiple-concept strategy in sequencing with a single-concept strategy for two lessons, science and language arts. In the multiple-concept sequencing,

simple descriptions of several related concepts are given first, followed by increasingly complex material about these same concepts. (This is similar to the spiral program, discussed earlier.) In the single-concept strategy, one concept is explained in a series of descriptions from simple to complex until the entire lesson on the concept has been presented. Then another concept is taken up, following a similar procedure.

Two versions were made of each of the two experimental lessons simply by rearranging the sequence of frames; thus content was held constant. In the abstract source necessarily cited here, no further information is given as to content of the lessons, but both were referred to as multiple-discrimination tasks. The results for the science lesson were superior for the multiple-concept sequence at the 1% level of confidence. The comparison of the two sequences for language arts did not yield significant differences. It is possible that if the materials were available for inspection, it might be suggested that the students could verbalize their responses more readily for the language arts material in such a way as to enable them to overcome some of the effects of the sequences as presented (see results for paired-associate materials presented later in this chapter).

Johnson, Parrott, and Stratton (53) raise the question whether one solution or many should be produced in a problem-solving situation before the S indicates his final choice of the alternate solutions he produced for consideration. A total of 600 Ss were assigned to one of two conditions (one solution or many solutions) for five kinds of tasks (plot titles; table titles; conclusions; sentences; cartoons). Taking as an example the task of supplying a title for the plot of a short story, the question is whether to train the S to (a) construct one title he thinks is appropriate, or (b) to list several, or (c) to list several and then indicate which one he considered best. These three conditions were employed, using judges as "blind" raters of the solutions; thus the quality of solutions for the three groups could be compared.

The results failed to show significant differences among the three groups. The preferred (single) solutions offered by group (a) were no poorer than the preferred solutions of group (c) Ss, who produced several solutions and then indicated which they thought was best. The authors concluded that in order to improve thinking, the next step would be to train Ss how to select solutions which have been produced.

Concept Acquisition

The experimental literature contains many reports concerning concept formation. The kinds of concepts to be acquired in these studies often do fit the definition of concepts as they are important in educational learning, but the experimental procedures employed and the particular kinds of stimuli presented in connection with concept formation in the laboratory are quite at variance with how concepts are dealt with in education. While these laboratory studies are instructive in regard to how people acquire concepts under the experimental conditions presented in the laboratory studies, they are not very instructive insofar

as the teaching procedures one would use to establish concepts more efficiently than they are established in the laboratory.

In short, a curious characteristic of concept formation as it has been studied in the laboratory is that the experimenter deliberately hides from S the cues and prompts which usually a teacher or programmer give to the students to facilitate their learning. Thus concept formation as studied in the laboratory requires a discovery type of learning in which knowledge already available to the experimenter is deliberately withheld from the learner. Thus S has to "guess what rule the experimenter is following in his administration of feedback to S after each response." This is similar to the game of Twenty Questions, or guessing what the experimenter has in mind as a controller of his behavior. Therefore prompted learning of concepts as taught in the classroom involves different behavior for both teacher and student than the same concept presented as a problem to solve in the laboratory. For these reasons no attempt is made here to review a representative sample of the experiments in concept formation. However, a few studies are cited to illustrate the findings for whatever value they may have when one considers their implications for the design of instruction.

Neisser and Weene (74) have pointed out that whether learning takes place in the laboratory or in the classroom, new ideas are indeed built upon former ideas and learnings. They indicate that much of cognitive activity is hierarchically organized. For example, in a laboratory experiment, for S to acquire the concept "three borders," he must first be able to identify a border when he sees one, such as in the stimulus objects presented, and to discriminate borders from non-borders. He must also be able to count and to make use of the yes/no feedback provided by the experimenter following each response S makes to the series of stimuli presented. These investigators thus have spoken of concept acquisition as involving three levels of learning within the hierarchy of learning needed to acquire certain concepts.

The simplest level involves experimental problems in which S must first react simply to the presence or absence of characteristic A of the stimulus objects. The second level in the hierarchy requires the discovery that the experimenter's rule is based on the property that either A or B must be present, but not necessarily both, as characteristics in the stimulus objects. The third level in the hierarchy would involve perception of more complex rules such as either A or B must be present but not ever both together.

Often the tasks used in the laboratory require identification of such hierarchical properties by use of the feedback provided by the experimenter; these are referred to as disjunctive properties. There is often a conjunctive property of the problem which requires S to combine some of the properties separately perceived earlier. They hypothesize that because of this type of hierarchy, concepts at level 3 are more complex and more difficult to attain than those at level 1, when difficulty is defined as time or number of trials to reach criterion.

(Difficulty in the context of educational sequencing is perhaps not the best description of this situation. A better description would be that the learner must first acquire each of the disjunctive properties and then combine them in some fashion to acquire the concept. The instruction should therefore be programed in that sequence. In an efficient teaching sequence, then, these separate achievements would be arranged for in order, thus making the mastery of the concept itself not a matter of difficulty but a matter of prerequisite learning.)

It is perhaps unsurprising that Neisser and Weene found evidence to support their hypothesis that it takes longer to learn concepts at level 3 than at level 1.

Haygood and Bourne (46) also have experimented with the three levels of concepts described by Neisser and Weene. They also mention that in a task analysis of the structure of concepts there are two main features, namely acquiring the relevant attributes or stimulus characteristics making up the concept, and the conceptual rule by which the attributes are combined to form the concept. Examples of attributes could be expressed in these experiments by the words "red" and "triangle," and the conceptual rule to be acquired through feedback after each response might be "all red triangles." These investigators recommended the separating out of these two components of learning in the further experimental study of concept formation so that more explicit information is gained in addition to the time and error data typically yielded.

Hunt and Hovland (52) indicated that the S in a visual concept formation task could derive a correct concept from the instances presented on the basis of either conjunctive, disjunctive, or relational characteristics. They used both conjunctive and relational situations more often than disjunctive. These experimenters found no interaction of the results with particular aspects of stimuli such as color, number or type of figure. They plan to use the results of their study for computer simulation of human concept learning.

Amster (2) referred to earlier findings by Podell (not otherwise cited here) that for college Ss deductive behavior occurs during intentional concept learning, while associative responding predominates during incidental concept learning. On a logical basis they theorized that when Ss employ deductive strategies a large variety of stimuli should provide an advantage over a small variety because the large variety permits false hypotheses to be rejected in fewer trials than a small variety does. Podell did find that under intentional instructions large variety was more effective in facilitating concept formation than was small variety, but when incidental instructions were used the small variety was more effective.

In Amster's study with five and ten-year-old children no variety effect was found, which was interpreted to mean that five-year-olds are not skilled in deductive strategies and thus are forced to respond in terms of associations. This observation may have importance in suggesting that the whole matter of sequencing of instruction would show marked

interactions with age levels because of the different cognitive manipulations which appear in the human repertoire at different ages in experiential development. Thus age differences may interact with instructional procedure due either to developmental or experiential factors.

Lee (59) observed that previous studies have shown bi-directional concepts to be the most difficult type to learn, followed by conditional, disjunctive, and conjunctive concepts in a descending order. However, this observed fact does not necessarily indicate that the behavioral concept belongs to a higher level than others in the hierarchy because the particular pre-experimental capabilities the Ss brought to the experiment as well as design flaws could interact with the findings. To illustrate the significance of this, Lee conducted pre-training sessions with different groups, employing pre-training in three rules, two rules, one rule, and no rule. The hypothesis was that if any of the three lower level concepts is not acquired, optimal transfer of learning to the biconditional concept will not be ensured. The greater the number of pretraining problems, he hypothesized, the more efficient the acquisition of the concept. The results were that the acquisition of all three lower level concepts facilitated the acquisition of the higher level concept, as compared to the prior learning of two or fewer lower level concepts, thus supporting the hypothesis of the learning hierarchy for concepts. In a second experiment he tested the hypothesis that there are two processes in concept formation, namely attribute detection and coding and rule formulation. The results supported the hypothesis but also indicated that for the materials used attribute coding accounted for the greater portion of the transfer.

