OE-34034
Bulletin 1964, No. 35

The Research On Programed Instruction

An Annotated Bibliography

by
WILBUR SCHRAMM
Institute for Communication Research
Stanford University

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education

Foreword

The Office of Education is pleased to make available this annotated bibliography of the research on programed instruction conducted in this country since 1954.

The bibliography was prepared under a contract with the Institute for Communication Research at Stanford University by a national authority on new instructional media. Dr. Wilbur Schramm, who has conducted several projects with the support of the new Educational Media Program of the Office of Education, has repeatedly demonstrated his wide range of capabilities in this field.

Because programed instruction is becoming an increasingly important and ever more effective teaching technique in American education, it is anticipated that this volume will be of major value.

RALPH C. M. FLYNT
Associate Commissioner,
Bureau of Educational Research and Development.

Contents

Introduction

Since B. Fred Skinner's now historic article of 1954, "The Science of Learning and the Art of Teaching," there have been approximately 190 reports of original research on programed instruction. More than 165 of these have appeared since 1959. No method of instruction has ever come into use surrounded by so much research activity; indeed for a time it seemed that there would be more research than programs. It seems appropriate now to look back over this flurry of research and review some of its conclusions.

A word of caution is necessary, however. The boundaries of this field are most unclear. In one sense, a very large part of the work in experimental psychology of learning belongs with the research on programed instruction, and in another sense this research should be limited to studies of programed texts and teaching machine programs. The first of these conceptions of research on programed instruction would make the field impossibly large; the second would unduly restrict it. We have taken a position nearer the latter than the former point of view, but have not insisted, before including an experiment, that it deal with programs in the narrow sense.[1] For example, an experiment on learning from a film (like the experiment of Weiss, Maccoby, and Sheffield, 1961) that contributes to our understanding of length of step, we have included. Similarly, we have included some experiments on paired-associate learning and some of Pressey's experiments on providing immediate knowledge of results after test questions, although some readers might not regard either of these topics as classifiable under programed instruction in the strictest sense. Our expectation is that the borders of this field of research will expand. As the generality of its problems becomes more apparent and the eccentricities of the method less diverting, we rather expect research on programed instruction to merge with the broader stream of research on instructional technology, to the benefit of both. The appearance in 1961 of *Student Response in Programed Instruction*, edited by A. A. Lumsdaine, which reported a number of years of Air Force research on training methods, using films, flashcards, recordings, and drill machines, as well as programs, and concentrating on such basic variables as

[1] We are not reviewing the numerous articles on how to write programs, or the books and articles that deal with research and theory but do not present new research.

practice, participation, prompting, motivation, feedback, transfer, retention, and the like, appears to be a step in the direction of the merger we anticipate. For the time being, however, it is still useful to review the research on programed instruction by itself, defining it rather closely, as we propose to do in the following pages.

In the second place, there is a somewhat disproportionate number of experiments in this field which report a finding of "no significant differences" (n.s.d.). It is perhaps unfortunate that this should occur so early in the history of a new method, because this is a time when people are looking to research for guidance and also because a finding of n.s.d. gives us less information than a finding of a difference. When a researcher finds a difference, he has statistical tools at hand to calculate the likelihood that his finding is due to chance; when he finds no difference, he has "no logically defensible basis," as A. A. Lumsdaine cogently points out in reference to drawing a realistic conclusion. "Any set of data compatible with a null hypothesis of *no* difference is also compatible with a number of alternative hypotheses that some difference *does* exist, even though it is not gross enough, in relation to the variability of the data to be significant."[2] R. A. Fisher made the point long ago that the null hypothesis can only be disproved; it cannot be proved.[3] The numerous experiments on programed instruction that do not succeed in disproving the null hypothesis may, indeed, be proving that no significant difference exists, but the suspicion arises that in many cases the programs are too short, the samples too small, the measuring instruments too dull, to pick up differences if they exist. Moreover, it is often very difficult to extrapolate from findings on short programs to the conditions of classroom use.

What the Research Has Disclosed

Now let us describe in gross terms the existing research on programed instruction, insofar as we have been able to discover and examine it. More than three-fourths of all the research papers in the field have appeared, as we have indicated, in the past 3 years. Nearly half the papers deal centrally with what we might call presentation variables—prompting and confirmation, branching, pacing, size of step, machine vs. text, programed television, and so forth. Nearly 30 percent deal chiefly with response modes—overt vs. covert, multiple-choice vs. constructed responses, and the like. Among the remaining experiments are a considerable number of evaluative tests,

[2] Lumsdaine, A. A. "Instruments and Media of Instruction," Gage, N., ed., *Handbook of Research on Education*. New York: Rand McNally, 1961.
[3] Fisher, R. A. *The Design of Experiments*, London. Boyd, 1961 (8th ed.).

which seek to compare the amount of learning from programs with the amount of learning from conventional classroom teaching of the same subject. A few experiments are concerned with special applications of programs—to slow learners, to deaf children, to industrial trainees, to voluntary and individual users, and so forth. A few others are concerned with the use of programs for special objectives, such as discovery teaching or transfer of training. It is interesting to note that only a handful of experiments make use of the "intrinsic" kind of programing, of which Norman Crowder is the chief exponent. Although there are experiments on programed films, programed television, paired-associate learning from flashcards, and Pressey tests, as well as the few on intrinsic programing, the great majority of the studies on programed instruction have been done with linear programs (or with linear programs fitted with branches or loops). As we might expect, about 2 out of 5 of the experiments have been done with college students as subjects, about 1 in 5 with secondary-school children, a little less than 1 in 5 with adult or military samples. About 1 in 8 has dealt with primary-school children, and a small scattering with preschool samples.

One final note by way of introduction: Perhaps as many as one-fourth of the circulating papers in this field have never been published. This is not uncommon in a new field, where scholars are anxious to learn of results and where the journals inevitably lag far behind the laboratories. Scholars, therefore, report their results to conferences or learned societies and duplicate a certain number of copies of their papers, which may not appear in print for 2 years or more, if ever. Obviously, these circulating papers should be annotated and listed if possible. But it is very difficult to know when one has seen them all; and in some cases a reviewer knows of the existence of papers, but has been unable to see them. Any such review as this, therefore, must necessarily be less than complete. We have tried to make it as complete as possible until about February of 1963, and the bibliography includes all the research reports we have found with the exception of a few whose authors asked us to omit them because they were "early work" or otherwise better forgotten. Nevertheless, there probably are some papers that should have been included in this list that are not, and for this we apologize to the authors and the readers.

What does the research say? Let us begin with the question most often asked:

Do Students Learn From Programed Instruction?

The research leaves us in no doubt of this. They do, indeed, learn. They learn from linear programs, from branching programs built on

the Skinnerian model, from scrambled books of the Crowder type, from Pressey review tests with immediate knowledge of results, from programs on machines or programs in texts.[4] Many kinds of students learn—college, high school, secondary, primary, preschool, adult, professional, skilled labor, clerical employees, military, deaf, retarded, imprisoned—every kind of student that programs have been tried on. Using programs, these students are able to learn mathematics and science at different levels, foreign languages, English language correctness, the details of the U.S. Constitution, spelling, electronics, computer science, psychology, statistics, business skills, reading skills, instrument flying rules, and many other subjects. The limits of the topics which can be studied efficiently by means of programs are not yet known.

For each of the kinds of subject matter and the kinds of student mentioned above, experiments have demonstrated that a considerable amount of learning can be derived from programs; this learning has been measured either by comparing pre- and post-tests or the time and trials needed to reach a set criterion of performance. But the question, how well do students learn from programs *as compared to how well they learn from other kinds of instruction*, we cannot answer quite so confidently.

Experimental psychologists typically do not take very seriously the evaluative experiments in which learning from programs is compared with learning from conventional teaching. Such experiments are doubtless useful, they say, for school administrators or teachers to prove to themselves (or their boards of education) that programs work. But whereas one can describe fairly well the characteristics of a program, can one describe the characteristics of a classroom teaching situation so that the result of the comparison will have any generality? What kind of teacher is being compared to what kind of program? Furthermore, these early evaluative experiments with programs are likely to suffer from the Hawthorne effect: that is to say, students are in the spotlight when testing

---

[4] These last two categories present another example of what nonsignificant differences do not tell us. The experimental comparisons of learning from machine programs and learning from programed texts are reviewed carefully by Goldstein and Gotkin, in the *Journal of Programed Instruction*, 1, 1 (1962), 29–36, and the verdict is clear and unanimous: n.s.d. Yet the important question is really not yet answered. These studies compare the crude teaching machines of today with today's programs. What could we accomplish if we really turned out technological skills to making better teaching machines? We know, for example, that it is possible to present a much wider variety of stimuli by means of a computer serving as a teaching machine, in fact to program the computer so that it will learn to be a better teacher and to allow for the individual needs and capabilities of the student in a way that printed programs could hardly be expected to do. On a simpler level, we know that machines have advantages over printed programs in some respects—in repeating incorrectly answered items, for example, in presenting programs to students not yet able to read, and in presenting moving stimuli. The question is, what are the *special* advantages of machines or programed texts, and how can we take advantage of them?

something new, and are challenged to do well.[5] It is very hard to make allowance for this effect. Therefore, the evaluative tests may be useful administratively, say many of the experimenters, but do not contribute much to science, and should properly be kept for private use.

These objections are well taken. And yet, do they justify us in ignoring the evaluative studies? The great strength of a program is that it permits the student to learn efficiently by himself. Is it not therefore important to know how much and what kind of skills, concepts, insights, or attitudes he can learn by himself from a program as compared to what he can learn from a teacher? Admittedly, this is a very difficult and complex research problem, but that should not keep us from trying to solve it.

In the light of those remarks, let us record some of the evidence comparing programs with conventional classroom instruction. We have tabulated 36 such reports. Sixteen of them were done in colleges, 4 in secondary school, 5 in primary school, 10 with adults, and 1 with retarded children. Of these 36 comparisons, 18 showed no significant difference when the two groups were measured on the same criterion test. But 17 showed a significant superiority for the students who worked with the program, and only 1 showed a final superiority for the classroom students. Eight of the experimenters mentioned a time advantage for the program students, and one (an industrial user) a cost advantage (see TEACHING MACHINES, INC., 1962, and KOPSTEIN AND CAVE, 1962a, on cost factors). These results, of course, must be considered in view of the possible Hawthorne effect, and of the possibility that the criterion tests (which are often the final examination in a course) do not measure *all* the significant outcomes. But, on the whole, the results should not discourage us about the amount of learning derived from programs.

The Characteristic Variables

If we had tried, 4 years and 165 research papers ago, to describe the characteristics of a presumably efficient linear program, we should probably have said that it has (1) an ordered sequence of items, through which the student works in (2) "short steps," therefore (3) making few errors, as he records (4) a constructed response to each item, and receives (5) immediate knowledge of results. The student

[5] Unfortunately the Hawthorne effect doubtless operates in many other experiments also—especially short ones, in which the program is viewed as essentially a test, challenging the student to do his best. On novelty effect, see POPHAM, 1962.

(6) works at his own pace, and (7) receives reinforcement for each correct response.

After 5 years, what does the research say about those characteristics of Skinnerian program design?

An Ordered Sequence

There have been five rather interesting experiments comparing a logically sequenced program with a presentation of the same items in random order. Strangely enough, three of these experiments showed no difference (see ZUCKERMAN, MARSHALL, AND GROESBERG, 1961; ROE, K., 1961; LEVIN AND BAKER, 1962). A fourth experiment (see ROE, A., 1962) showed significantly more learning for students using the logically sequenced program, and a fifth experiment (see GAVURIN AND DONAHUE, 1960) found that students who worked with the logically sequenced program made fewer errors during trial and required fewer trials to reach a set criterion, but 1 month later scored no higher in a retention test than did the students who worked with the randomized sequence. These results are somewhat puzzling, but it must be remembered that the experiments were done with very short programs. The longest of the five was 180 frames (and showed n.s.d.). All the others were less than 100 frames, the shortest being only 29 frames (this was the one that showed superiority for the logical order on criterion test, but no difference on retention). It is reasonable to suppose that longer programs would have greater need of ordered sequence.

A more positive contribution to the understanding of sequence has been made by Gagné and his collaborators, who have analyzed a number of tasks (e. g., solving algebraic equations) and have broken them down into hierarchies of subordinate learning sets (see GAGNÉ, 1961; GAGNÉ AND PARADISE, 1961; GAGNÉ, MAYOR, GARSTENS, and PARADISE, 1962; and SILBERMAN, COULSON, GUNN, AND MELARAGNO, 1962). Subordinate learning sets for a given class of tasks are defined as the answer to the question, If he were given instructions only, what would the individual have to know how to do in order to be able to perform this (new) task? Beginning with the final task, the question is applied successively to each subordinate learning set, and thus it is possible to identify a hierarchy of learning sets which grow increasingly simple and general the further they are from the final task. Gagné has tested this kind of analysis with encouraging success. It seems to provide a learning-based "logic" for designing programs, and also to make it possible for any student to begin "where he is," because his mastery of any level of learning set can quickly be tested.

MAGER (1961) made another thought-provoking contribution to the area of sequencing when he permitted six trainees in electronics to control their own learning program throughout a course. They were taught individually. At the beginning they were told only that they were to learn as much as possible about electronics, and were given some examples of the kind of end behavior to be expected of them. Then they were permitted to ask any questions they wished. All of the students learned a great deal about electronics, but the instructors noted that the sequence by which they proceeded bore little resemblance to the "logical" sequence in which the course was usually taught. Whereas the instructors usually worked from parts to wholes, these students typically moved from smaller wholes to larger wholes. It was noted that these students had very high motivation, apparently because they were in control of their own program.