Paired Associate Learning

As observed above for concept formation, laboratory studies of paired associate learning also are not entirely instructive in suggesting how instances of such instruction in education might most efficiently be arranged. Rather, the tradition has been that a fixed series of stimulus pairs are presented in order, trial after trial, until the learner somehow learns to repeat the total task. The learner is aided only by confirmation or correction after each overt response he makes or, when a response is not made within a two-second interval in the anticipation method, by the prompting for the following trial represented in the exposure of the response term at the end of the interval. Thus the sequence of events in the anticipation method of laboratory research is designed more to standardize conditions than to promote efficient learning. It is now known that prompting before S attempts to respond is often more effective for rote learning than is confirmation (Briggs and Hamilton, 16) and that selective item dropout is preferred to repeated presentation of the whole series on each trial, and that the response interval should be decreased for later trials (see Briggs, in (61), chapter 25).

Prompting thus has a somewhat unique temporal significance in the anticipation method in that it appears after the response attempt, while in programed learning the prompt is given before the student is expected

to attempt to make the response. Thus the detailed sequence of events involved in presentation of a single stimulus pair is quite different for prompting in the case of the anticipation method than for effective study techniques or for programed instruction. Confirmation, on the other hand, in both types of situations is typically given rather immediately after each overt response made, thus giving knowledge of right or wrong to S immediately after his response.

Other than providing prompting after a failure to respond correctly on time and confirmation after a timely response, most laboratory studies have left the learning of the task up to the learner's own effort and have not seriously considered, as have programmers, how to make the instruction more efficient by giving the learner more appropriate help. Even so, some points of interest may be derived from a highly selective few studies chosen from among the immense literature in paired associate learning.

Gagne (28) wished to determine the effect of sequence of presentation on the learning of 12 nonsense form - nonsense syllable paired associates, containing four dissimilar sets of three highly similar forms. Four sequences were arranged as follows:

1. The most similar forms were presented together in sequence.

2. Each member of similar form was given maximum separation from other members by interspersal with dissimilar forms.

3. Sub-groupings were sequenced in such a way that there were four sub-groups of three similar forms, the members of each sub-group appearing together in sequence.

4. Random arrangement of the 12 pairs.

Each group was given 14 trials, and the criterion measures consisted of number of correct responses each trial. A significant difference was found among the performances of the groups on the last two of the 14 trials. Condition 1 was found to provide the most effective learning. The interpretation was that when similar items are grouped together confusion is high at the outset, but it declines as the number of trials progresses. Ss made more overt responses, including wrong responses, because they would succeed in anticipating within the time interval required because they would think they were correct even though they were wrong when they were confused among similar items. They thus received more instances of reinforcement and were able to correct their errors and learn better in future trials.

The application suggested by Gagne for these findings was that in such problems as aircraft recognition and code reception one should present together the groups of similar items and arrange the instruction to give emphasis on distinguishing among the highly similar items. Such additional prompting as compared to the experimental situation could accelerate overcoming the initial confusion and lead to more effective

learning than if dissimilar items appear in sequence, thus requiring less precise discrimination.

Rotberg and Woolman (87) also studied paired associate learning as a function of the grouping of stimuli. They referred to Gagne's study (28) which used stimuli subject to verbalization, and pointed out that the discrepancy between Gagne's findings and those of Rothkopf (88) who used Morse code stimuli were accountable for by the difference in the amenability of the stimuli to verbalization. Thus whereas Gagne found, for the stimuli subject to verbalization, that grouping was best by similar items, Rothkopf found that grouping was best by dissimilar items. The comment relating to verbalization, however, pertained only to Gagne's stimulus terms, not to the response terms. The stimulus terms were nonsense syllables which could be verbalized by spelling them, whereas the response terms were geometric forms which could not readily be verbalized. Rotberg and Woolman therefore employed stimulus terms and response terms both of which were subject to verbalization. Grouping by similar and dissimilar was employed as in the following example.

Similar	Dissimilar
VKIW - yard	VKIW - yard
VKIH - star	LDAK - mint

The results were that learning was better when grouped by similar stimuli rather than by dissimilar stimuli with reference to the stimulus terms. This finding did not occur when the similarity or dissimilarity in grouping was based on response terms. The interpretation given was that the similar grouping provides better opportunity both for discrimination and for coding.

A summary of the many conditions other than sequencing or grouping of items in paired associate studies which have been studied in experiments has been presented in the well-known text by McGeoch and Irion (67, pp.499-507).

Serial Verbal Learning

Learning of verbal lists of items (words, phrases, nonsense syllables, and the like), along with the learning of pairs of such items (as summarized immediately above), taken together represent the kind of experimental task most exhaustively studied in learning laboratories over the past century. Literally thousands of experiments have been conducted using these forms of rote learning materials. Such experiments have dealt with numerous variables of which rate of learning is a function, such as the meaningfulness of the material, the degree of similarity among items or lists of items, massed vs. distributed practice, rate of presentation, length of list, and many others. It is the research on these variables for verbal serial learning and paired-associate learning which represents the major kind of research treated by McGeoch and Irion, referenced above.

More recent work along this line has explored some added variables in short-term memory, distributed practice, and other issues of theoretical importance. Almost all these issues involve sequencing variations. However since this area is so vast, it was decided to omit it from this review rather than to give it superficial treatment.

Part-whole Learning

This is another issue involving sequencing. The question arises whether an entire memorization task should be practiced, or only parts of it at a time until all parts have been mastered. Materials used in study of this issue have included poems, serial and paired associate lists, prose, lists of numbers, and motor performance tasks of various sorts.

Part-whole issues arise in learning of many real-life jobs or tasks, such as flying an airplane. In this case, the issue also arises as to whether to practice the whole task or separate parts of it in a real airplane, or with a simulator or part-task trainer. The answers in most cases relate to degree of relevant prior skills mastered, the relative degree of time-sharing or overlapping in time in the performance of different parts of the task, the trade-off of skill acquisition rate with safety and cost, and other relevant factors. Some of the military research on this problem has been reviewed by Adams (1).

Another type of task of military and industrial importance is represented in fixed procedures in performing manual and intellectual tasks, such as assembling a carburetor or finding trouble sources in the functioning of electronic equipment. Much research has been conducted in how these tasks should be analyzed into parts, and how the sequencing of instruction on the parts should be arranged. Involved and interacting with sequences of parts of the instruction is the issue of the sequencing of the media employed; that is, if a film is used to show the steps in the task, how long a sequence should the learner observe before he practices the steps on the equipment or mentally rehearses them? In the book edited by Lumsdaine (61) much attention is given to these matters, especially in the chapters by Sheffield and Maccoby (99) and by Margolius, Sheffield and Maccoby (65). Research such as this also is dealt with in the text edited by Edwin A. Fleishman (Studies in Personal and Industrial Psychology (Revised Edition), Homewood, Ill.: The Dorsey Press, 1967) and in the text edited by Robert M. Gagne (Psychological Principles in System Development, New York: Holt, Rinehart and Winston, 1962).

Again, these research areas are vast in themselves, and are mentioned here to acknowledge their relevance to the topic of this report. Like the topics of serial and paired-associate learning, however, they must be left here with this brief mention.

Sequential Learning of Motor Chains

While this aspect of sequencing of a particular kind of task is closely related to all the other issues mentioned in this chapter, and tasks described by this section heading overlap with tasks mentioned above, one specific is reported here because of recent theoretic interest relating to programed instruction.