Short Steps and Few Errors

Size of step has never been quite satisfactorily defined. In some cases it has been expressed as the reciprocal of the number of steps used to cover a given body of material, but a program with fewer steps does not necessarily require longer leaps; it may merely have less practice or fewer examples. Size of step has also been measured as the amount of material in a frame or an item; thus, intrinsic programing would typically be said to have very long steps. In other cases, step size has been measured in terms of the average number of errors made in the program; this is apparently based on the circular argument that long steps should produce more errors, and therefore if there are more errors there must be long steps.

However, when significant differences have been found in learning from programs of different step size (measured in *any* of these ways), they have usually been in favor of the programs with short steps. Thus, for example, EVANS, GLASER, AND HOMME (1960), using 4 alternate versions of a program, taught 20 graduate students how to convert numbers to bases other than 10. These alternate versions contained respectively 30, 40, 51, and 68 steps. The students with 51 and 68 steps did significantly better than the others on both an immediate post-test and a delayed retention test. COULSON AND SILBERMAN (1960) got similar results. SHAY (1961), inferring step size from number of errors made on given items, found no significant relation between intelligence and step size in amount of pupil learning. But SMITH AND MOORE (1961) found no significant differences in the rate of learning to spell when pupils were taught 166 words by means of programs of 1,128, 830, and 546 steps, respectively. The task of learn-

ing to spell may be essentially different from that of learning to convert to unfamiliar number bases (EVANS, GLASER, AND HOMME) or of learning psychology items (COULSON AND SILBERMAN).

An ingenious way to measure step size was employed in a series of experiments with training films (see MACCOBY AND SHEFFIELD, 1958; MARGOLIUS AND SHEFFIELD, 1961; and WEISS, MACCOBY, AND SHEFFIELD, 1961). The experimenters simply varied the length of film sequence before permitting practice. They found that more learning came from gradually *increasing* the step size than from maintaining either very short or very long steps. They also discovered that when the students were permitted to select their own size of step before practice, they chose a gradually increasing length. These experimenters noted that performance gradually deteriorated among students who were permitted to practice short steps only. This observation may help to explain some of the reports of impatience and boredom which have come in on occasion from students who have worked through long programs.

Constructed Response

On the question of whether students should write out their response, rather than merely "thinking" it or selecting one from a multiple choice of answers, the evidence is not clear. The great majority of the studies find no significant differences between the amount of learning from overt and covert responses (see, for example, ALTER AND SILVERMAN, 1962; EVANS, 1960; EVANS, GLASER, AND HOMME, 1960; FELDHUSEN AND BIRT, 1962; GOLDBECK AND CAMPBELL, 1962; GROPPER AND LUMSDAINE, 1961c; HUGHES, 1961; KAESS AND ZEAMAN, 1960; KANNER AND SULZER, 1961; KEISLAR AND MCNEILL, 1962; KORMONDY, 1960; LAMBERT, MILLER, AND WILEY, 1962; MICHAEL AND MACCOBY, 1953; ROE, MASSEY, WELTMAN, AND LEEDS, 1960; SILVERMAN AND ALTER, 1961; and STOLUROW AND WALKER, 1962). Since in most cases, the covert responses take less time, it is possible to say that the covert response mode is more "efficient."[6]

Let us look at the minority of studies in which overt or covert response *does* seem to make a difference. CUMMINGS AND GOLDSTEIN (1962) found that a group which wrote answers scored higher on both immediate and delayed tests than a group which was told merely to "think" its answers. Their program was 110 frames, teaching a medical diagnosis, and it may be that the complexity of the subject

[6] Taking into account both time and amount of learning, SIDOWSKI, KOPSTEIN, and SHILLESTAD (1961) found that covert responders scored higher than overt ones on an "efficiency" measure. MCGUIRE (1961) found covert participation superior to overt at a fast rate, and superior to no participation at both fast and slow rates. Overt participation, on the other hand, resulted in more learning than no participation at the slow rate, but less than either covert or no participation at a fast rate.

matter made for an advantage to the students who took longer and practiced writing the response. KRUMBOLTZ AND WEISMAN (1962b) found no significant difference on an immediate post-test between a group which wrote answers and a group which only composed answers mentally; but on a test 2 weeks later there was a significant advantage to the overt responders. It may be, therefore, that the additional practice of writing an answer contributes to the retention of some subject matter (KRUMBOLTZ AND WEISMAN used a program on educational testing). SUPPES AND GINSBERG (two studies in 1962) found that a group of 5- and 6-year-olds who were required to make a correct response after each error, rather than merely being told the correct answer, did better on a post-test than a group which was told the answer but not required to repeat it. There is support elsewhere in the literature for the view that it is important for the student to actively make a correct response (overt or covert) before going on to the next step (see ANGELL AND LUMSDAINE, 1961a). GROPPER AND LUMSDAINE (four studies in 1961) found that active response made no difference when a televised lesson was not "programed," but when it was sequenced like a teaching machine program, active response made for significantly higher scores. Thus, we have at least a few guidelines as to when and under what conditions active responses may be important.

The studies comparing the practice of active response versus the mere reading of items (usually with the responses filled in) are somewhat equivocal. HOLLAND (1960) found that a group which merely read complete statements made more errors than either active or covert responders on a final test, but also took less time to complete the program. In two experiments, SILVERMAN AND ALTER (1962) found nonsignificant differences between a group which read complete items and another group which read and then responded actively. On a third experiment, however, they found significant differences in favor of the group that merely read the items. FELDHUSEN AND BIRT (1962) and GROPPER AND LUMSDAINE (1962c) found nonsignificant differences between reading and responding groups, although the latter two authors found the responding group superior when the television lesson was programed. GOLDBECK AND CAMPBELL (1962) tested materials at different levels of difficulty. They found that overt responding was superior to mere reading at an intermediate level of difficulty, but, of course, took more time. There were no significant differences at the most difficult and easiest levels. In a second experiment, however, they found the reading group superior on a 10-week retention test.

PRESSEY (1963) tested the first unit of the Holland-Skinner psychology program (only 54 frames) against the same material rewritten

into good prose paragraphs. He found no significant differences in learning, although reading the prose took less time than working through the program. When he added some of his typical review questions to the prose selections, the amount of learning was higher, but non-significantly so, than the learning from the program.

These results are going to cause researchers in this field to do a great deal of thinking about the principles behind programed instruction. For one thing, it is undoubtedly true that human beings are capable of learning in ways other than the step-by-step conditioning which is characteristic of the Skinner program. Again, it may be that when one reduces the step size and error level to a minimum, then overt responding is hardly necessary. Again, in *some* situations, with *some* subject matter, the additional practice of overt responding must be helpful; one example may well be foreign language study.

The comparisons of constructed versus multiple-choice response have so far indicated no clear superiority for either one, although there must be *some* subject matters and *some* learning tasks for which one or the other method works better.[7] Ordinarily, multiple-choice responding saves some time.

HOLLAND (1960) reports an interesting finding concerning the *cruciality* of the required response. He prepared three versions of a program in addition to the original version. One of the new versions left blanks for responses of a trivial and easy nature; a second left blanks which made the items ambiguous and difficult; a third filled in the blanks so that the student read complete statements. Thus, he was able to compare the original version, in which the blanks required the student to notice the critical material which he was supposed to learn, with the other three versions. The group using the normal version of the program did better than the others. This appears to demonstrate that items should be so written that the ability to respond correctly to them depends on the student's noticing the critical information in the item (see also KRUMBOLTZ, 1963).

Immediate Knowledge of Results

The majority of the studies support the idea that immediate knowledge of results contributes to learning. Four studies found no significant differences attributable to knowledge of results (FELDHUSEN AND BIRT, 1962; HOUGH AND REVSIN, 1963; MCDONALD AND ALLEN, 1962; and MOORE AND SMITH, 1961), and one found no significant

[7] For example, COULSON AND SILBERMAN (1960), HOUGH (1962a), BURTON AND GOLDBECK (1962), and ZUCKERMAN, MARSHALL, AND GROSSBERG (1961) all found n.s.d. PRICE (1962), teaching arithmetical skills to mentally retarded students, found in certain situations a slight advantage for multiple choice; FRY (1960), teaching Spanish vocabulary, found an advantage for constructed responses, although they required significantly greater training time.

difference between immediate and slightly delayed knowledge of results (EVANS, 1960). On the other hand, ANGELL (1949) using Pressey-type questions, and MEYER (1960) using a linear program, both found significantly more learning by a group that received immediate knowledge than by a group that waited for the next meeting of the class to find out the results. MICHAEL AND MACCOBY (1953) found that immediate knowledge of results contributed significantly to learning from an instructional film with questions inserted.

A number of other studies have also reported significant differences in favor of immediate knowledge of results, but have been concerned chiefly with the method of giving the knowledge. In one experiment, BRYAN AND RIGNEY (1956) found that a group which received knowledge of results *plus an explanation* did better on a test 1 week later than groups which received no knowledge of results or knowledge of results without explanation. BRYAN, RIGNEY, AND VAN HORNE (1957) followed up this earlier finding, but could find no difference in the amount of learning resulting from three forms of explanation: A correct definition or description, the reason why a chosen alternative was the correct one, or the probable consequences of action represented by a chosen alternative. KRUMBOLTZ AND BONAWITZ (1962) found that a group which received knowledge of results in the context of complete sentences was better able to *apply* the principles learned than a group that received only the correct response. KANNER AND SULZER (1961) experimented with covert responses, and found that a covert response plus feedback of results brought about more learning than a covert response *without* feedback.

GLASER AND TABER (1961) found no significant difference in learning from receiving 100 percent continuous knowledge of results, from receiving such knowledge only 50 or 25 percent of the time, or from receiving it in a variable ratio. They suggest cogently that knowledge of results is doubtless more important when the probability of error is high. When the probability of error is kept low, as in a typical linear program, it becomes less important to have immediate knowledge of results.

At His Own Pace

To most of us, it makes sense intuitively that a student will learn more efficiently at his own pace. Somewhat surprisingly, the experimental literature has not been able to demonstrate as much advantage for individual pacing as might be expected. FOLLETTIE (1961) found self-pacing better on an efficiency measure incorporating test score, training time, and testing time. MACCOBY AND SHEFFIELD (1958) found that self-pacing worked best for superior students in learning

783-259 O—64——2

from a training film interspersed with practice. On the other hand, no fewer than seven studies have found no significant difference attributable to individual or external pacing, whether the students are taught by teaching machines, programed texts, or television (see CARPENTER AND GREENHILL, 1963; BRIGGS, PLASHINSKI, AND JONES, 1955; ALTER AND SILVERMAN, 1962; FELDHUSEN AND BIRT, 1962; MOORE AND SMITH, 1961; SILVERMAN AND ALTER, 1961; and BRIGGS, 1961b).

Indeed, one of the significant findings seems to be that it is possible to teach efficiently with programed materials on television or films.[8] Carpenter and Greenhill compared an externally paced television program with self-paced teaching machine programs in three experiments, and externally paced films with a self-paced programed text in another. In each case they found no significant difference attributable to the pacing. In one experiment they found it possible to vary the pace 20 percent below and 10 percent above the average of class self-pacing without significantly decreasing the amount of learning. GROPPER AND LUMSDAINE (1961a) also found evidence that a lesson on television could be programed with active responses, and would result in more learning than a nonprogramed television lesson.[9]

FRYE (1963) contributed to the further understanding of the use of external pacing by comparing groups that were more or less homogeneous in ability. He found that a heterogeneous group took longer on the average to master a program when it was externally paced than when it was individually paced. A homogeneous and externally paced group, however, took no longer than a self-paced group.

Reinforcement

When Skinner applied his experience with animal training to human learning, he defined response-confirmation as reinforcement in the sense that it increases the probability that desired behaviors will appear at the proper time.[10] In other words, to tell a student that his answer is right is a form of reinforcement. This is one reason

[8] KIMBLE AND WULFF (1961a) demonstrated that a motion picture, when so designed as to require the student to make guided responses and to minimize error, resulted in more learning than unguided attention to such a picture (see also MICHAEL AND MACCORY, 1953, MACCORY, MICHAEL, AND LEVINE, 1961, HOVLAND, LUMSDAINE, AND SHEFFIELD, 1949; and KANNER AND SULZER, 1961).

[9] KLAUS AND LUMSDAINE (1960) demonstrated that a program could be used effectively to supplement instructional television (see also SHUNNY, 1961; PETERSON, 1931; GOLDBECK, SHEARER, CAMPEAU, AND WILLIS, 1962; BARCUS, HAYMAN, AND JOHNSON, 1963; HATCH AND FLINT, 1962; and HICKEY, AUTOR, AND ROBINSON, 1962, all of whom are concerned with programed materials as supplements.

[10] See SKINNER'S 1954 article, previously cited, and also his "Teaching Machines," *Science*, 128 (1958) 969–77. See also the thoughtful presentation by UGELOW, A. *Motivation and the Automation of Training: A Literature Review*. Technical Documentary Report MRL-TDR-62-15. Wright-Patterson Air Force Base, Ohio: Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, Aerospace Medical Division, Air Force Systems Command, 1962.

why it has seemed desirable to keep the error rate low. However, Skinner's own experience with schedules of reinforcement in animal learning would lead him to prefer a variable-ratio schedule of reinforcement—a high reward occasionally, small rewards fairly frequently, and, after many trials, an absence of reward.[11] Amsel has suggested that reinforcement should be constant during the period when the response is being shaped and should be variable afterward, so that the persistence of the response can be insured.[12]

If such information is to be rewarding, however, it must meet some need or reduce some drive in the student. Gagné and Bolles have suggested that motivation—by which they presumably mean the desire to achieve, the desire to learn, the desire to solve a problem, or something of the kind—may be intrinsic to the task of learning. Motivation may also be built up by the teacher, by competition, or by the knowledge of some such delayed reward as a better job, admission to college, or passing the course. Or it may be built in, to some extent, by skillful writing in the program.

The experiments so far have not been too successful in identifying the incentive in programed instruction that supposedly makes response-confirmation an act of reinforcement. For example, when MOORE AND SMITH (1962) varied the response-confirmation from (a) zero to (b) knowledge of the correct response to (c) a flashing light for each correct response to (d) a small monetary reward for each one, they found no significant differences. ALTER, EIGEN, AND KING (1962) likewise found n.s.d., using trinkets as reinforcers. The studies of attitudes are not very helpful, although, in general, students are more often reported bored with long programs than with short ones, and with continued use of short steps than with increasing size of step or longer steps (see FELDHUSEN, RAMHARTER, AND BIRT, 1962; REED AND HAYMAN, 1962; NAUMANN, 1962; VAN ATTA, 1961; and EIGEN, 1963).

There have been a number of studies on prompting versus confirmation. In a confirmation mode, the stimulus is presented first, then the student writes a response, and finally he is told the correct response. In a prompting mode, the stimulus and correct response are shown simultaneously to the student, after which he repeats the response if that is part of the procedure. Although a very large proportion of these prompting studies have dealt with paired associates (for example, Russian-English word pairs or the military phonetic alphabet) rather than concept learning or problem-solving, still it is

[11] FERSTER, C. B., AND SKINNER, B. F. *Schedules of Reinforcement*. New York: Appleton-Century-Crofts, 1957.

[12] AMSEL, A. "Error Responses and Reinforcement Schedules in Self-Instructional Devices," in LUMSDAINE, A. A., AND GLASER, R., eds. *Teaching Machines and Programed Learning*. Washington: National Education Association, 1960. See also UGELOW, op. cit.

very impressive to note that almost all of them have shown a superiority for prompting. These studies include several in which Cook was either senior or sole author (Cook, Miller, Grier, and Staman, 1962; Cook and Brown, 1963; Cook and Kendler, 1958; and Cook, 1958) and also studies by Irion and Briggs (1957), Briggs (1961), Kopstein and Roshal (1955 and 1961), Silberman, Melaragno, and Coulson (1961), Shettel and Lindley (1961), Sidowski, Kopstein, and Shillestead (1961), and Stolurow and Lippert (1961). All of them showed significant superiority for prompting over confirmation. Certain other studies qualified the finding. For example, Angell and Lumsdaine (1961a) found partial prompting (3 prompting items followed by 1 response-confirmation item) better than prompting alone (see also Guthrie and Lumsdaine, 1961). Stolurow (1961), in teaching vocabulary to retarded children, found that with a shorter practice period prompting was superior to confirmation, but, with a longer time for practice, confirmation proved superior to prompting in bringing about retention of what was learned.

These studies indicate that prompting is a very powerful method for some kinds of learning. They also implicitly raise the question of whether the contiguity hypothesis (notably in the form advanced by Guthrie) does not describe better than the reinforcement hypothesis the process by which learning takes place in programed instruction (Lumsdaine also proposed the contiguity hypothesis as a model for programed instruction in a 1961 paper).

Some Other Topics

This review is necessarily incomplete and fragmentary. Many interesting studies and findings are not mentioned, and some topics that have concerned scholars in this field are not touched on. If space permitted, we might talk, for example, about studies of branching (see, for example, Campbell, 1961 and 1962; Coulson, Estavan, Melaragno, and Silberman, 1962; Roe, A., 1962; and Beane, 1962), which has been used successfully to modify linear programs, although the art of its use is still in an early stage; or studies of individual differences (see Lumsdaine, Sulzer, and Kopstein, 1961; Feldhusen and Eigen, 1963; Lambert, Miller, and Wiley, 1962; and Price, 1962), in which students of almost any ability from the mentally retarded to those with a very high IQ seem to learn from programs, but not a great deal has been done yet to understand how to modify programs for different levels of ability; or studies of review (see Holland and Porter, 1961; Lumsdaine, Sulzer, and Kopstein, 1961; Dowell, 1955; Ferster, 1960; Glaser and Reynolds, 1962; Kimble and Wulff, 1961b; Rothkopf and Coke, 1963; and

SHEFFIELD, MARGOLIUS, AND HOEHN, 1961), which is necessary with this method as with other methods, although there are few guides except empirical ones on how much is necessary; or studies of cueing and vanishing (see ANGELL AND LUMSDAINE, 1960, 1961b, and 1962; GLASER AND TABER, 1961; and MCNEILL AND KEISLAR, 1962), among which should be noted the study by Angell and Lumsdaine in which the effect of vanishing of cues showed up only after 2 weeks, suggesting that delayed retention tests may in some cases be necessary to evaluate programs.

But let us conclude by noting two small but provocative studies, respectively, by MCNEILL (1962) and ROTHKOPF (1963). McNeill's study had for its subjects 132 kindergarten children, 91 of whom were later studied as first-graders under female teachers. Under programed instruction in word recognition, the boys in this sample did significantly better than the girls; under female teachers in the classroom, the girls did significantly better than the boys. The author suggests that perhaps female teachers in the early grades fail to adjust themselves or their teaching procedures as well to the traits of boys as to those of girls, and raises the possibilities that greater use of programed instruction in the early grades might be beneficial to boys and that a study of the features of auto-instruction might help in developing teacher behavior more appropriate for boys.

Rothkopf's study had for its subject 12 high school teachers or principals who had just experienced a summer seminar in programed instruction. They were shown seven versions of a short program and asked to predict the relative effectiveness of these versions. The *actual* effectiveness had previously been determined by testing with subjects. The rank-order correlation between the empirically determined effectiveness and the predictions of the high school teachers and principals was MINUS .75!

WILBUR SCHRAMM,
Institute for Communication Research,
Stanford University.

The Bibliography

Note. The following annotations are listed alphabetically, by senior author. The annotations have been submitted to their authors for checking (or to one author, if there were several) in every instance where an address was available for the author. In some cases, the manuscript was not returned, and in some cases it may have gone astray. For any errors that, despite this precaution, have crept into the annotations and for any omissions of useful facts that have been dictated by the policy of short annotations, we apologize to the authors and the readers. We have tried to make the listing as complete as possible, except for articles which authors have asked us to omit. However, in a field where the rate of publishing has been as great as it has been in this one in the last few years, and where at least a fourth of all the papers are circulating in mimeographed or offset form, it must be expected that some titles will be omitted. For this, too, we apologize in advance, and hope that authors of omitted articles will send us copies, reprints, or references, so that these additional titles may appear in later listings.

ALTER, MILLICENT. *Retention in Programed Instruction.* New York: Center for Programed Instruction, 1962.

Problem: Measures that predict retention of programed material.
Sample: 236 students in grades 7 through 11.
Program: 235 frames; linear program on sets, relations, and functions.

After working through the program, subjects were tested twice: immediately and again after a retention interval of 2 to 30 weeks. More intelligent students performed better on the retest than less intelligent students. When immediate post-test performance was held constant, this difference was greatly reduced, but still significant. Faster workers performed better than slower workers on the retest, but when immediate post-test performance was held constant, this difference disappeared. In addition, no significant differences were found in the contours of the retention curves of high, middle, and low immediate post-test students; students of high, middle, and low intelligence; or fast, middle, and slow workers.

ALTER, MILLICENT; EIGEN, LEWIS; and KING, SHIRLEY. *The Effectiveness of Confirmation Plus Trinket Reinforcers in Young Children.* New York: Center for Programed Instruction, Inc. 1962. 22 p.

Problem: Effect of different kinds of reinforcement.

Sample: Sixteen 5- and 6-year-old children.

Program: Preliterate program designed to teach numerals and concepts of one-ness through nine-ness, by means of matching pictures showing different numbers of stars, scissors, balls, and so forth.

In both experimental groups, correct responses were confirmed by telling the students they were correct. One group was, in addition, given small trinkets for the first five correct responses, and thereafter, for about 25 percent of the correct responses. No significant differences in learning or test performance were found.

Control data suggested that the program was not teaching what it was designed to teach. As a result of this, the goals and techniques of the program were discussed critically.

ALTER, MILLICENT, and SILVERMAN, ROBERT. "The Response in Programed Instruction." *Journal of Programed Instruction*, 1, 1 (1962), 55–78.

Experiment I:

Problem: Covert, spoken, written, and written-spoken responses, compared with one control group that read the program in the form of statement, and another control group that was not given the material.

Sample: 90 university students.

Program: 87 frames on basic electricity; linear; constructed response.

Results: All groups that read the program learned significantly more than the uninstructed group, but there was no significant difference among them on a post-test.

Experiment IIa:

Problem: Written responding vs. reading; external pacing vs. self-pacing, in a factorial design.

Sample: 60 university students.

Program: 87 frames on basic electricity; linear; constructed response.

Results: Neither main effects nor interaction significant on total test, but on constructed response portion, reading was significantly superior to written responding.

Experiment IIb:

Problem: Written responding vs. reading; teaching machine vs. programed text, in a factorial design.

Sample: 60 elementary psychology students.

Program: 90 frames on binary numbers; linear; constructed response.

Results: Neither main effects nor interaction significant on a post-test.

See also SILVERMAN AND ALTER, 1961.

ANGELL, DAVID, and LUMSDAINE, ARTHUR A. *Prompted Plus Unprompted Trials versus Prompted Trials Alone in Paired-Associate Learning.* Pittsburgh: American Institute for Research, 1960. 19 p. (mimeo)

Problem: Complete vs. partial prompting.

Sample: 48 volunteers, 16 to 35 in age.

Program: 10 paired-associate items—letters paired with arbitrary 3-line figures.

Using the same material as COOK AND KENDLER (1958), these investigators compared a condition in which the subject was prompted on every item with one in which he responded without prompting on every fourth item. The condition of partial prompting resulted in significantly more efficient learning. This condition was identical with the "prompting" condition which Cook had found superior to "confirmation." The results supported the theoretical expectation that the learner should not only be helped to respond correctly (by being prompted) but should also be given some practice in responding without the help of prompts. A similar argument underlies the rationale for "vanishing." See also ANGELL AND LUMSDAINE, 1962.

ANGELL, DAVID, and LUMSDAINE, ARTHUR A. *The Effects of Prompting Trials and Partial-Correction Procedures on Learning by Anticipation.* San Mateo, Calif.: American Institute for Research, 1961a. 47 p. (mimeo)

Problem: Prompting vs. confirmation, in different conditions of correction.

Sample: 25 high-school students.

Program: 16 lists, each consisting of 12 English words paired with digits; and (second experiment) lists of 16 pairs of French and English words.

Using pairings of English words with digits, found that prompting was more effective than confirmation when partial correction was used (e.g., subject given two or more possible response terms and told one of them is right), but *not* when full correction is used; and full cor-

rection was superior to partial correction when there had been no initial prompting trials, but not when there had been two or more prompting trials.

A second experiment, using French and English word pairs, and requiring the subjects in partial correction situations to keep on responding until they found the right response, failed to show a significant difference between prompting and confirmation.

Authors suggest that trainers should consider prompting a powerful procedure for many kinds of paired-associate learning, and also that it is important to get correct response from learner before going on to next item.

ANGELL, DAVID, and LUMSDAINE, ARTHUR A. *A Study of Subject-Controlled Partial Cueing in Paired Associate Learning.* San Mateo, Calif.: American Institute for Research, 1961b. 13 p. (mimeo)

Problem: Effectiveness of partial cueing.

Sample: 40 eleventh- and twelfth-graders.

Program: 12 city names paired with their airline code letters.

Subjects were tested so as to compare learning with standard confirmation technique (as in typical anticipation procedures) and partial cueing. By partial cueing is meant here that the student could see at will one, two, or three letters of the three-letter symbol which constituted the response term. Little overall difference was found between the two methods. When the items were dichotomized into difficult and easy ones, the partial-cueing procedure was found to be somewhat more effective for slow learners with hard items, somewhat less effective for fast learners with easy items (interaction, $P. < 10$).

ANGELL, DAVID, and LUMSDAINE, ARTHUR A. *Retention of Material Presented by Autoinstructional Programs Which Vanish and Which Do Not Vanish Verbal Cues.* Palo Alto, Calif.: American Institute For Research, 1962. 13 p. (mimeo)

Problem: Advantage of "vanishing" of prompts.

Sample: 174 fourth-, fifth-, and sixth-grade students.

Program: Alternate forms of a 54-frame program to teach a short-cut method for squaring numbers ending in 5.

Individual students in several classes were randomly assigned either to (1) a version of the program that employed "vanishing" (progressive reduction) of prompts, in more or less customary linear-program style, or to (2) a "nonvanishing" version, in which full prompting was continued throughout the program. When students were tested immediately after finishing the program, scores for the

two versions failed to differ significantly. However, scores on a delayed retention test (2 weeks later) were significantly higher for the vanishing than for the nonvanishing treatment. The fact that effects of a theoretically important factor such as vanishing showed up only after a lapse of time (see also ROTHKOPF, 1962), suggests that delayed retention tests may be needed for proper evaluation of programs.

ANGELL, G. W. "The Effect of Immediate Knowledge of Quiz Results and Final Examination Scores in Freshman Chemistry." *Journal of Educational Research*, 42 (1949), 391–94.
Problem: Immediate vs. delayed knowledge of results.
Sample: 162 college students.
Program: Test items on chemistry, designed for punchboard.

Students who secured immediate knowledge of results through use of the Angell and Troyer punchboard were compared with students who learned of results, through IBM score sheets, at the next meeting of the class. Final examination scores were significantly higher for students getting immediate knowledge of results.