First it should be said that learning of sequential tasks, or chains, may be either verbal tasks, or motor performances which are done in a chain-like series but are supported by prior learning of concepts, fixed verbal sequences, and other kinds of learning. Thus this kind of task which appears to consist of motor performance as viewed from the outside, actually viewed from the learner's history of competency-development may have called for prior learning of all other types mentioned in this report. Thus this type of task, seen as a whole, is a very complex hierarchy, involving not only many subordinate components but also many kinds of learning (called elsewhere by the reviewer "mixed" tasks). In such tasks, it is not the end motor performance which is hard to learn, but rather the real learning task is in mastery of each knowledge element and in knowing or remembering what to do next.

Gilbert (39), employing a type of task analysis and approach to learning he calls "mathetics," has proposed that the way to learn a chain is to "learn it backwards." (Reviewer's quotation marks.) The rationale is based on reinforcement theory, and will not be detailed here. Briefly, however, Gilbert proposes breaking up the task performance into units defined by the "operant span," which, to put it inexactly, is similar to "memory span for a series performed in reverse." According to Gilbert, the last span is to be learned first, then the next-to-last, etc., until the whole task can be performed in the "natural" sequence, from beginning to end.

Cox and Boren (23) refer to a study by Slack (103)[1] designed to test Gilbert's hypothesis. Both these experimenters appear to be concerned about the applicability to human learning of a theoretical model of backward chaining in learning a sequential task based upon observation of the behavior of animals. However, animals cannot be instructed effectively by verbal statements, and they presumably do not have the ability of the human to keep a temporally remote goal in mind from the very beginning of a lengthy task. Humans, unlike other animals, are capable of anticipating reinforcement to come at the successful conclusion of a lengthy task. Therefore it appears unsafe to depend upon extrapolations of theory based upon animal behavior when predicting and designing sequences of instruction for the human learner.

As a test of the hypothesis that such extrapolation is undesirable,

---

1 The report by Slack was not available to the reviewer. The comments here about that study are based on the report by Cox and Boren, which was available to the reviewer.

Cox and Boren trained 30 men to perform a 72-action procedure on Nike Hercules equipment. Ten men were trained in the reverse chronological order recommended by Gilbert, the tasks being broken up into seven operant spans as Gilbert recommends. Ten other men were trained in chronological order by a part method involving the same seven operant spans as parts. The third group of ten men practiced the task by the whole method in chronological order without grouping of tasks into operant spans but in terms of the maximum sequences they could perform on each trial before requiring correction. All three groups continued practice until they could perform one perfect trial without prompting. There was no significant difference in effectiveness of the learning among the three groups, thus not confirming Gilbert's hypothesis but agreeing with the findings of Slack (103), who employed poetry as the experimental task.

Chapter VI: Discussion, Conclusions, and Recommendations

Discussion

The research studies reviewed in this report employed mainly materials (programed instruction for the most part) designed to promote attainment of mastery of skills in core subjects (like mathematics, science, theory of musical notation, etc.) for which clear-cut behavioral objectives and criteria can be developed readily. Thus the structure of the learning was somewhat easier to come to grips with than would be the case, for example, in social science or in materials designed to foster creativity. Consequently the knowledge available from the research reviewed pertains to "teaching what the teacher, programmer, or experimenter already knows and can do and can describe in behavioral terms," not to the larger and more intangible objective of "preparing students to make discoveries not yet made by anybody else."

With the above limitation in mind, it may be said that a number of studies have supported the view that many courses or portions of them, when analyzed carefully, display an inferred hierarchical structure whose validity can be supported by comparing sequences of instruction designed to follow the hierarchy with some other strategy of sequencing. Several experiments provided various kinds of evidence of the validity of the hierarchy, because when learning was sequenced accordingly, learning progressed better than under other sequencing procedures.

Unfortunately, practical constraints have led most experimenters to employ learning tasks requiring only a few minutes or a few hours of study time. While positive results were often found even with these short learning programs, many "no difference" findings from other experiments had to be attributed by the investigators to factors related to the brevity of the programs. There exists the dilemma, then, that few studies of year-long experimental learning programs are conducted on the topic of this report, because of cost and other practical constraints, and because of the difficulty of controlling and evaluating the outcomes of such experiments. Thus better control is often exercised over shorter learning periods, but if the periods are too short, students can mentally correct for sequence factors in the experimental treatments in a way they could not continue over a one-year course. On the other hand, with a year-long effort, it is hard to be sure what conditions are really responsible for the results obtained.

Nevertheless, several studies of learning periods of intermediate length (eight hours, for example) did yield positive findings, and thus did benefit from the greater carefulness employed in analyzing the task and preparing instructional sequences than these investigators could have managed to achieve for a year-long course.

Another limitation of the research reviewed is that most experiments which used eduational materials rather than artificial laboratory tasks used materials from science and mathematics. These two subject-matter areas have for a long time been considered to have a highly organized

structure, and it is reasonable that the idea of hierarchies was first intensively explored with areas which were thought to be highly structured. Findings from experiments such as those by Gagne and his associates not only confirm the long-held idea that learning of mathematics and science is highly structured, but it goes on to show that this structure is hierarchical in nature and that, when learning is sequenced accordingly, learning is facilitated through transfer from subordinate skills to superordinate skills, as the hierarchy predicts. One next step clearly is to continue to explore the nature of the structure of mathematics and science for larger segments of instruction, and another need is to continue this type of research for subject matters other than mathematics and science.

A persistent problem in learning research is the failure of investigators to identify more clearly the type of learning which their tasks represent. One potential value of the book by Gagne (13) on the types of learning and their associated conditions is that it could provide a framework which experimenters could use to identify the categories into which their tasks belong. Then when the results are available it would be known more precisely how far these results can be generalized, e.g., to other tasks falling within the same category of learning. This is a very old problem in learning research which has long been recognized (Melton, 69) but not yet fully solved. Until experimenters do reach some sort of agreement on the classification of types of tasks, many supposedly conflicting findings will continue to be produced, whereas if the task were properly categorized the findings might not be so much in conflict since the conditions for each type of learning are supposedly different.

Another problem complicating much research in learning is the distinction between the type of task and the general learning conditions corresponding to it on one hand, and on the other hand the skill with which these learning conditions are implemented in the programing of the instruction. For example, in the series of studies on logical vs. scrambled sequence of frames in programed instruction, the extent to which each frame actually contains needed characteristics can get lost in the analysis of the sequencing effects within series of frames. One advantage which the research of Gagne and his associates has is that when the total learning is carefully broken up into subordinate competencies, one can correct to some degree for faulty programing by keeping the student working at each competency until he passes a test over it before he goes to the next, as Gagne and others have done. Errors in sequencing of frames or errors in their other characteristics within a unit of instruction on a competency then need not have so serious an adverse effect on instruction in the next competency to be acquired. The data from the testing of each competency, of course, can also be used to revise the program.

Considering all of the above problems often encountered in learning research, including research in sequencing, it would appear desirable that certain minimum conditions be regarded as characteristics which an experiment must have before it is conducted; then one could place

greater dependence upon the results. The minimum experimental conditions meant are listed under Recommendations. The importance of the consequences of not doing this is illustrated again in the group of studies concerning logical vs. random sequencing of frames. In those studies the experimenters hypothesized that sequencing does make a difference, then they designed experiments inadequately for the testing of the hypothesis, and they consequently attributed the negative results (appropriately) to the use of too short a program to represent classroom learning. This has happened so often as to make it apparent that future investigators should be warned to insure that the length of program selected is adequate to test the hypothesis, so that the results can be accepted rather than necessarily explained away insofar as their applicability to programs of lengths typically used for practical purposes.