ARNOULT, M. D. *A Comparison of Training Methods in the Recognition of Spatial Patterns*. Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, 1956.
Problem: Reproduction vs. verbal practice of an essentially visual task.
Sample: Air Force basic trainees.
Program: Spatial patterns composed of meaningful or nonsense elements.

Several different methods were used to train basic airmen to recognize certain complex visual patterns. Comparing effect of responses constructed by student (i.e., drawn from memory) with responses consisting of verbal descriptions of stimulus, former were found to be more effective.

ASHER, JAMES J. *Sensory Interrelationships in the Automated Teaching of Foreign Languages*. San Jose, Calif.: Psychology Department, San Jose State College, 1962. 99 p. (mimeo).
Problem: Effects of foreign language vocabulary learning through learning first from visual vs. auditory materials.
Sample: 80 college undergraduates.
Program: 92 Spanish words with English equivalents.

One subsample learned the Spanish words first visually, then relearned them from auditory stimuli. The other subsample learned first from auditory material, relearned from visual. In each case,

the following comparisons were tested: (a) paired associate vs. recognition, (b) pictures vs. words as stimuli, (c) simultaneous vs. sequential presentation. Active acquisition was required in that all responses had to be in the foreign language. Throughout all three conditions listed above, superior learning was found to result from learning the words first visually. Superior performance means (a) significantly less perseverative error in initial learning, (b) significantly less unique error in relearning, (c) significantly greater transfer to syntactic comprehension.

A one-page review of this article can be found under the same title in *Perceptual and Motor Skills*, 14 (1962), 38.

AUSTWICK, KENNETH. *Programmed Learning—Some First Impressions.* In *Teaching Machines and Programmes*, a bulletin prepared by the Department of Psychology, University of Sheffield. Sheffield, England: January 1962 (mimeo).

Problem: Learning and retention, programed vs. conventional instruction.

Sample: 2 groups of 18 English secondary school students, matched on IQ and Ballard arithmetic tests.

Program: Algebra.

Experimental group taught by simple program; control group, by conventional classroom methods. Author says this was his first program, which may explain why immediate post-test scores were better for control group. However, retention several weeks later was relatively better for experimental group than for controls. In general, the experimental group scored as well and retained more on questions actually covered in the program, but did less well in applying learning to new questions not covered in the program. Experimenter found also that making few errors in the program did not necessarily predict high scores on the post-test; he speculates that what was understood at the time may not have been understood well enough for the student to apply it.

BARCUS, DELBERT; HAYMAN, JOHN L., JR.; and JOHNSON, JAMES T. "Programing Instruction in Elementary Spanish." *Phi Delta Kappan*, 44, 6 (1963), 269–72.

Problem: Effectiveness of programed text vs. teaching machines vs. conventional classroom instruction, under control of teachers with varying amounts of training and experience, and with timing in the course as a factor.

Sample: 6,000 sixth-grade students.

Program: 2,200 frames; linear program on reading and writing Spanish.

Half the students were taught by program after a year of oral-aural Spanish, the other half after three semesters. Results indicated that the time of introducing the program was an essential element in the result: in the first semester the students getting automated drill did significantly less well than the students getting conventional instruction, whereas in the second semester there was no significant difference between the conventionally taught group and the group using the programed text, but the group using the teaching machine did significantly better than either. The level of teachers' training was directly proportional to the amount of learning from the programed text, indicating, the authors believe, that motivation supplied by the teacher is an important element in the students' individual work.

BEANE, DONALD G. *A Comparison of Linear and Branching Techniques of Programed Instruction in Plane Geometry.* Urbana, Ill.: University of Illinois, 1962. (mimeo)

Problem: Comparison of a linear and a branch program covering a unit on parallel and perpendicular lines in plane geometry.

Sample: Two experimental classes and one control class (65 students) in high school.

Program: Branch and linear programs on "Parallel and Perpendicular Lines;" linear program contained 951 frames in 5 booklets; branch program contained 852 frames in 7 booklets.

Control group had conventional classroom teaching from the text mentioned. Experimental groups were divided into four groups, each receiving different treatment: two groups used the linear and branch programs exclusively and two groups switched from one program to the other halfway through the experiment. All five treatments resulted in a significant amount of learning during the experiment; in each treatment group the high-ability students exceeded the low-ability students in achievement and retention. The branch program was more efficient than the linear program timewise. The students expressed attitudes more favorable to the linear program.*

BENSON, EUGENE W., and KOPSTEIN, FELIX F. *Machine Teaching of Basic Electronics at Keesler Air Force Base: An Experiment.* Paper presented at convention of Department of Audio-Visual Instruction, National Education Association, 1961. 5 p. (mimeo)

See KOPSTEIN and CAVE (two studies in 1962).

BIRCH, JACK W., and STUCKLESS, E. ROSS. *The Development and Evaluation of Programed Instruction in Language for Children*

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

with Auditory Disorders. Pittsburgh, Pa.: University of Pittsburgh, 1962 (mimeo).

Problem: Programed instruction vs. conventional classroom instruction, for deaf children.

Sample: 99 deaf children, aged 7 to 10, drawn from 6 schools for the deaf.

Program: Linear program on language; 534 frames using pictorial as well as verbal cues.

Subjects lacked any appreciable written or vocal language, and the program was therefore designed to initiate symbolic communication. The subjects were divided into experimental and control groups, the former taught by program, the latter by conventional classroom methods, over a 20-day period. Criterion was ability to construct 50 sentences in response to 50 pictures. No significant difference was found between the groups on five of six language variables, but on the sixth the control group was superior. The experimental group required only 386.5 minutes to complete the program as compared to 900 minutes for the control group.

BLYTH, JOHN W.; GERE, BREWSTER H.; GODCHARLES, CHARLES A.; LIEDKE, OTTO K.; and MORAUD, MARCEL I. *The Hamilton College Experiment in Programed Learning*. Clinton, N.Y.: Hamilton College, 1962. 91 p.

Problem: Effectiveness of programed materials in college teaching.

Sample: 68 college students in French, 37 in German, approximately 100 in logic.

Programs: French, German, and logic. French program consisted of 35 visual units of 60 frames each, with correct answers provided after student response; 35 audio units integrated with visual ones; and 35 lessons in a workbook, corresponding to the visual and audio ones, and providing examples, drill, and questions. German programs provided both visual and audio lessons for three 20- or 30-minute laboratory periods. Logic program had over 8,000 frames, and was used as basic preparation for class meetings. (Programs in mathematics and psychology were made also, but no research results reported.)

Compared results of French course with results of former, traditional course given in 1953–55 (apparently without controlling IQ). Students taking programed course averaged about 20 percent higher on a test of written French, grammar, and translation. With the new course, it is now possible (and was not formerly possible) to teach second- and third-year courses entirely in French.

Students in programed German course averaged about 20 points higher on a test than did students in former, traditional course.

Students in programed logic course gained average of 10 points in final grade over previous year's class, which had had no programed materials. Some well-motivated students used *only* programs and made *A* scores on final exam. There were great individual time differences. Time per lesson correlated 0.641 with average number of errors, 0.604 with college entrance examination score, 0.728 with course grades—quicker students, better grades. This same program was also used successfully with gifted high school students.

BRIGGS, LESLIE J. "Prompting and Confirmation Conditions for Three Learning Tasks Employing the Subject-Matter Trainer." In LUMSDAINE, ed., 1961a, 375–87.

Problem: Prompting vs., and in combination with, confirmation modes.

Sample: 24 high school students.

Program: Learning map symbols from subject-matter trainer (a form of teaching machine).

Prompting, when employed as the sole practice condition, was found to be superior to various confirmation conditions. When prompting was followed by confirmation practices, there seemed little difference in effectiveness between single-try confirmation, multiple-try confirmation, or a combination of the two.

BRIGGS, LESLIE J. "Self-pacing versus Automatic Pacing in Paired-Associate Learning." In LUMSDAINE, ed., 1961b, 399–409.

Problem: Self-pacing vs. automatic pacing.

Sample: 77 airmen, in two groups.

Program: Subject-Matter Trainer program intended to teach association of names with line drawings of real and fictitious electron tubes.

Subjects were first given a prompting trial (in which they were shown the stimulus and response terms but not required to "respond"), followed by 13 minutes of practice with confirmation of correct answers. One group worked at its own pace through the confirmation part of the experiment; the other group was automatically paced at a rate of about 13 seconds per item. No significant differences were found in the post-test performance of the two groups. Author says, "This finding does not rule out the possibility that variable automatic pacing, slow at first, and faster as learning trials progress, could be superior to self-pacing."

BRIGGS, LESLIE J.; PLASHINSKI, D.; and JONES, D. L. *Self-Pacing versus Automatic Pacing of Practice on the Subject-Matter Trainer.* Lowry Air Force Base, Colo.: Armament Systems Personnel Research Laboratory, 1955, 6 p. Laboratory note.

See BRIGGS (1961b).

BROOKS, LLOYD O. *Response Latency in Programmed Learning: Latency Related to Error Rate.* Ph. D. dissertation, University of Houston, 1961.

Problem: Relation of latency of answers to errors in programed materials.

Sample: 16 university students.

Program: Remedial English, linear program, 13 one-hour lessons.

The data support the view that longer latencies tend to go with errors. Reading times, reading-plus-answer times, score and turn times, and total frame (but not overt answer) times, for frames on which a Subject made an error, all tended to be above Subject's own average. Median latency in all categories studied tended to be greater for error than for correct answer in the case of a given frame. There was also a strong association between error and latency in review tests at end of lesson. Latency is therefore suggested as more sensitive than error rate as a measure of difficulty.

BROWN, O. ROBERT, JR. *A Comparison of Test Scores of Students Using Programed Instructional Materials With Those of Students Not Using Programed Instructional Materials.* Urbana, Ill.: University of Illinois, 1962 (mimeo).

Problem: Programed instruction vs. conventional classroom teaching.

Sample: An experimental group of 147, and a control group of 183, eighth- and ninth-grade students from 7 different schools.

Program: Linear programed booklets prepared from the material in Unit 1 of the University of Illinois Committee on School Mathematics (UICSM) *High School Mathematics.*

Control group had conventional classroom teaching from the text mentioned. Experimental group had a combination of conventional teaching plus programed instructional materials prepared by the UICSM project members. The experimental group proved to be significantly superior to the control in a test of general ability, and about the same level of superiority was maintained in eight out of nine achievement tests given during the school term. The author concludes that "no student was penalized in his level of mathematics achievement because of having used programed materials."

BRYAN, G. L., and RIGNEY, J. W. *An Evaluation of a Method for Shipboard Training in Operations Knowledge*. Los Angeles: Department of Psychology, University of Southern California, 1956.

Problem: Effectiveness of different modes of reinforcement.
Sample: 48 college senior R.O.T.C. students.
Program: Multiple-choice items in shipboard operations.

Group A answered multiple-choice questions without knowledge of results. Group B was given immediate knowledge of results. Group C was given immediate knowledge plus an explanation of why the answer was right or wrong. On a multiple-choice examination 1 week later, group C did significantly better than the other groups.

BRYAN, GLENN L.; RIGNEY, JOSEPH W.; and VAN HORNE, CHARLES. *An Evaluation of Three Types of Information for Supplementing Knowledge of Results in a Training Technique*. Los Angeles: Electronics Personnel Research Group, Department of Psychology, University of Southern California, 1957. 23 p. (mimeo).

Problem: Relative effectiveness of three different types of explanations when employed in the multiple-choice trainer format.
Sample: 48 university students.
Program: Information not available.

The three types of explanation, given to rationalize the scoring key for each alternative, were as follows: (1) Explanations which gave the trainee the correct *definition* or *description* of the chosen alternative; (2) explanations which indicated the *principal reason* why the chosen alternative was keyed as correct or incorrect; (3) explanations which pointed out the probable *operational consequences* of the course of action represented by the alternative. Analysis of the data revealed that significant learning occurred as a result of the 1-hour training session regardless of which of the three types of explanation was employed. No one type of explanation was found to be superior.

BURTON, BENJAMIN B., and GOLDBECK, ROBERT A. *The Effect of Response Characteristics and Multiple-Choice Alternatives on Learning During Programed Instruction*. San Mateo, Calif.: American Institute for Research, 1962. 16 p.

Problem: Interaction of response mode with student aptitude.
Sample: 21 ninth-grade students.
Program: 35 frames dealing with such topics as characteristics of different animals.

783-239 O—64——3

Program was presented in three versions—constructed responses, multiple-choice questions with difficult responses (i.e., responses which could reasonably be made to the stimulus), and multiple-choice questions with easy responses (responses that could more easily be discarded than the "difficult" ones). Results did not support the idea that constructed response is superior to multiple choice (n.s.d.). Major finding was a significant interaction among learning methods, student aptitude, or presence or absence of the desired response in the (previously elicited) repertory of the student. Easy alternatives on the multiple-choice questions produced better learning of common responses for students with high verbal-reasoning aptitudes. Difficult alternatives produced relatively better learning of these responses for students with low verbal-reasoning aptitudes.

BUSHNELL, DAVID S. *Technological Change and the Journeyman Electrician: An Experimental Study in Continuing Education.* Menlo Park, Calif.: Stanford Research Institute, 1963.

Problem: Effect of teaching machines on attitudinal and behavioral change in an adult, voluntary education program.

Sample: 96 journeyman electricians.

Program: Electricity for journeymen, a program for a branching-type teaching machine.

Three instructional modes were employed: auto-instruction, auto-instruction plus live discussion and review, and conventional instruction. The above-average students on a pretest of knowledge performed significantly better in the auto-instructional mode. However, the combination of auto- plus live instruction produced higher student satisfaction and higher subsequent enrollment than did the other methods.

CAMPBELL, VINCENT N. *Adjusting Self-Instruction Programs to Individual Differences: Studies of Cueing, Responding, and Bypassing.* San Mateo, Calif.: American Institute for Research, 1961. 56 p.

Problem: Effect of different prompting and cueing, response mode, and "bypassing" (branching).

Sample: 3 experiments with 31, 57, and 31 junior high school students, respectively.

Programs: First experiment, linear program on static electricity (30 frames) and optics (33 frames); second, program on optics (33 frames); third, program on set theory (78 frames).

First experiment presented two versions of a program, one with responses rather fully prompted, the other with more indirect cueing.

Small, non-significant differences were found (of the order of 2 or 3 percent on a criterion test).

Second experiment provided four conditions: (1) Subjects wrote in the response only when sure of its correctness, (2) wrote answer in every blank, (3) did not write answers but mentally composed answers, (4) read the same frames with no words omitted—no blanks to fill. On an immediate post-test, there was no significant difference among the means of the four groups; 10 weeks later, condition (4) was better than others at 0.05 level. Expectation that option form (1) would make for more learning and save time was not supported. Condition (1) resulted in fewer errors, but these did not show up in learning effect.

A third experiment evaluated the usefulness of "bypassing," a process by which a subject is permitted to skip auxiliary loops of instruction as long as he is progressing successfully through the basic sequence of material. A group of 31 students were randomly assigned to three experimental conditions, using a program on set theory: (1) Short form of program with no bypassing, (2) long form of program with no bypassing, (3) long form of program with bypassing permitted. Students taking forms (2) and (3) learned significantly more than those taking (1). Students in (1) and (3) took significantly less time than students in (2).

This experiment was replicated at another school, with the addition of a group using a program of medium length. Differences in learning time were in the same direction as in previous experiment and were significant; there were no significant differences in the amount of learning. Difference in motivation at two schools, and resulting failure to learn much from latter parts of longer forms, are suggested as reasons.

CAMPBELL, VINCENT N. *Studies of Bypassing as a Way of Adapting Self-Instruction Programs to Individual Differences*. San Mateo, Calif.: American Institute for Research, 1962. 68 p.

Problem: Effect of bypassing under different conditions of program form and at different grade levels.

Sample: 8 experiments, using respectively fifty-four 9th-grade, twenty-seven 5th-grade, twenty-two 9th-grade, fifty-eight 8th-grade, forty-seven 4th-grade, fifty-four 12th-grade, forty-seven 8th-grade students, and 780 students in the 4th, 6th, 8th, 10th, and 12th grades.

Programs: Experiments 1, 4, 5, and 7 used a linear program on set theory; experiments 2 and 3 used a linear program on the United Nations; experiments 6 and 8 used both the U.N. and

the set theory programs (student works remedial loop if he gives wrong answer).

Experiment 1 tested (1) bypassing vs. (2) long linear form, (3) medium linear form, and (4) short linear form, on ninth-grade students. This was a replication of last experiment in CAMPBELL (1961) with addition of program's being divided between 2 class days to minimize fatigue. No significant difference in criterion test scores, but learning times differed significantly by length of program, with bypass program coming in between short and medium.

Experiment 2 tested fifth-grade class on three forms of program: (1) basic steps only—19 paragraph-size steps, (2) basic steps and one page of 2 to 4 multiple-choice questions following each basic step page, and (3) basic steps, evaluative questions, and remedial loops. Learning times differed significantly by length of program, but test scores were only marginally significant (0.10).

Experiment 3 tested bypass, long forms, and short forms. Learning times differed significantly, but overall scores on learning not significantly different. When only subjects scoring above 50th percentile on DAT (Differential Aptitude Test) numerical test were included, however, bypass treatment was significantly better than others.

Experiment 4 tested a short linear program against two bypass versions: One in which student was permitted to skip auxiliary loop when his basic step response was correct, and another in which student was instructed to skip auxiliary loop when he understood the page *and* when his answer was right. Subjects were three eighth-grade classes. No significant differences on amount of learning, even when the subjects were divided at the 50th percentile of DAT numerical score. Time differences significant as usual.

Experiment 5 tested the bypass version as compared to the short program, the long program, and a long reversed program in which the remedial loop appeared before the basic step. Analysis of covariance showed no significant difference. However, a t-test of bypass vs. short form showed bypass made for significantly more learning. Learning times significantly different: Short, bypass, and long groups in order.

Experiment 6 tested the bypass vs. short linear form. On sets program, resultant means nearly identical, although bypass program took nearly twice as long. On U.N. program, bypass group scored significantly higher than short form (0.05) but took three times as long.

Experiment 7 tested order of presentation: (1) Basic step preceding remedial loop vs. (2) remedial loop preceding basic step. The result: n.s.d.

Experiment 8 tested the bypass, long forms, and short forms of the sets program, and the bypass vs. short forms of the U.N. program.

Samples or sets program were classified in a 3-factor design: Learning method, grade, and ability group. Considering grades 4 through 12, grade, ability, and method effects were all significant, and grade x ability interaction was also significant. Considering grades 4 through 8, only the three main effects were all significant, and the grade x ability and grade x method interactions were also significant. In 12th grade, short and bypass means were almost identical; in 10th grade, bypass was higher, especially in high-ability group. In eighth grade, long form was highest, then bypass, then short, and so on for all ability groups. In sixth grade, same order obtained, except with low-ability group, where bypass method n.s.d. from short form. In fourth grade, long highest; short and bypass about same. With U.N. program, analysis of covariance (adjusted to DAT verbal) showed highly significant differences between bypass and short groups, bypass higher. Time as usual: three times as long for bypass.

Author concludes: "One should apparently be prepared to supply remedial loops of instruction which may take 20 or 30 times as long as the basic steps." Bypassing is most useful when a topic has a hierarchical structure, so that it is necessary or easier to learn one step before learning the next higher step. On that type of task, "the simplicity and economy of bypassing justify its use on even a small scale."

CANTOR, J. H., and BROWN, J. S. *An Evaluation of the Trainer-Tester and Punchboard Tutor as Electronics Troubleshooting Training.* Port Washington, N.Y.: Special Devices Center, Office of Naval Research, 1956.

Problem: Effectiveness of programed instruction.
Sample: 690 Navy men.
Program: Electronics.

Navy's traditional method of training in electronics using mockups of actual equipment was compared with same material conveyed either by a punchboard tutor, or by a trainer-tester, which was a large sheet with silver overlay printed above the answers. The latter two groups of naval trainees proved to be superior in certain intellectual aspects, while the traditionally taught group in some cases proved superior in laboratory work.

CARPENTER, C. R. and GREENHILL, L. P., and others. *Comparative Research on Methods and Media for Presenting Programed Courses in Mathematics and English.* University Park, Pa.: University Division of Instructional Services, The Pennsylvania State University, 1963.

Experiment I—*Comparison of External Pacing and Self-Pacing of*

Programed Instruction in Mathematics Using Different Methods of Instruction, by R. T. Heimer.

Problem: Effectiveness of programed course presented by (1) teaching machine (self-paced), (2) programed textbooks (self-paced), and (3) filmstrips (externally paced), compared with (4) conventional classroom teaching.

Sample: 113 college students enrolled in Math 2.

Program: Contemporary Algebra, 15 units, 2,055 frames.

Results: No significant differences were found among learning outcomes, except that on the unit tests (not the final tests) the programed treatments produced higher scores and lower variance. Variance (not means) was also significantly lower for the programed treatments on the final test.

Experiment II—*An Experimental Comparison of Differential Rates of Pacing Programed Math*, by J. J. Lottes, G. Palmer, and H. Oakes.

Problem: Effectiveness of different rates for externally paced programs.

Sample: 180 college Math students.

Results: No significant differences in student performance or attitudes toward programed learning resulted when pacing was varied as much as 20 percent below and 10 percent above the average time required by a group of self-pacing students.

Experiment III—*Adaptation and Evaluation of a Programed Mathematics Course for Televised Instruction*, by A. Pagano and J. J. Lottes.

Problem: To determine the effectiveness of televised programed instruction vs. individual teaching machine programed instruction of the same material.

Sample: 63 college students enrolled in Math 2.

Program: Contemporary Algebra, 28 units, 3,512 frames, linear type.

Results: No significant differences in learning were found.

Experiment IV—*Paired vs. Individual Study of Programed Instruction in Contemporary Algebra*, by W. Dick.

Problem: Effectiveness of paired vs. individual study of programed Algebra.

Sample: 70 college students enrolled in Math 2.

Program: Contemporary Algebra, 28 units, 3,512 frames.

Results: No significant differences were found between students who studied in pairs and those who studied individually, but verbal ability accounted for more of the variance than did quantitative ability for the paired groups, whereas the reverse

was true for the individuals. Paired group took significantly longer time to complete work.

Experiment V—*Comparisons of Televised with Teaching Machine and Televised with Instruction Presentations of English Grammar*, by R. E. Spencer.

Problem: Effectiveness of televised, externally paced programed instruction vs. teaching machine presentations, and television vs. lecture-discussion method in the classroom.

Sample:

TV vs. teaching machine study, 43 volunteer high school seniors and juniors.

TV vs. classroom presentation, 30 remedial English college freshmen.

Program: English Grammar, 14 units, linear program, 1,522 frames.

Results: No significant difference was found in students' performance, but attitudes were significantly in favor of instructor presentation.

Experiment VI—*Effects of Personality-Pairing on the Performance of Students in a Programed Course in English Grammar*, by W. Dick and E. L. Seguin.

Problem: Personality characteristics related to student performance in pairs on programed instruction.

Sample: 56 entering college freshmen enrolled in remedial English.

Program: English Grammar, 14 units, linear, 1,522 frames.

Results: One group of students were paired on basis of similarity in dominance-submissiveness score; other group on basis of dissimilarity. No significant differences were found in the programed learning results.

CARTWRIGHT, G. P. *Two Types of Programed Instruction for Mentally Retarded Adolescents*. Unpublished Master's Thesis, University of Illinois, 1962.

Problem: Effects of two sequences on learning, retention, and transfer.

Sample: 40 educable mentally handicapped children.

Program: Concept of a fraction; 612 frames.

In spite of equivalent learning scores, the groups under the two sequences earned significantly different mean scores on the measures of retention and transfer. The students taking Sequence 1 made significantly higher mean scores on retention tests than the students taking Sequence 2. The reverse relationship between the two groups was true for the regression scores on the transfer test; those taking

Sequence 2 obtained scores that exceeded their predicted scores, whereas those taking Sequence 1 obtained scores that were below expectation. Sequence 1 was designed so that a definite pattern of fractions and concepts would unfold from page to page and Sequence 2 was designed to reveal no such pattern. Sequence 1 can be thought of as having compensated more for general ability than did Sequence 2, but it did not compensate as well for specific abilities as did Sequence 2 as determined by the pattern of correlation between ability test scores and post-program learning-achievement test scores.*

CUMMINGS, ALLANA, and GOLDSTEIN, LEO S. *The Effect of Overt and Covert Responding on Two Kinds of Learning Tasks.* New York: The Center for Programed Instruction, 1962 (mimeo).

Problem: Effectiveness of overt vs. covert responding on learning of verbal and of pictorial material.

Sample: 68 female college students.

Program: 119 frames on diagnosis of myocardial infarction.

One group was told to write answers in the answer booklet; the other group, merely to "think" their answers. Significantly higher scores were made by the overt responders on both the verbal and pictorial materials, and on both the immediate and the delayed post-tests. The covert responders took an average of 50 minutes to complete the program; the overt responders, about 96 minutes.

COOK, JOHN OLIVER. "Supplementary Report: Processes Underlying Learning a Single Paired-Associate Item." *Journal of Experimental Psychology*, 56 (1958), 455.

Problem: Prompting vs. confirmation.

Sample: 160 airmen.

Program: Stimulus terms were first 10 letters of the alphabet. Response terms were nonrepresentational "code lines."

Subjects were tested on prompting (stimulus term on and off, response term on and off, then subject responds overtly) vs. confirmation (stimulus term presented, then subject responds overtly, then response term presented). A test was given after every third one of 36 training trials, so that the whole learning curve could be examined. Prompting proved to be superior throughout the learning curve.

COOK, JOHN OLIVER, and BROWN, JAMES EDWIN. "Familiarity and Novelty of Stimulus and Response Terms in Paired Associate

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

Learning." To appear in *Psychological Reports*, 12 (1963), 535–45.

Problem: Relative effectiveness of different combinations of novel and familiar terms as stimulus and response terms in learning of paired associates.

Sample: 45 college students.

Program: Familiar terms were letters of the alphabet; novel ones were dots randomly placed in an 8 x 8 matrix.

In order from highest to lowest, the four treatments were FF, NF, FN, and NN (F=familiar, N=novel). All differences were significant. Taking these and other dependent variables into account, author concludes that paired associate learning is a complex set of interdependent processes, and that altering one process by manipulating an independent variable appears to affect one or more other processes.

Cook, John Oliver, and Kendler, Tracy S. "A Theoretical Model to Explain Some Paired-Associate Learning Data." In *Symposium on Air Force Human Engineering, Personnel, and Training Research*, ed. by Glen Finch and F. Cameron, National Research Council Publications, 455. Washington, D.C.: NRC, 1958. p. 90–98.

Problem: Prompting vs. confirmation.

Sample: 90 male undergraduates.

Program: Stimulus terms were first 10 letters of the alphabet. Response terms were nonrepresentational "code lines."

Prompting proved more effective than confirmation in learning paired associates. (Prompting=presentation of stimulus and response terms together. Confirmation=presentation of stimulus, response, *then* correct response.) Theoretical model involves delay in stimulus-response process, and possible interference of overt responding.