Another important source of error in some of the experiments reviewed is that the variable sequencing conditions intended by the experimenter were not actually the conditions under which the Ss were responding. That is, when the purpose of the experiment is to test hypotheses about the sequencing of events during learning, if one does not exercise control (by machine or otherwise) to be sure that the intended sequence has been followed, the interpretation of the results is subject to gross error. In the type of sequencing studied by Gagne and his associates rigid control over the sequence of frames within a unit is not as critical as would be rigid control over the sequence of frames in the studies on scrambled vs. logical sequence of frames. In these latter experiments a teaching machine or some equivalent should be used so that a student cannot change the sequence involved in the experimental condition to which he has been assigned.

Finally, as has been shown earlier, when one sets out to test the hypothesis that a given task has a given type of structure by evaluating the results from a sequence of instruction fitting the inferred structure, if the results of the learning are contrary to the inferred structure, it is the hypothesis about the structure of that task which must be rejected, unless one believes it was poorly programed. In case of such an experimental finding, the rejection of the hypothesis that that task was of a given type of structure is appropriate; but the rejection of the hypothesis that other tasks have a given structure requiring given sequencing of instruction is inappropriate. Nobody whose work the reviewer has cited has said that all tasks have a particular structure, nor has anyone said that a given approach to sequencing is suitable for all tasks. Furthermore in such experiments, it is just as important to show the hypothesized structure of the experimental tasks as it is to report the learning results accurately.

The above discussion has been limited to factors which influence the interpretation of many of the experiments reviewed in this report. Since the reviewer's discussion has been integrated, in earlier chapters, with his account of the experiments reviewed for each of the nine types of sequencing study, the present discussion has been restricted to these general characteristics of research in the structure of courses and the sequencing of instruction.

Conclusions

The following conclusions and observations are numbered to correspond with the nine types of experiments reviewed in this report.

Type I: Maximum Learner Control

Independent study programs represent the maximum degree and variety of freedoms which the learner has in controlling the objectives of his study effort, as well as the sequence of study efforts made in an attempt to reach the objectives. This represents the highest degree of control over the learning by the learner, and correspondingly, the lowest degree of control by the teacher or other persons who have developed the materials selected by the student for use.

In such independent study programs there have been no known experiments which have precisely recorded the sequence in which the student went about his study efforts. Therefore there are no results to cite from experiments which would compare the results from varying courses of actions as adopted by several students attempting to achieve the same objectives. Evaluations of independent study programs then do not provide data sufficiently similar to the other types of data from other studies reviewed here to warrant citation, but mention is made of these programs to recognize their place in the continuum existing in learner control vs. experimenter control over the sequence of instruction.

Evaluations of such study programs often are based upon examination of the extent to which the objectives were achieved in order to determine how useful such programs are as compared to more conventional methods of instruction, but such evaluations do not add to our knowledge of how the structure of a course relates to the sequencing of instruction.

Type II: Learner-controlled Content and Sequencing

The experiment by Mager in student-sequenced learning represents the second most open type of experiment since the only restriction placed upon the learner was that he must be trying to learn something about electronics. Thus the learner was free to choose his specific objectives, and he was forced to ask for the information he wanted in whatever sequence he wanted. While this experiment was not designed primarily to determine how much learning resulted from this procedure, as compared to a programmer-controlled procedure, the results are cited in this report to show that when the trainee determines the sequence of the information he wants, this sequence differs from the sequence in which authors present materials in textbooks and from the sequence in which programmers arrange materials.

An implication of this study is that the now well-known procedure of accomplishing empirical revisions of first draft instructional programs does not institute the empirical procedure early enough in the

course development activity. The implication is that one should first state objectives, then give a test to determine how many of the objectives the entering trainees can already pass, and then find out which order they would like to employ in using the study materials. This order could then be compared to the order in which a programmer has prepared the materials, and then a criterion test could reveal whether a student can structure a course better than the programmer or not.

Type III: Learner Selection of Materials and Procedures

Student option in materials and study procedures used, as represented by the experiments summarized by Campbell, represent a variation from independent study programs in that the materials available were limited, were packaged into identified groups, and represented the total resources available to the student other than conferring with the teacher. This more definite identification of the materials available to the students in the self-directing groups does make it possible to describe better the materials and procedures available to them than is the case for independent study programs.

Nevertheless, there are two important reasons why the results of these studies cannot be interpreted explicitly in terms of the present topic. First, although the packages of material for the self-directed groups covered in general the same subject matter as in the programed instruction control condition, it cannot be said that the same content was presented to the two groups. Thus content differences were confounded with differences in the sequence of study steps the students used, so the results do not enable us to determine whether skill in preparing the programed instruction as compared to skill in preparing the self-instructional packages was the main source of variance, or whether the main source of variance was the learner's freedom to determine his own sequence of study activity. This is not necessarily a criticism of those experiments, because they were designed to determine from performance on a final examination which group learned the most overall.

This third type of experiment, then, was included to show how often sequencing is a variable in educational experiments although the experiments may not always be designed to enable one to interpret the results in terms of sequencing *per se*.

Type IV: Adjunct Autoinstruction; Mixed Experimenter and Learner Control

The early work by Pressey in adjunct autoinstruction is a still different variation in the kinds of sequencing control which are assigned to the learner as compared to the textbook writer or the instructor. In this case the sources of information are fixed as defined by the textbook, lecture, or other information sources. In the case of a textbook, there is no essential control over the order in which a student proceeds, whereas in the case of a lecture the

instructor would have control over the presentation of the material. The use of adjunct autoinstructional questions, however, enables the student in the case of the textbook to go back and review in whatever sequence he wishes, and in the case of a lecture permits the student to attempt to recall information given in whatever order he chooses.

Again, while there is no detailed record of the sequence the student uses, there is evidence that the adjunct questions result in better learning than instructional conditions which do not employ consistent and immediate feedback to the silent recitation efforts the student is required to make by responding to the self-test questions. Another advantage of the adjunct autoinstructional procedure is that it can be used with any basic medium of instruction, e.g., textbooks, lectures, films, field trips, demonstrations, etc.

Type V: Experimenter-determined Sequencing of Instruction in Accordance with Hierarchies of Competence

The studies by Gagne and his associates in sequencing of instruction in accordance with subordinate and superordinate competencies implied in a logically derived hierarchical structure for the course represents the group of experiments which in a sense yield the most information on the topic at hand. By most information, here is meant the most information on how the learning is structured and the most information confirming the hypothesis that the units of instruction should be presented in an order corresponding with the hierarchy. Gagne's studies thus address themselves primarily to the sequencing of units of instruction corresponding to individual competencies to be acquired. Thus frame sequencing *per se* was not a variable in these experiments, although the type of frames employed was a variable in some experiments such as the one on number and variety of examples.

An important implication of this series of studies as a whole is that the arrangement of the units of instruction in accordance with the order in which competencies need to be acquired appears to be a more powerful influence determining criterion scores than do such other variables such as number and type of examples which represent characteristics of the instruction within a unit of instruction. This is not to say that such variables as number and type of examples are not important in learning, but it is to say that in a hierarchically structured course, if the units are arranged in the wrong order it may not matter how skillfully the instruction is programed in the frames comprising the unit. It is considered that in the face of the difficulties mentioned in the discussion section in conducting learning experiments on this topic, the results of Gagne and his associates which confirmed hypotheses in from 90 to 100% of the cases analyzed is actually a remarkable finding in this kind of research. While one reviewer has said that the type of data analysis Gagne employed does not establish the idea of sequencing in accordance with hierarchies, the overwhelming consistency of his findings argues strongly for the acceptability of the underlying hypotheses, even though the present reviewer recommends also the conduct of other kinds of experiments in the search for additional confirmation of these hypotheses.