Cook, John Oliver; Miller, Howard G.; Grier, James Brown; and Staman, James W. *A Generalized "Plan" for Serial Learning*. Raleigh, N.C.: North Carolina State College, 1962. 33 p. (mimeo).

Problem: Relative effectiveness of several guidance procedures in prompting serial learning, and of a confirmation (no guidance) procedure.

Sample: 60 college students.

Program: A 12 x 12 electronic punchboard maze, with one correct light in each row. Problem is to learn the correct sequence of lights.

Treatments were (1) self-paced prompting, in which the light over the correct button in the first row was turned on by the experimenter; when the subject had pressed the button under that light, the light turned off and the correct lights in the second row went on, etc.; (2) simultaneous-display prompting, in which all correct lights came on and stayed on for 36 seconds while subject observed but made no overt response; (3) sequential-display prompting, in which subject made no overt response, but simply observed as correct light in first row went on for 3 seconds, then correct light in second row for 3 seconds, etc.; (4) sequential-display prompting, which was similar to previous method except that, as each light came on, it stayed on until maze was completed, and subject responded in each case by pressing buttons; (5) confirmation, in which subject tried buttons in each row until he found the right one, which was signaled by the light going on; (6) automatically paced prompting, which was like (3) except that, during the 3 seconds when the light was on, the subject responded by pushing buttons. Groups (1) and (2) scored significantly better than the others; group (6) scored significantly lower than the others. A mediational model of serial learning is proposed.

COOK, JOHN OLIVER, and SPITZER, MORTON E. "Supplementary Report: Prompting versus Confirmation in Paired-Associate Learning." *Journal of Experimental Psychology*, 59. (1960), 275–7[illegible]

Problem: Prompting vs. confirmation.

Sample: 35 male college students for each of 4 groups.

Program: Stimulus terms were first 10 letters of the alphabet. Response terms were nonrepresentational "code lines."

Four treatments were used: (A) prompting, with no overt practice (e.g., the stimulus term was presented and removed, then the response term was presented and removed; (B) confirmation, with no overt practice (e.g., the subject was shown the stimulus term, allowed to make a covert response, then shown the response term); (C) prompting with overt practice (e.g., written practice after seeing both terms); and (D) confirmation with overt practice (e.g., the ordinary Skinner program, stimulus term with blank, filling in the blank, then knowledge of results).

Condition A proved to make for fastest learning, D for slowest. Condition A was assumed to be best because of the short stimulus-response delay and because no intervening practice interfered. Condition D was judged to be worst because of the long stimulus-response delay and the interfering effect of overt practice.

Conclusions: Under some conditions overt practice interferes with learning a response term (in paired associates) and connecting it with

proper stimulus; stimulus-response delay interferes with process of connecting a response to its proper stimulus, but has no very reliable effect upon learning the response term as such.

COULSON, JOHN, E.; ESTAVAN, DONALD P.; MELARAGNO, RALPH J.; and SILBERMAN, HARRY F. "Effects of Branching in a Computer Controlled Autoinstructional Device." *Journal of Applied Psychology*, 46, 6 (1962), 389–92.

Problem: Branching vs. fixed sequence in a computer-controlled program.

Sample: 2 groups of 15 high school students.

Program: 233 items on logic for the fixed sequence program; more or less than this number, as determined by the computer, for the branching sequence.

In the branching sequence, the computer was instructed to select sequences of items on the basis of (a) the number of errors accumulated by the student on a given topic, (b) answers indicating a particular kind of misunderstanding, (c) answers given by the student on whether he feels ready to advance to a new topic or needs further review. Post-test scores were significantly higher for the branching group than for the fixed-sequence group. Training times for the two groups were not significantly different.

COULSON, JOHN E., and SILBERMAN, HARRY F. "Effects of Three Variables in a Teaching Machine." *Journal of Educational Psychology*, 51 (1960), 135–43.

Problem: Response mode and size of step, as effecting learning.

Sample: 80 experimental, 104 control, college students in psychology course.

Program: Psychology items (104 frames in small-step form; 56 in large-step; number varied with individual in branching form).

No significant difference in performance on a post-test (requiring constructed responses) between two groups which had studied, respectively, multiple-choice and constructed-response programs. Multiple-choice program required significantly less time. Small-step forms were significantly superior to large steps in amount learned, but required significantly more time. Branching subjects learned as much as fixed-sequence subjects, but branching required significantly less time.

DELLA PIANA, GABRIEL. "An Experimental Evaluation of Programmed Learning." *Journal of Educational Research*, 55 (1962), 495–501.

Problem: Effect of different degrees of prompting and motivation.

Sample: 175 juniors in education.

Program: Introduction to counseling, 280 frames.

Different groups were given different program treatments, including (a) the usual constructed response, (b) initial letters of required response inserted in program in red, (c) entire response, in red, inserted into program, so that the student merely copied it, (d) use of the common linear program, as in (a), with an instructor saying "right" or "wrong" after each response. No significant differences were found in learning among the several treatments, but treatments (b) and (c) proved somewhat more efficient than the others in terms of time and errors. Degree of motivation was found to have an effect on within-group variance in constructed response programs.

DOWELL, E. C. *An Evaluation of Trainer-Testers.* Keesler Air Force Base, Miss.: Headquarters, Technical Air Force, 1955 (mimeo).

Problem: Relation of practice on verbal programed material and practice on actual equipment to learning of skills with equipment.

Sample: Large number of airmen (26,000 sets of programed materials used); exact number not available.

Program: Troubleshooting with the superheterodyne circuit.

Group (1) studied the programed material and practiced on actual equipment only. Group (2) studied and practiced on both programed material and equipment. Group (3) studied and practiced on the programed material only. Group (4) received general instructions on troubleshooting, and did not practice. Group (2) learned most. Group (1) did better than group (4), but (3) did no better than (4).

EIGEN, LEWIS D. "High School Student Reaction to Programed Instruction." *Phi Delta Kappan*, 44, 6 (1963), 282–85.

Problem: Student reaction to programed instruction.

Sample: 72 high school students (mean IQ, 126), 39 of whom used a Skinner-style write-in machine, the other 33 a programed textbook.

Program: Linear program on sets, relations, and functions.

Subjects who performed well on achievement test were significantly more likely than others to say that programed instruction is "best method of learning" for good students because they are not held back by the rest of the class, and also that students learn more from automated teaching because they never get left behind the class. In general, students using the programed text had a more favorable

attitude toward programed instruction than did those using the teaching machine. Students' total attitude toward automated teaching, however, appeared to have no relationship to how much they had learned by the method. Author concludes that "it is difficult if not impossible" to conceive of a typical reaction to controversial statements about programed instruction after a student's first exposure. Attitudes are vastly different from student to student.

EIGEN, LEWIS D.; FILEP, ROBERT T.; GOLDSTEIN, LEO S.; and ANGALET, BRUCE W. *A Comparison of Three Modes of Presenting a Programed Instruction Sequence.* New York: The Center for Programed Instruction, 1962. 40 p. (mimeo).

Problem: Teaching machine vs. horizontally programed text vs. vertically programed text.

Sample: 77 eighth-grade students, in 3 groups.

Program: Numbers and numerals, 65 frames, linear.

One group used a teaching machine, a second used a textbook program with horizontal progression, a third a textbook program with vertical progression. No significant difference was found in amount of learning either on immediate or delayed post-test. Time for textbooks was less than for machine, but not significantly so. There appeared to be an interaction between IQ and method, but not quite significant. ($p<0.10$)

EIGEN, LEWIS D., and KOMOSKI, P. KENNETH. *Automated Teaching Project* (Research Summary No. 1). New York: Collegiate School, 1960. 7 p. (mimeo).

Problem: Teaching machine vs. programmed text; learning from same program at different grade levels.

Sample: Twenty-five 9-grade, twenty-three 10-grade, and twenty-six 11-grade students.

Program: Sets, relations and functions; 707 frames; linear.

No significant differences were found between learning resulting when the program was presented on a machine, and when it was presented as a mimeographed programed book. No significant effects of grade level on learning were found, suggesting, say the authors, that the belief that certain materials are best taught at a certain grade level needs serious reexamination.

EIGEN, LEWIS D., and MARGULIES, STUART. "Response Characteristics as a Function of Information Level." *Journal of Programed Instruction* (Spring 1963) Vol. II, No. 1.

Problem: Effectiveness of relevant vs. incidental responses, and covert vs. overt responses, in the case of answers requiring different amounts of information.

Sample: 362 high school students, 179 of whom were told to write their responses, 183 to "think" them.

Program: 12 items designed to teach groups of three-letter words.

Each program was designed to teach a string of three-letter words, some of which were more difficult than others to remember, or, as the authors preferred to say, required more information in order to be reproduced. For example, reproducing the nonsense word CVX (high information level) requires more "information" than the nonsense word BOF (intermediate information level), which in turn requires more than the word FAT (low information level). The authors found that overt persons resulted in better post-test performances at both intermediate and high information level, but not for low information level. Relevant responses resulted in better post-test responses than incidental ones.

EVANS, JAMES L. *An Investigation of "Teaching Machine" Variables Using Learning Programs in Symbolic Logic.* Doctor's dissertation, University of Pittsburgh, 1960. (Also submitted as a report to U.S. Office of Education, under authorship of Evans, Robert Glaser, and Lloyd E. Homme.)

Problem: Effects of different response modes, reinforcement, and stimulus presentation.

Sample: 60 college students.

Program: Linear programs designed to teach how to construct short deductive proofs employing 15 rules similar to those used in symbolic logic. "Ruleg" program of 125 frames; "less systematic" program of 72 frames.

Employing different versions and treatments of program, found that (a) students not required to make overt responses to items completed program in about 65 percent of time required by overt responding students; (b) neither overt vs. covert responding nor multiple-choice vs. constructed responses nor immediate vs. slightly delayed feedback made any significant difference in criterion performance; (c) program constructed by "ruleg" method resulted in the same performance in less time than less systematic program; (d) overt responders took less time than covert responders on an immediate post-test, but showed no difference after a week. Performances on retest after 1 week differed as a function of the measure: True-false scores better, recall less, actual construction of proofs unchanged.

EVANS, JAMES L., GLASER, ROBERT, and HOMME, LLOYD E. "An Investigation of 'Teaching Machine' Variables Using Learning Programs in Symbolic Logic." *Journal of Educational Research*, 55 (1962), 433–52.

See EVANS, GLASER, and HOMME (1960) and GLASER and TABER, 1961.

EVANS, JAMES L.; GLASER, ROBERT; and HOMME, LLOYD E. "An Investigation of Variation in the Properties of Verbal Learning Sequences of the 'Teaching Machine' Type." In *Teaching Machines and Programmed Learning: A Source Book*, ed. by Lumsdaine and Glaser. Washington, D.C.: Department of Audio-Visual Instruction, National Education Association, 1960. 446–51.

Problem: 3 experiments, respectively testing size of step, overt vs. covert-response, and programed instruction vs. conventional text.

Sample: first experiment, 4 groups of 5 graduate education students; second, 2 groups of 5 undergraduate psychology students; third, 33, 13, 10 college undergraduates, each group divided into 2 for experimental purposes.

Programs: conversion to number bases other than 10 (4 forms), measures of central tendency (approximately 50 frames), fundamentals of music (approximately 60 frames): linear programs.

First experiment compared performance on retention tests of groups studying 4 versions of a program with, respectively, 30, 40, 51, and 68 steps. Groups who had small-step programs (51 and 68) made significantly fewer errors on program and performed significantly better on immediate and delayed retention tests than did other groups. Group with 51 steps did slightly better than group with 68.

Second experiment compared two groups instructed respectively (1) to write answers to each step (2) to respond covertly to each step. Found that group (2) gained more in performance, but nonsignificantly; also took less time to complete sequence (also nonsignificantly); and displayed less variability in performance scores (nonsignificantly).

Third experiment compared learning sequence (1) with conventional text materials (2) on three different subjects. Found that groups with learning sequence had higher performance scores on post-test, but significant only with program on music. On two statistics programs, however, found that text-reading groups displayed more variability in performance scores (differences significant).

FALCONER, G. W. "A Mechanical Device for Teaching Sight Vocabulary to Young Deaf Children." *American Annals of the Deaf*, 106 (1961), 251–57. (Doctoral dissertation, University of Illinois, 1959.)

Problem: Effectiveness of a stimulus-response, multiple-choice, immediate reinforcement, electromechanical teaching device.

Sample: 8 profoundly deaf children.

Program: 15 nouns taken from a standard primary wordlist, 90 frames.

Teaching material for use in the machine was developed. The machine was a drum housed in a box which held 12 cards, only 1 of which showed through a window. The machine was pretested on 43 children with normal hearing. The subjects in the experiment were 8 deaf children, who were taught 15 words through working at the machine 5 minutes a day for 10 consecutive school days. Immediate comprehension was high and after 2 weeks, testing revealed nearly perfect retention.*

FELDHUSEN, JOHN F., and BIRT, ANDREW. "A Study of Nine Methods of Presentation of Programed Learning Material." *Journal of Educational Research*, 55 (1962), 461–66.

Problem: Amount of learning from different types of programs.

Sample: 270 college students.

Program: Linear program on teaching machines and programed texts; 55 blanks compressed to 37 frames, in linear format.

Comparison groups were given the same linear program (a) in a teaching machine, (b) in a simple cardboard folder, (c) set up for class rather than individual administration, (d) with knowledge of results withheld altogether, (e) with knowledge of results withheld until completion of the entire program, (f) with the answer blanks filled in, (g) with pace predetermined rather than individually set, and (h) with narrative material derived from the program. No significant differences were found in learning among the different groups.

FELDHUSEN, JOHN F., and EIGEN, LEWIS D. *Interrelationships Among Attitude, Achievement, Reading, Intelligence, and Transfer Variables in Programed Instruction*. Paper presented to the Midwestern Psychological Association, May 1963. 23 p. (mimeo).

---

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

Problem: Relationship of attitudes toward the method, IQ, general achievement level, reading ability, attitude toward school, characteristic error patterns, and study methods, to learning from programed instruction; and interaction of mode of presentation (teaching machine or programed textbook) with other variables.

Sample: 2 studies using same program; Study A, twenty-four 9th-grade students; Study B, twenty-four 9th-grade, twenty-three 10th-grade, and twenty-five 11th-grade students.

Program: Linear program on sets, relations, and functions.

This is a correlation study, and the correlations and partial correlations are too many to be reported here. The authors conclude that "attitudes of students toward programed instruction are not consistently related to the youngsters' levels or amounts of learning . . . However, a general pattern of increasing correlations for acquisition and transfer with attitude is noted in the progression from ninth to eleventh grade groups . . . The variance in learning which may be attributed to IQ seems less essential than that which may be attributed to general achievement level . . . in neither study is IQ *per se* found to be the fundamental learner variable in programed instruction: . . . The variance attributable to reading ability [was] not essential in accounting for learning in programed instruction . . . The teaching machine presentation seems to magnify the relationships of time with all other variables except attitude."

FELDHUSEN, JOHN F.; RAMHARTER, HAZEL; and BIRT, ANDREW T. "The Teacher vs. Programed Learning." *Wisconsin Journal of Education*, 95 (1962), 8–10.

Problem: Programed vs. conventional classroom instruction.

Sample: 2 matched groups, each one composed of thirteen 7-grade students.

Program: 14 weeks of linear programed instruction on fundamental arithmetical skills.

One group taught by program, other by teacher, for first 7 weeks; procedures reversed for the second 7 weeks. No significant differences were found between the groups either on tests of achievement given after the 7th week or on others given after the 14th week. Authors say, "Under the carefully planned regimen provided by a live teacher and under the programed instruction, each group gained approximately 1 full year in arithmetic skills during only 14 weeks of instruction." An attitudinal scale was given. Responses were highly varied, although there was a greater concentration on responses favorable to programed learning.

782-229 O—64——4

FERSTER, CHARLES B. *The Role of Review Material in Continuous Programing With Teaching Machines.* Indianapolis: Indiana University Medical School, 1960. 4 pp.

Problem: Effect of repeating whole lesson in programed instruction.

Sample: 49 medical school students.

Program: First 61 lessons of Holland-Skinner program; linear.

Different versions of the program provided for repetition or non-repetition of certain entire lessons. Repetition brought about no significant difference in amount of learning. Authors concluded that the important repetition was that which was already built into the lessons.

FERSTER, CHARLES B., and SAPON, STANLEY M. "An Application of Recent Developments in Psychology to the Teaching of German." *Harvard Educational Review*, 28 (1958), 58–69.

Problem: Effectiveness of a programed course.

Sample: 6 college students.

Program: German language, approximately 800 frames.

Six students studied a programed course in German carrying "an amount of German comparable to that presented in a first-semester course." Learning was considerable. Authors found no correlation between aptitude and achievement in learning from program, but say that in a mean time of 47.5 hours, the six students learned an amount of German comparable to that presented in a first-semester course.

FOLLETTIE, JOSEPH F. *Effect of Training Response Mode, Test Form, and Measure on Acquisition of Semi-ordered Factual Materials.* Fort Benning, Ga.: U.S. Army Infantry Human Research Unit, Research Memorandum No. 24, 1961. 62 pp. (mimeo).

Problem: Effectiveness of learning from (1) live vs. taped lecture, (2) taped lecture vs. force-paced reading of a book, (3) force-paced book vs. self-paced book, (4) self-paced book vs. scrambled book program.

Sample: About 375 Army basic combat training troops, in 5 groups.

Program: Various treatments of two topics: first aid and rules of land warfare.

Author found no significant difference between learning from live and taped lecture, a significant advantage of read material over heard material, a significant advantage of self-paced over class-paced reading (at not more than seven-tenths of a standard deviation from

a student-set pace), and a significant advantage of plain book over scrambled book format. Using two forms of test—recognition (three choices) and recall (sentence completion)—found generally stable relationship between them. Also used two measures of acquisition: accuracy per unit training time and accuracy modified by test responding per unit training time.

ERY, EDWARD B. "A Study of Teaching Machine Response Modes." In *Teaching Machines and Programed Learning: A Source Book*, ed. by Lumsdaine and Glaser. Washington, D.C.: Department of Audio-Visual Instruction, National Education Association, 1960, p. 469–74.

Problem: Constructed vs. selected responses.
Sample: 153 ninth-grade students.
Program: Spanish vocabulary items.

Multiple-choice and constructed modes of responding were compared under three conditions: (1) Students worked with a masked program until achieving a fixed criterion of two correct responses—time of achievement was recorded; (2) students were stopped before mastery to insure equal working time; (3) items were presented on group flashcards to insure equal presentation-time and equal responding-time. On immediate and delayed post-tests, multiple-choice test items approached maximum possible score, hence no significance in training was noted. However, on constructed response post-test items, the constructed response training group did significantly better for all conditions. Multiple-choice training was faster; hence if multiple-choice test items are the only criterion, they are more efficient.

FRYE, CHARLES H. *Group vs. Individual Pacing in Programed Instruction*. Oregon State System of Higher Education (undated—probably written in 1963)*.

Problem: Effectiveness of individual vs. group pacing, as related to homogeneity or heterogeneity of group.
Sample: 44 high school freshmen, divided into 4 groups.
Program: Linear program; 47 frames on "Completing the Square" and 88 on "Quadratic Formula."

The four groups were (1) homogeneous—determined on the basis of IQ and predicted algebra ability—and group-paced, (2) homogeneous and individually paced, (3) heterogeneous and group-paced, (4) heterogeneous and individually paced. It was found that the time required by the heterogeneous and group-paced subjects (group

*As noted previously, some bibliography entries are unpublished. Among the unpublished entries, some are also undated, and the dates given here for such entries are based on the author's best recollection.

3) was significantly greater than that required by the heterogeneous and individually paced (group 4). In the homogeneous groups, however, the time required by the group pacing (1) was not significantly more than that needed by the individual pacing (2). The time required by the heterogeneous and group-paced subjects (3) was significantly greater than that required by the homogeneous and group-paced subjects (1).

GAGNÉ, ROBERT M. "The Acquisition of Knowledge." *Psychological Review*, 69, 4 (1962), p. 355–65.

Problem: Effectiveness of learning program designed to transfer training from component learning sets to a new activity which incorporates these previously acquired capabilities.

Sample: 7 ninth-grade boys.

Program: Program on finding formulas for sum of n terms in a number series.

Theory of task analysis into hierarchical learning sets is discussed, and is illustrated by experiment on seven boys. Each began at the point of his lowest successful learning set achievement; six of the seven were brought to successful achievement through ascending hierarchies to the final task goal.

See GAGNÉ and PARADISE, 1961; and GAGNÉ, MAYOR, GARSTENS, AND PARADISE, 1962.

GAGNÉ, ROBERT M., and BROWN, LARRY T. "Some Factors in the Programing of Conceptual Learning." *Journal of Experimental Psychology*, 62 (1961), pp. 313–21.

Problem: Effectiveness of different program strategies in teaching concepts.

Sample: 33 boys in ninth and tenth grades, divided into 3 groups.

Programs: Linear programs designed to teach users to devise formulas for the sum of terms in different kinds of series. Introductory: 89 frames; R + E: 36; D: 7, plus 0–14 "hints"; GD: 40 (abbreviations explained below).

Each group learned the basic concepts from an initial small-step program. Then each group was given a different program designed to teach them certain concepts. The three programs were, respectively, a "ruleg" program (R + E), in which each of the concepts was stated and then illustrated with examples; a "discovery" program (D), in which the student was given material from which he could actively produce the desired concepts; and a "guided discovery" program (GD), in which the student was guided down the path of discovery and was given progressively more and more information until he could actively produce or "discover" the concept.

All three programs produced significant gain (0.01). The GD program was best, D next, R+E worst, with differences of individual means significant at 0.02 level or better.

GAGNÉ, R. M., and DICK, W. "Measures in a Self-Instructional Program." *Psychological Reports*, 10 (1962), p. 131–46.

Problem: Different measures of learning from programed instruction.

Sample: 52 seventh-graders.

Program: 8 programed booklets on solving equations, 245 frames.

The experimenters used a test of ability to reproduce verbally the concepts learned, immediately and again 6 weeks after working through the program. They also applied a test of performance in solving the sort of equations with which the program dealt; again, an immediate post-test and another 6 weeks later. Finally, the experimenters gave a test of transferability of the knowledge learned, i.e., the ability to solve equations with new and unfamiliar symbols and structures.

The immediate post-tests, both verbal and performance, revealed a considerable amount of learning, a little over 90 percent of which was retained after 6 weeks. The correlation between verbal and performance tests was 0.69, and between the verbal and performance retests, 0.73. The correlations between these test grades and the students' previous mathematics grades ranged from 0.50 to 0.53. The correlations between performance measures and *internal* program measures (errors and time required) were low and insignificant. Mean score in the transfer test was 2.07, out of a possible 50. Transfer scores were correlated significantly with both verbal and performance scores, but not with previous mathematics grades. The authors discuss implications of these results for measuring "what is learned."

GAGNÉ, ROBERT M.; MAYOR, JOHN R.; GARSTENS, HELEN L.; and PARADISE, NOEL E. "Factors in Acquiring Knowledge of a Mathematical Task." *Psychological Monographs*, 76, 7 (1962), No. 526.

Problem: Effectiveness of learning program requiring learner to proceed through a hierarchy of learning sets "which mediate positive transfer of learning in a unidirectional fashion from one to another, and ultimately to the performance" of the task-goal.

Sample: 136 seventh-grade students.

Program: Four versions of a program aimed at teaching two tasks: (1) Stating the series of steps necessary to formulate a definition of addition of integers, (2) adding integers. Ver-

sions contained, respectively, low repetition, low guidance (see below), 99 frames; low repetition, high guidance, 125 frames; high repetition, low guidance, 105 frames; high repetition, high guidance, 134 frames.

"Repetition" referred to the insertion into the program of many or of few problems for each newly attained learning-set task. "Guidance" referred to the use of a few or of many frames to state a logical principle, in other words, to how completely the steps on thinking that were necessary to attain each new learning set were stated.

Relatively small effect was found for the experimental variables, i.e., repetition and guidance. For task 2 and for a test of transfer, no significant effects of the experimental variables, of ability, or of their interactions were found. For task I, there was a significant effect for the combination of the experimental variables: High repetition and high guidance produced performance superior to that produced by low repetition and low guidance. Tested singly, these variables produced effects that were not significant.

Addition or omission of learning sets in the hierarchy, however, was found to have powerful effects. Acquisition of new knowledge at successively higher stages was found to be dependent on prior mastery of subordinate learning sets. When mediating effects were examined for learning sets intervening between lower and higher ones, proportions of achievement on the higher learning sets indicated substantial amounts of positive transfer.

See GAGNÉ and PARADISE, 1961.

GAGNÉ, ROBERT M., and PARADISE, NOEL E. "Abilities and Learning Sets in Knowledge Acquisition." *Psychological Monographs*, 75, 15 (1961), No. 518.

Problem: Relation of certain abilities to learning from a program which leads a learner through a hierarchy of learning sets "which mediate positive transfer of learning in a unidirectional fashion from one to another, and ultimately to the performance" of the task-goal.

Sample: 118 seventh-grade students.

Program: 247 frames on learning to solve linear equations.

The task was analyzed into a hierarchy of 22 learning sets, plus 3 additional ones at the low end of the hierarchy, corresponding to common ability factors sometimes called number ability, symbol recognition, and "Integration I" (of which an example would be "following directions"). Two factors thought to be irrelevant were also introduced into the program: Speed of symbol discrimination, and vocabulary. It was found that the correlations of the theoretically relevant

basic abilities were higher than those of the irrelevant basic abilities with measures of final performance, transfer of training scores, number of learning sets achieved, and rate of learning of the total program. Correlations of relevant basic ability with *rates of attainment* of learning sets at progressively higher levels of the hierarchy decreased steeply, whereas their correlations with *achievement* increased at higher levels. On the other hand, the correlations of the irrelevant abilities with these measures remained nearly constant. Evidence of several kinds (for example, positive transfer to each learning set from subordinate relevant learning sets) was found to confirm the learning set analysis.

See GAGNÉ, MAYOR, GARSTENS, and PARADISE, 1962.

GAVURIN, EDWARD L., and DONAHUE, VIRGINIA M. *Logical Sequence and Random Sequence Teaching Machine Programs.* Burlington, Mass.: Radio Corporation of America, 1960. 16 pp. Also in *Automated Teaching Bulletin*, 1 (1961), 4 p. 3–9.

Problem: Logical vs. random sequence of items.

Sample: 40 RCA employees, in 2 experimental groups.

Program: 29 items on psychology.

Found significantly fewer errors during trial and significantly fewer trials to criterion, when items were logically sequenced than when randomly sequenced. Test of retention of information 1 month later showed no significant difference between groups.

GLASER, ROBERT, and REYNOLDS, JAMES H. *Investigation of Learning Variables in Programed Instruction.* Pittsburgh: Programed Learning Laboratory, University of Pittsburgh, 1962. 182 p.

Part I. Repetition and Spaced Review Variables.

Experiment I:

Problem: Effects of combined repetition and spaced review upon learning and retention (linear vs. spiral programing) under fixed-time conditions (one semester allowed, which was not enough to finish program).