Type VI: Experimenter-determined Sequencing of Frames in Programed Instruction

In this report, frequent mention has been made of deficiencies in the design of experiments comparing logical vs. scrambled sequencing of frames within units of an instructional program or among units in such programs. Considering the difficulties and criticisms already pointed out in connection with this group of studies, it should be said that often the investigators, in the reports of their experiments, offer very helpful interpretative comments even though the experiments for the most part were inadequately designed to test their hypotheses.

Outstanding among this group of experiments in this respect is the study by Hamilton. Her interpretation of the factors which simultaneously are brought into action when sequencing has been changed represents a useful caution that when an investigator changes the sequencing of frames in a program, the results are likely to be a function of an interaction of sequencing with other characteristics of the program such as response mode, degree of prompting, provision for feedback, and the extent to which major concepts remain in the same sequence even though the frames relating to each concept are scrambled. For this reason the study reported by this investigator can be instructive for others intending to conduct further research on frame sequence.

Type VII: Learner-determined Branching in Autoinstruction

Branching in autoinstruction is an interesting variation in the context of this report because the basic or "main line" program is essentially a linear program arranged in a fixed sequence by the programmer. Unlike regular linear programs, however, the fixed sequence is departed from whenever the student's performance so indicates. This departure from the prepared sequence may take place at the very beginning of the instruction because a test has revealed that the student does not need all of the instruction. In other instances this departure from the fixed sequence takes place because it is determined after instruction has begun that the student can bypass information which it was not realized earlier he could bypass, or the student is branched into remedial loops of instruction because of errors made, and furthermore, the content of this branching remedial loop may vary in accordance with the particular errors made.

Thus in branching procedures there is the interesting combination of predetermined, pre-sequenced, main line programs, with departures from them made either based on the objective performance of the student or on the basis of the student's subjective belief that he should depart from the fixed sequence. Since the instruction is divided into frames and since exact records can be kept as to when the student departs from the linear program, experiments comparing linear and branching programs have produced more explicit data for analysis than do the results of some of the earlier types of sequencing experiments reported here.

The characteristics of branching programs appeal to most people for reasons discussed earlier. In brief, they represent attempts to adjust to individual differences in prior entering competencies, in different routes of progress toward the goal, in different misconceptions made apparent during the learning, and in different reasons for errors made. All of these individual considerations are added to the characteristic that linear programs have, namely that the learner is allowed to progress at his own rate.

It appears strange that since everyone believes and knows that certain kinds of individual differences do exist and therefore believes also that instruction should adjust to these individual differences, in spite of this the results showing branching superior to linear programs are quite meager. The reasons for this have been reviewed earlier, and it may be expected that branching programs will ultimately prove quite superior to linear programs, provided of course that the basic programming itself is done with equal skill in the two cases. However, one safeguard to be observed when branching procedures are employed is that in the effort to branch in accordance with the student's immediate past performance one does not neglect the important requirement to introduce redundancy and review in order to achieve good retention.

Type VIII: Experimenter-prepared Advance Organizers

Research in advance organizers represents an unusual application of sequencing matters for two reasons. First, it is based on a cognitive theory of learning rather than a behavioral or eclectic view of learning, and consequently it deals with ideas and their organization as achieved by the learner rather than with competencies the students achieve. For this reason it has been said in this report that cognitive theory may be the more usable theory in the less desirable of two educational situations. That is, it has been recommended that behavioral objectives be stated for a course, and when this is done, analyses of the structure of the course in terms of required competencies (of whatever structure) are therefore possible. But when behavioral objectives are not stated and one regards the student's task as acquiring the substance of the content contained in the course materials, cognitive theory is perhaps more capable of being utilized. If it is correct to recommend that behavioral objectives should be stated for all courses, the apparent implication is that the kind of theory and procedure employed by Gagne in regard to the nature and sequencing of instruction would then appear more precise and useful than the cognitive theory underlying the utilization of advance organizers.

Nevertheless, the research using advance organizers constitutes a source of information on the possibility that presenting the student with a brief abstract of the substance of the learning material in the most general form in which the ideas can be expressed may have the advantages of precluding rote learning and providing a conceptual framework facilitating the learning of the materials in the lesson itself.

Type IX: Sequences Pre-planned by the Experimenter to Test Hypotheses about Effective Characteristics of Learning Programs

The miscellaneous experiments categorized as Type IX are concerned primarily with laboratory type tasks not consisting of sequences of material prepared for educational purposes. While often such tasks, if properly identified in terms of type of learning, permit some controlled investigations otherwise not easy to arrange with actual educational material, the reviewer's bias is that too often the experimenter contrives a task and learning procedures which are not meaningful in terms of any type of learning of educational importance, and therefore the results are not generalizable to any educational objective.

The experiments classified as Type IX involve experiments designed not so much to test alternate strategies of sequencing but rather to simply employ sequencing as a means to put into effect variations in another independent variable the effects of which the investigator wishes to examine. Thus sequencing becomes involved as a mechanism enabling one to study some other point of theory or practice which is of interest.

Recommendations

Research Recommendations

Considering the positive results which were found from some of the experiments reviewed, continued research is believed worthwhile on the topic of course or task structure as it relates to the sequencing of instruction.

The reviewer's recommendation concerning the conduct of further research may be briefly summarized as follows:

1. There is a great need to extend research such as that by Gagne and his associates to cover larger segments of a curriculum (or an entire course of study) so that the hypotheses concerning hierarchical structures and effective sequencing can be tested in the context of larger blocks of instruction.

2. Since most research on the topic has involved mathematics and science as the subject matter, this research should be extended into social sciences and other areas so that it can be determined the extent to which courses do have hierarchical learning structures as compared to other types of structure which may exist but for which little data is available. Possible other types of course structure were identified and discussed in Chapter I.

3. Whether future experiments involve large or small segments of instructional courses, the specific learning objectives and their subordinate competencies should be analyzed in such a way that not only the structure of the learning is shown but the type of learning involved for each objective or competency also is identified.

4. Then studies of the relationship of structure of objectives to sequencing of the instruction could be related to the as-yet-uninvestigated question of whether some aspects of sequencing should vary for example for two adjacent competencies representing the same type of learning as compared to two adjacent competencies of different learning types.

5. Experiments in sequencing should involve control over the skillfulness of the programing, so that one sequence of well-programed material is not compared with another sequence of poorly-programed material. Some experiments here reported have achieved this, such as by using the same words (content) in different sequences; but other experiments have reported comparisons of two grossly different packages of material which surely must vary in many unknown ways other than sequencing.

6. Effective programs should be used in sequencing experiments. They should be tested and revised until they are effective; then various ways of sequencing the same materials could be undertaken to maximize effectiveness.

7. Entering competency should be measured, so that the results can be used to distinguish clearly effect of recent learning from effects of old learning. These distinct portions of the data may have much relevance for sequencing to ensure retention and transfer.

8. Experiments studying sequencing in relation to task or course structure should make both explicit in their reports.

9. Degrees of learner vs. experimenter control should be investigated for the same task; both as to how units are sequenced and as to how frames are sequenced.

10. Textual materials should be written on basis of a defined task structure enabling experiments to be done to compare reading of well-sequenced prose with well-sequenced programed instruction, adjunct auto-instruction, and other teaching approaches. Such materials could result in better experiments in optimum ratio of words read to responses required, thus defining size of step for methods other than the prevailing type of linear programs.

11. The well-known procedure of submitting draft programs to empirical test as a basis for revision should be enlarged and extended to sequencing and other issues, to see if the learner can help design programs more effective than programs not so defined.

12. Results of alternate sequencing strategies should be analyzed to search for interactions with age of the learner and any other relevant learner characteristic, so that sequence, size of step, mode of response, etc., can be analyzed as a function of learner characteristics as well as of task characteristics.

13. Interaction studies of a complex nature are recommended, e.g., task x learner characteristic x media of instruction x sequence x size of step, etc.

Some of the research recommended above could best be pursued by independent investigators planning small-scale single-purpose studies or interaction studies. Some could best be done by a larger "programatic" research effort, in which individual investigators jointly plan out the areas to be studied. Some could be best done in conjunction with a large-scale curriculum-development project. And some, the reviewer suggests, could be done by one or more "grand slam" experiments designed somewhat along the following lines:

Step 1. Identify and describe the characteristics of several kinds of learning so that members of a team of researchers will recognize them as distinct kinds of learning.

Step 2. Identify at least two tasks of each type of learning.

Step 3. Construct a curriculum that requires many tasks representing all kinds of learning, and at least two tasks of each type.

Step 4. Construct a diagram showing the structure of the entire curriculum.

Step 5. Construct a diagram showing the structure of each task.

Step 6. Program the entire curriculum (a) in the sequence implied by the structures, (b) in the reverse sequence of (a), and (c) in a random sequence.

Step 7. Find out which sequence is learned and retained best.

Step 8. Then, utilizing the best sequence, do studies of response mode, size of step, conditions of feedback, etc., for the entire curriculum and for each task representing each kind of learning.

Step 9. Continue such experiments until results on each variable, for the tasks chosen to represent different kinds of learning, can be meaningfully interpreted.

Step 10. Test the generalizability of each finding to other tasks for the same kind of learning.

Step 11. Then show what principles, if any, apply for all kinds of learning, and which apply to only one or more types of learning.

Recommendations for Curriculum Design

It should be apparent that the research reviewed in this report has many implications for the design of improved curricula. It was not intended that this report emphasize these as heavily as the implications for needed future research. However, it is also apparent that the steps

outlined immediately above for a recommended large-scale experimental study could, with a few modifications, become the basis for a very analytic approach to curriculum development, which even though representing a large effort, might cost no more than existing projects which innovate in content areas rather than in the areas of instructional method. Ideally, the same new curriculum development projects should forge new frontiers in both areas.

This is not to say that every possible experiment should be repeated over again as a part of every curriculum development effort. Rather, the results of the recommended research should guide the curriculum innovation by careful selection of those results applicable to each curriculum project.

Just as experiments are recommended which would integrate heretofore separate research areas (structure, sequence, type of learning, media, and learner variables), so might curriculum design be directed in the future.

There appears to be emerging in recent years the outlines of a systematic approach for the development of curricula which, though laborious, time-consuming, and expensive, could, if put into practice, result potentially in far improved educational efforts. It is simply a matter of the resources which are available to apply the total state of the art in implementing curriculum development.

Thus a new curriculum-project task-force might seek to apply in practice the implications of research reviewed here and in other such reviews. In so doing, here are a few of the key elements in the approach:

1. A performance-oriented view of curriculum design (Gagne, 32).

2. Attention to the kinds of learning required (Gagne, 31).

3. Attention to structure and sequence (the present report).

4. Attention to selection of the instructional media (Briggs *et al.*, 15)

5. Attention to more skillful programing techniques for media (all analytical research in instructional media).

6. Expansion of present empirical program-revision technique (Markle, 66).

7. Overall course evaluation, feedback, revision, and re-evaluation (a field of its own).

Such a recommendation as the above one seems already to have occurred to many others. Recent program plans in the Office of Education reflect the general intent to pursue similar research and development plans. It is hoped that this report will be useful to those programs.

References

1. Adams, J. A. "Part Trainers." In G. Finch (Ed.), Educational and Training Media: A Symposium. Washington, D. C.: National Academy of Sciences--National Research Council, Publication 789, 1960. p. 129-149.

2. Amster, Harriett. "Effect of Instructional Set and Variety of Instances on Children's Learning," Journal of Educational Psychology. 57, 1966. p. 74-85.

3. Anderson, R. C. "Educational Psychology." In P. R. Farnsworth (Ed.), Annual Review of Psychology. Palo Alto, Calif.: Annual Reviews, Inc., 1967. p. 129-164.

4. Ausubel, D. P. "The Use of Advance Organizers in the Learning and Retention of Meaningful Verbal Material," Journal of Educational Psychology. 51, (5) October 1960. p. 267-272.

5. Ausubel, D. P. The Psychology of Meaningful Verbal Learning. New York: Grune & Stratton. 1963.

6. Ausubel, D. P., and Fitzgerald, D. "The Role of Discriminability in Meaningful Verbal Learning and Retention," Journal of Educational Psychology. 52, October 1961. p. 266-274.

7. Ausubel, D. P., and Fitzgerald, D. "Organizer, General Background, and Antecedent Learning Variables in Sequential Verbal Learning," Journal of Educational Psychology. 53, December 1962. p. 243-249.

8. Ausubel, D. P., and Youssef, M. "Role of Discriminability in Meaningful Parallel Learning," Journal of Educational Psychology. 54, December 1963. p. 331-336.

9. Belcastro, F. P. "A Comparison of Four Techniques of Programing Algebra," Science Education. Vol. 47 (1963). p. 388-393.

10. Belcastro, F. P. "Relative Effectiveness of the Inductive and Deductive Methods of Programming Algebra," Journal of Experimental Education. 34, Spring 1966. p. 77-82.

11. Briggs, L. J. "The Development and Appraisal of Special Procedures for Superior Students and an Analysis of the Effects of 'Knowledge of Results'," Abstracts of Doctoral Dissertations. No. 58, The Ohio State University Press. 1949.

12. Briggs, L. J. "Two Self-instructional Devices," Psychological Reports. 4, 4 December 1958. p. 671-676.

13. Briggs, L. J. "Teaching Machines." In G. Finch (Ed.), Educational and Training Media. Washington, D. C.: National Academy of Sciences--National Research Council, Publication 789, 1960. p. 150-195.

14. Briggs, L. J., and Angell, D. "Programed Instruction in Science and Mathematics," Review of Educational Research. 34, No. 3, June 1964. p. 354-373.

15. Briggs, L. J.; Campeau, Peggie L.; Gagne, R. M.; and May, M. A. Instructional Media: A Procedure for the Design of Multi-media Instruction. A Critical Review of Research, and Suggestions for Future Research. Pittsburgh, Pa.: American Institutes for Research, Monograph No. 2. 1967.

16. Briggs, L. J., and Hamilton, Nancy Russell. "Meaningful Learning and Retention: Practice and Feedback Variables," Review of Educational Research. 34, No. 5, December 1964. p. 545-558.

17. Bruner, J. S. "Needed: A Theory of Instruction," Educational Leadership. 20, May 1963. p. 523-532.

18. Buckle, C. F. "Relations Between Attainment of Component Skills, Retention, and Transfer," Programmed Learning and Educational Technology. 4, July 1967. p. 179-184.

19. Campbell, V. N. "Bypassing as a Way of Adapting Self-instruction Programs to Individual Differences," Journal of Educational Psychology. 54, 1963. p. 337-345.

20. Campbell, V. N. "Self-direction and Programed Instruction for Five Different Types of Learning Objectives," Psychology in the Schools. 1, October 1964. p. 348-359.

21. Coulson, J. E., and Silberman, H. F. "Automated Teaching and Individual Differences," AV Communication Review. 9, January - February 1961. p. 5-15.

22. Coulson, J. E., and others. "Effects of Branching in a Computer Controlled Autoinstructional Device," Journal of Applied Psychology. 46, December 1962. p. 389-392.

23. Cox, J. A., and Boren, L. M. "A Study of Backward Chaining," Journal of Educational Psychology. 56, October 1965. p. 270-274.

24. Crowder, N. A. "Automatic Tutoring by Means of Intrinsic Programing." In E. H. Galanter (Ed.), Automatic Teaching: The State of the Art. New York: Wiley & Sons. 1959. p. 109-116.

25. Evans, J. L.; Homme, L. E.; and Glaser, R. "The RULEG System for the Construction of Programed Verbal Learning Sequences," Journal of Educational Research. 55, June - July 1962. p. 513-518.

26. Fitzgerald, D., and Ausubel, D. P. "Cognitive versus Affective Factors in the Learning and Retention of Controversial Material," Journal of Educational Psychology. 54, April 1963. p. 73-84.

27. Fowler, W. "The Design of Early Developmental Learning Programs for Disadvantaged Young Children," Supplement to the IRCD Bulletin. New York: ERIC Information Retrieval Center for the Disadvantaged, Vol. III, 1A. FERKAUF Graduate School, 55 Fifth Avenue, New York. (Undated)

28. Gagne, R. M. "The Effect of Sequence of Presentation of Similar Items on the Learning of Paired Associates," Journal of Experimental Psychology. 40, February 1950. p. 61-73.

29. Gagne, R. M. "The Acquisition of Knowledge," Psychological Review. 69, July 1962. p. 355-365.

30. Gagne, R. M. "The Learning Requirements for Enquiry," Journal of Research in Science Teaching. 10, Issue 2, June 1963. p. 144-153.

31. Gagne, R. M. The Conditions of Learning. New York: Holt, Rinehard and Winston. 1965.

32. Gagne, R. M. "Curriculum Research and the Promotion of Learning," Monographs on Curriculum Research and Evaluation. Chicago: Rand-McNally, American Educational Research Association. No. 1, 1967.

33. Gagne, R. M. "Training and Principles of Learning." In E. A. Fleishman (Ed.), Studies in Personnel and Industrial Psychology. Homewood, Ill.: The Dorsey Press. 1967. Chapter 19, p. 175-181.

34. Gagne, R. M., and Bassler, O. C. "Study of Retention of Some Topics of Elementary Nonmetric Geometry." Journal of Educational Psychology. 54, June 1963. p. 123-131.

35. Gagne, R. M.; Mayor, J. R.; Garstens, Helen L.: and Paradise, N. E. "Factors in Acquiring Knowledge of a Mathematical Task," Psychological Monographs: General and Applied. 76, 1962. (Whole No. 526).

36. Gagne, R. M., and Paradise, N. E. "Abilities and Learning Sets in Knowledge Acquisition," Psychological Monographs: General and Applied. 75, 1961. (Whole No. 518).

37. Gagne, R. M., and Staff, University of Maryland Mathematics Project. "Some Factors in Learning Nonmetric Geometry," Mongraphs of the Society for Research in Child Development. 30, 1965. p. 42-49.

38. Gavurin, E. I., and Donahue, Virginia M. "Logical Sequence and Random Sequence," Automated Teaching Bulletin. 1, Spring 1961. p. 3-9.

39. Gilbert, T. F. "Mathetics: The Technology of Education," Journal of Mathetics. 1, January 1962. p.7-73.

40. Glaser, R. "Some Research Problems in Automated Instruction; Instructional Programming and Subject-matter Structure." In J. E. Coulson (Ed.), Programmed Learning and Computer-based Instruction. New York: Wiley & Sons. 1962. p. 67-85.

41. Goldbeck, R. A., and Briggs, L. J. "An Analysis of Response Mode and Feedback Factors in Automated Instruction." In W. I. Smith and J. W. Moore (Eds.), Programed Learning. New York: D. Van Nostrand Co. 1962. p. 166-191.

42. Goldbeck, R. A., and Campbell, V. N. "The Effects of Response Mode and Response Difficulty on Programed Learning," Journal of Educational Psychology. 53, June 1962. p. 110-118.

43. Guthrie, E. R. The Psychology of Learning. New York: Harper & Row. 1935.

44. Hamilton, Nancy Russell. "Effects of Logical versus Random Sequencing of Items in an Autoinstructional Program under Two Conditions of Covert Response," Journal of Educational Psychology. 55, (No. 5) 1964. p. 258-266.

45. Harlow, H. F. "The Formation of Learning Sets," Psychological Review. 56, 1949. p. 51-65.

46. Haygood, R. C., and Bourne, L. E., Jr. "Attribute- and Rule-learning Aspects of Conceptual Behavior," Psychological Review. 72, 1965. p. 175-195.

47. Hershberger, W. "Self-evaluational Responding and Typographical Cueing: Techniques for Programing Self-instructional Reading Materials," Journal of Educational Psychology. 55, October 1964. p. 288-296.

48. Hilgard, E. R. Theories of Learning (2nd Ed.) New York: Appleton-Century-Crofts. 1956.

49. Holland, J. G. "A Quantitative Measure for Programed Instruction," American Educational Research Journal. 4, March 1967. p. 87-101.

50. Holland, J. G., and Skinner, B. F. The Analysis of Behavior: A Program for Self-instruction. New York: McGraw-Hill Book Co., Inc. 1961.

51. Hull, C. L. Principles of Behavior. New York: Appleton-Century-Crofts. 1943.

52. Hunt, E. B., and Hovland, C. I. "Order of Consideration of Different Types of Concepts," Journal of Experimental Psychology. 59, April 1960. p. 220-225.

53. Johnson, D. M.; Parrott, G. R.; and Stratton, R. P. "Productive Thinking: Produce One Solution or Many." Proceedings of the 75th Annual Convention of the American Psychological Association. Washington, D. C.: American Psychological Association. 1967. p. 299-300.

54. Kaess, W., and Zeaman, D. "Positive and Negative Knowledge of Results on a Pressey-type Punchboard," Journal of Experimental Psychology. 60, 1960. p. 12-17.

55. Karraker, R. J. "Knowledge of Results and Incorrect Recall of Plausible Multiple-choice Alternatives," Journal of Educational Psychology. 58, February 1967. p. 11-14.

56. Kersh, B. Y. "Programing Classroom Instruction." In R. Glaser (Ed.), Teaching Machines and Programed Learning, II: Data and Directions. Washington, D. C.: National Education Association of the U. S. 1965. p. 321-368.

57. Krathwohl. D. R. "Stating Objectives Appropriately for Program, for Curriculum, and for Instructional Materials Development," Journal of Teacher Education. 16, 1965. p. 83-92.

58. Krumboltz, J. D., and Weisman, R. G. A program on educational measurements. (Further identification not found).

59. Lee, S. S. Transfer from Lower-level to Higher-level Concept. Unpublished doctoral dissertation, University of California, Berkeley. 1967.

60. Levin, G. R., and Baker, B. L. "Item Scrambling in a Self-instructional Program," Journal of Educational Psychology. 54, Issue 3, 1963. p. 138-143.

61. Lumsdaine, A. A. Student Response in Programmed Instruction. Washington, D. C.: National Academy of Sciences--National Research Council, Publication 943. 1961.

62. Lumsdaine, A. A., and Glaser, R. (Eds.). Teaching Machines and Programmed Learning. Washington, D. C.: Department of Audio-Visual Instruction, National Education Association. 1960.

63. Mager, R. F. "On Sequencing of Instructional Content," Psychological Reports. 9, October 1961. p. 405-413.

64. Mager, R. F., and Clark, C. "Explorations in Student-controlled Instruction," Psychological Reports. 13, August 1963. p. 71-76.

65. Margolius, G. J.; Sheffield, F. D.; and Maccoby, N. "Timing of Demonstration and Overt Practice as a Function of Task Organization." In A. A. Lumsdaine (Ed.), Student Response in Programed Instruction. Washington, D. C.: National Academy of Sciences--National Research Council, Publication 943. 1961. p. 101-105.

66. Markle, D. G. The Development of the Bell System First Aid and Personal Safety Course: An Exercise in the Application of Empirical Methods to Instructional System Design. Palo Alto, Calif.: American Institutes for Research. April 1967.

67. McGeoch, J. A., and Irion, A. L. The Psychology of Human Learning. New York: Longmans. 1952.

68. McNeil, J. D. "Programed Instruction versus Usual Classroom Procedures in Teaching Boys to Read," American Educational Research Journal. 1, March 1964. p. 113-119.

69. Melton, A. W. (Ed.) Categories of Human Learning. Ann Arbor: University of Michigan. 1962.

70. Merrill, M. D. "Correction and Review on Successive Parts in Learning a Hierarchical Task," Journal of Educational Psychology. 56, October 1965. p. 225-234.

71. Merrill, M. D., and Stolurow, L. M. "Hierarchical Preview vs. Problem Oriented Review in Learning an Imaginary Science," American Educational Research Journal. 4, November 1966. p. 251-261.

72. Meyers, D. E. Adjunct to Self-study for Aircrew Refresher Training under Operational Conditions in the Air Defense Command. Dayton, Ohio: Wright-Patterson Air Force Base. (AMRL-TR-65-83). Behavioral Sciences Laboratory. March 1965.

73. Morrisett, L., Jr., and Hovland, C. I. "A Comparison of Three Varieties of Training in Human Problem Solving," Journal of Experimental Psychology. 58, July 1959. p. 52-55.

74. Neisser, U., and Weene, P. "Hierarchies in Concept Attainment," Journal of Experimental Psychology. 64, December 1962. p. 640-645.

75. Newton, J. M., and Hickey, A. E. "Sequence Effects in Programed Learning of a Verbal Concept," Journal of Educational Psychology. 56, June 1965. p. 140-147.

76. Odom, R. R. "Sequence and Grade Placement of Capitalization Skills," Elementary English. 38, 1961. p. 118-121.

77. Odom, R. R. "Sequence and Grade Placement of Punctuation Skills," California Journal of Educational Research. 13, 1962. p. 179-185.

78. Payne, D. A.; Krathwohl, D. R.; and Gordon, J. "The Effect of Sequence on Programmed Instruction," American Educational Research Journal. 4, March 1967. p. 125-132.

79. Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores--and Teaches," School and Society. 13 (No. 586), March 20, 1926. p. 373-376.

80. Pressey, S. L. "A Machine for Automatic Teaching of Drill Material,"School and Society. 25 (No. 645), May 7, 1927. p. 549-552.

81. Pressey, S. L. "A Third and Fourth Contribution Toward the Coming 'Industrial Revolution' in Education," School and Society. 36 (No. 934), November 19, 1932. p. 668-672.

82. Pressey, S. L. "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-instruction," Journal of Psychology. 29, April 1950. p. 417-447.

83. Pressey, S. L. "Re-program Programing?", Psychology in the Schools. 4, July 1967. p. 234-239.

84. Pressey, S. L., and Kinzer, J. R. "Auto-elucidation without Programing!", Psychology in the Schools. 1, Issue 4, October 1964. p. 359-365.

85. Roe, A. "A Comparison of Branching Methods for Programmed Learning," Journal of Educational Research. 55, 1962. p. 407-416.

86. Roe, K. V.; Case, H. W.; and Roe, A. "A Scrambled versus Ordered Sequence in Autoinstructional Programs," Journal of Educational Psychology. 53, April 1962. p. 101-104.

87. Rotberg, Iris C., and Woolman, M. "Verbal Paired-associate Learning as a Function of Grouping Similar Stimuli or Responses," Journal of Experimental Psychology. 65, January 1963. p. 47-51.

88. Rothkopf, E. Z. "Stimulus Similarity and Sequence of Stimulus Presentation in Paired-associate Learning," Journal of Experimental Psychology. 56, 1958. p. 114-122.

89. Rothkopf, E. Z. "Learning from Written Instructive Material: An Exploration of the Control of Inspection Behavior by Test-like Events," American Educational Research Journal. 3, November 1966. p. 241-250.

90. Scandura, J. M. "An Analysis of Exposition and Discovery Modes of Problem Solving Instruction," The Journal of Experimental Education. 33, Winter 1964. p. 149-159.

91. Scandura, J. M. "The Basic Unit in Meaningful Learning--Association or Principle?" (Mimeographed). (Based on a paper read at the American Educational Research Association Convention in Chicago on February 19, 1966. Also submitted to The School Review).

92. Scandura, J. M. "Precision in Research on Mathematics Learning: The Emerging Field of Psycho-mathematics," Journal of Research in Science Teaching. 4, 1966. p. 253-274.

93. Scandura, J. M. "Prior Learning, Presentation Order, and Prerequisite Practice in Problem Solving," Journal of Experimental Education. 34, Summer 1966. p. 12-18.

94. Scandura, J. M. "Teaching--Technology or Theory," American Educational Research Journal. 3, March 1966. p. 139-146.

95. Scandura, J. M. "The Basic Unit in Meaningful Learning--Association or Principle?" Proceedings of the 75th Annual Convention of the American Psychological Association. Washington, D. C.: American Psychological Association. 1967. p. 275-276.

96. Scandura, J. M., and Wells, J. N. "Advance Organizers in Learning Abstract Mathematics," American Educational Research Journal. 4, May 1967. p. 295-301.

97. Schutz, R. E.; Baker, R. L.; and Gerlach, V. S. "Teaching capitalization with a Programed Text," AV Communication Review. 10, 1962. p. 359-363.

98. Schutz, R. E.; Baker, R. L.; and Gerlach, V. S. Measurement Procedures in Programmed Instruction. Tempe, Ariz.: Classroom Learning Laboratory, Arizona State University. 1964.

99. Sheffield, F. D., and Maccoby, N. "Summary and Interpretation of Research on Organizational Principles in Constructing Filmed Demonstrations." In A. A. Lumsdaine (Ed.), Student Response in Programed Instruction. Washington, D. C.: National Academy of Sciences--National Research Council, Publication 943. 1961. p. 117-131.

100. Short, J., and Haughey, Betty. "Study of Sequencing Strategies," Proceedings of the 75th Annual Convention of the American Psychological Association, Vol. 2. Washington, D. C.: American Psychological Association. 1967. p. 317-318.

101. Skinner, B. F. The Behavior of Organisms; An Experimental Analysis. New York: Appleton-Century-Crofts. 1938.

102. Skinner, B. F. "Teaching Machines," Science. 128, October 24, 1958. p. 137-158.

103. Slack, C. W. Lesson Writing for Teaching Verbal Chains: Order in Which Elements are Taught. Stamford, Conn.: TOR Educations. 1964.

104. Smith, W. I., and Moore, J. W. Learning Sets in Programmed Instruction. Final Report, U. S. Office of Education Grant No. 7-48-0070-208. March 1965.

105. Stolurow, L. M. Teaching by Machine. Washington, D. C.: U. S. Department of Health, Education, and Welfare (Cooperative Research Monograph No. 6). 1961.

106. Suppes, P. "Modern Learning Theory and the Elementary-school Curriculum," American Educational Research Journal. 1, March 1964. p. 79-93.

107. Thorndike, E. L. Animal Intelligence. New York: Macmillan. 1911.

108. Tolman, E. C. "There is More Than One Kind of Learning," Psychological Review. 56, 1949. p. 144-155.

109. Traub, R. E. "Importance of Problem Heterogeneity to Programed Instruction," Journal of Educational Psychology. 57, Issue 1, 1966. p. 54-60.

110. Tversky, A. "On the Optimal Number of Alternatives at a Choice Point," Journal of Mathematical Psychology. 1, No. 2, July 1964. p. 386-391.