Sample: Junior high school students, 47 using linear program, 50 using spiral program, and 41 traditional teaching.

Program: General science program; linear version about 7,000 frames, spiral version about 10,000

Experiment II:

Problem: Same as in experiment I, but under fixed-material conditions (each group allowed to finish the program, regardless of time required).

Sample: Junior high school students, 31 in a nonreview treatment, 24 in a review treatment.

Program: Same as in experiment I.

Experiment III:

Problem: Effects of repetition and review upon retention in a linear program.

Sample: 75 junior high school students.

Program: 1,280-frame biology chapter from linear general science program.

The authors conclude, "These experiments showed (1) that variations in amount of repetition required of the subject during learning had just transitory effects upon the amount learned and no significant effects upon retention, (2) that spaced review had significant facilitating effects upon learning and retention when used in a linear programing format, (3) that retention of materials learned by programed instruction is equivalent to retention from traditional instructional methods over periods at least as long as 15 weeks, and (4) that reminiscence, under certain testing conditions at least, is a significant factor contributing to long-term retention of sequential learning materials. Other results . . . implied that intelligence and achievement measures may not be as predictive of amount of learning resulting from a linear programed sequence as they are of learning acquired by other instructional methods."

Part II. Learning Set Formation.

Experiment IV:

Problem: Formation of learning sets making for efficient use of programed instruction.

Sample: Matched groups from among 120 junior high school students.

Program: 3 chapters from general science program.

Concerning this experiment, the authors conclude, ". . . a learning set may show up as a change in error response rate but is not reflected in final achievement. Such learning sets can occur only in programs having a relatively high error rate. A conclusion of this study is that the difference between low and high error rate programs may not be as potent a factor in programed instruction as generally believed."

Part III. Programing Method and Response Mode in a Visual-Oral Response Task (Csanyi, Attila P., and Glaser and Reynolds).

Problem: Prompting vs. confirmation; oral vs. nonoral response.

Sample: 40 junior high school students.

Program: 4 versions, each having 480 frames, of a program designed to teach pronunciation of 12 phonetic symbols.

The authors conclude, "No differences in learning or retention were found between programing methods [prompting vs. confirmation], but oral responding during learning resulted in higher retention performance on oral tests given after the learning trials. Further evidence that reminiscence is a factor contributing to retention performance was obtained."

GLASER, ROBERT, and REYNOLDS, JAMES H. *Multi-Tracking in Programed Instruction: Studies with a Program on Mathematical Bases for Management Decision Making.* Pittsburgh: Programed Learning Laboratory, University of Pittsburgh, 1962.

Part I. (tests in connection with development of a program):

Problem: Linear vs. multitrack programing (multitrack program permitted student to call for additional prompts if needed to respond correctly to a frame).

Subjects: 14 college students, 12 Air Force personnel, and 20 industrial personnel.

Program: Mathematical decision making.

Results of testing indicated little difference in efficiency between the linear and the multitrack program. College-educated subjects did considerably better than noncollege. In the course of developing the program, small sample tests indicated no significant difference between effect of home study and of classroom study of the program, a significant difference in favor of brief review of program before criterion test, no significant relation of error rate to criterion performance, and a higher score for students using the program than for students taught by conventional classroom methods.

Part II.

Problem: Formal test of linear small-step program vs. a multitrack version in which large steps were employed; when a subject made an incorrect response to one of these steps, he branched to a sequence of small steps covering same material.

Subjects: 34 mathematics teachers.

Program: First 9 sections, 1,060 frames, of mathematical decision-making program (teaching matrix theory summation symbolism, and industrial applications).

Results indicated "no difference in instructional efficiency of the two programming procedures even though error rate was greater for the large-step branching program. Both groups obtained the same level of performance on criterion tests and both worked at an average rate of about 100 frames an hour."

GLASER, ROBERT, and REYNOLDS, JAMES H. *Some Insights about Instructional Objectives Arising from Work on Programed Instruction.* Pittsburgh: Programed Learning Laboratory, University of Pittsburgh (undated—probably written in 1963).

Problem: Implications for instructional objectives derived from process of constructing and testing a program.

Sample: First-grade students (n=n.a.)

Program: How to tell time.

Authors conclude, "1. In order for an instructional sequence to be adequately prepared, instructional objectives need to be stated in terms of observable events . . . 2. All lessons cannot teach all conceivable learning outcomes. The existence of primary . . . and secondary instructional objectives must be recognized . . . 3. The teacher works directly with terminal instructional behavior, and incidentally with long-term educational goals. The teacher must, for the most part, obtain information about his success in terms of the behavior the student displays at the end of a course of instruction. 4. Attainment of terminal behavior is achieved by teaching sub-objectives which comprise successively finer approximations to mastery. The appropriate teaching sequences and strategy for each achievement level should be determined and evaluated empirically. 5. The diagnosis of the entering behavior necessary to begin the instructional sequence is as important to specify as the terminal objectives . . . 6. Transfer and generalization cannot be assumed. They must be made explicit instructional objectives and must be taught and . . . evaluated. 7. Objectives should be measured in terms of the content of the achievement obtained and not in terms of the relative achievement of one student with another. 8. Motivation needs to be considered as a specific instructional objective and a specific instructional problem."

GLASER, ROBERT, and TABER, JULIAN I. *Investigations of the Characteristics of Programed Learning Sequences.* Pittsburgh: Programed Learning Laboratory, University of Pittsburgh, 1961.

Problem: Effectiveness of different strategies in programing. Chapter 4 deals with schedules of reinforcement, chapter 5 with a Skinner-Cook-Pressey comparison, chapter 6 with a comparison of programed text vs. lecture, chapter 7 with printed vs. auditory cues in learning prounciation of a foreign language, chapter 8 with an ingenious technique of fading out cues, chapter 9 with discriminative transfer.

Sample: Chapter 4 uses 84 high school juniors; chapter 5, 96

high school juniors; 6, 49 pharmacy students; 7, 30 junior high students; 8, 3 students; 9, 15 five- and six-year olds.

Programs: Mentioned below in connection with each chapter.

Chapter 2 consists of the specifications for the Ruleg method of program construction. Chapter 3 is a summary of the Evans dissertation (See Evans, 1960). Chapter 4 describes an M.A. thesis by Eugenia S. Scharf on the effects of partial reinforcement in a learning program. The symbolic logic program was rewritten for high school students and presented in four alternate treatments with, respectively, 100 percent continuous reinforcement, 50 percent fixed ratio reinforcement, 50 percent variable ratio reinforcement, and 25 percent variable ratio reinforcement (reinforcement was the correct answer on the back of the card on which the question was written). On a multiple-choice criterion test, no effect could be found which was attributable to any schedule of reinforcement. It is suggested that reinforcement by confirmation is a crucial event only when probability of correct response is low; but in Skinner programs the probability is deliberately kept high.

Chapter 5 is a summary of a Ph. D. dissertation by Robert B. Hessert. Students were assigned randomly to three versions of the logic program, one version constructed by the "Skinner" method of short steps and continuous confirmation; one by the "Cook" method of prompted paired associates (stimulus and response presented together); and a third by a "Pressey" method of no overt response and frequent multiple-choice questions to test response. No difference was found on learning criterion test, but the Skinner type, or prompted and reinforced version, took significantly less learning time.

Chapter 6 reports a master's thesis by Carole Finelli. Pharmacy students were taught the desired subject matter about two vitamins by traditional lecture methods, and the desired subject matter about two other vitamins of comparable complexity by programed textbook. On an immediate post-test and again 6 weeks later, learning results were significantly in favor of the programed presentation.

Chapter 7 is an experiment by Attila P. Csanyi, in which students in three experimental groups were taught the pronunciation of a list of Spanish words in three ways: by being allowed to practice the pronunciation (1) as suggested by English cues of approximately the same sound, (2) as suggested by the usual tape recording of native speakers, pronouncing the words while students read the words on a card, and (3) as suggested by a combination of hearing the tape recorder and reading the English cues. At the end of this procedure, the students read the words into a tape recorder, and four Spanish teachers and native speakers graded the performance.

No significant difference was found between the performance of students trained in the different ways.

Chapter 8 is a paper by Halmuth H. Schaefer describing an attempt to teach German vocabulary by progressively introducing more German equivalents of Common English words into a passage of English fiction where the redundancy of the English permitted the student easily to assign meaning to the foreign words. Three students using this kind of program learned 66 percent of the words, as shown by a meaning recognition test.

Chapter 9 describes the making and trial of a program which attempts to teach color names by discriminative transfer by gradually withdrawing the color prompts (actual use of the color in lines or in printed letters). At the end the student is left with only the written word, to which, it is hoped, the verbal discriminative response previously associated only with the color itself, will have been transferred. Program was used with 5- and 6-year olds, and apparently worked well when a teacher was present to give constant and prompt reinforcement.

Chapter 10 gives a series of editorial notations that have proved useful in training programers and in marking copy.

Chapters 2, 3, and 9 have appeared in *The Journal of Educational Research* (55, 9, 1962) on pp. 513–18, 433–52, and 508–12, respectively. Chapter 8 has appeared in *Phychological Reports* (8, 1961) on p. 398.

GOLDBECK, ROBERT A. *The Effect of Response Mode and Learning Material Difficulty on Automated Instruction.* Santa Barbara, Calif.: American Institute for Research, 1960. 24 p. (mimeo).

See GOLDBECK and CAMPBELL (1962).

GOLDBECK, ROBERT A., and CAMPBELL, VINCENT N. "The Effects of Response Mode and Response Difficulty on Programed Learning." *Journal of Educational Psychology*, 53 (1962), p. 110–18.

Problem: Interaction between learning from different response modes and difficulty of program.

Sample: 63 and 62 junior high school students in first and second experiments, respectively.

Programs: 35 independent factual items; brief continuous program on light.

First experiment compared overt, covert, and reading (i.e., no blanks to fill in) response modes on versions of program constructed so as to differ in difficulty. Found significant interaction of difficulty with response mode, with overt response group performing *below* others at low difficulty level and *above* others at intermediate difficulty level. At most difficult level, the reading mode was highest, but not significantly so.

Second experiment tested the three response modes plus a fourth optional mode on continuous program rather than discrete items. Reading group was superior but not significantly so on immediate testing; was significantly superior on 10-week retention test.

In both experiments, the reading group learned most per time spent.

GOLDBECK, ROBERT A., CAMPBELL, VINCENT N., and LLEWELLYN, JOAN E. *Further Experimental Evidence of Response Modes in Automated Teaching.* Santa Barbara, Calif.: American Institute for Research, 1960. 24 pp.

See GOLDBECK AND CAMPBELL (1962).

GOLDBECK, ROBERT A.; SHEARER, JAMES W.; CAMPEAU, PEGGIE L.; and WILLIS, MARY B. *Integrating Programed Instruction with Conventional Classroom Teaching.* San Mateo, Calif.: American Institute for Research, 1962.

Problem: Effectiveness of programed material integrated with classroom teaching vs. classroom teaching without programed material.

Sample: 72 general and 78 college preparatory students in high school.

Program: Three segments of a course on the U.S. Government; linear program.

This is the final report of a project of which the earlier results have been reported elsewhere. The final field tests of the project indicated that a few minutes a day of programed instruction integrated with conventional classroom teaching could raise student performance on two of the three units (or "segments") significantly higher than conventional classroom teaching alone. Furthermore, student attitudes were favorable to programs used this way, and tended to become more favorable with longer acquaintance.

GOLDSTEIN, LEO S., and GOTKIN, LASSAR G. "A Review of Research: Teaching Machines vs. Programed Textbooks as Presentation Modes." *Journal of Programed Instruction*, 1, 1 (1962), 29–36.

Problem: Effectiveness of programed textbook vs. teaching machine.

Sample and programs: See table on p. 32 of article.

This review analyzes eight studies, six previously available in published form or as institutional reports (see EIGEN and KOMOSKI, 1960; EIGEN, FILEP, GOLDSTEIN, and ANGALET, 1962; HOLT and HAMMOCK, 1962; ROE, MASSEY, WELTMAN, and LEEDS, 1960; SILVERMAN and ALTER, 1962; and two others, by GOLDSTEIN and GOTKIN, reported in a

previous paper. Both experiments by GOLDSTEIN and GOTKIN used a program of 3,100 frames on spelling, in one case with 87 fourth-graders and in another case with 118 fourth-graders. Surveying all eight studies, the authors found no significant differences in mastery of subject matter between machine and programed text treatments. In four of the five studies, in which time to complete the program was a variable, significant differences were found in favor of the programed text.

GOTKIN, LASSAR G., and GOLDSTEIN, LEO S. *Programed Instruction for the Younger Learners: A Comparison of Two Presentation Modes in Two Environments.* New York: The Center for Programed Instruction, 1962. 19 p. (mimeo).

Problem: Relative effectiveness of teaching machine and programed textbook with and without supervision.

Sample: 215 fourth-grade students in 4 experimental groups.

Program: First 100 words of TMI (Teaching Machines, Inc.) linear spelling program.

In home (unsupervised) as well as classroom (supervised), no significant differences were found between machine and programed text presentation, for post-test, learning gain, or attitude toward the experience. In school environment, students working with the programed text in two schools completed 50 percent and 59 percent more frames respectively than students using the machine. Students using the machine at home completed as many frames as students in the classroom, whereas students using the programed text at home completed fewer frames than students in the classroom, but still more on the average than machine users.

GROPPER, GEORGE L., and LUMSDAINE, ARTHUR A. *An Experimental Comparison of a Conventional TV Lesson with a Programed TV Lesson Requiring Active Student Response. Studies in Televised Instruction, Report No. 2.* Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED–WQEX, and American Institute for Research, 1961a. 21 p.

Problem: Programed television lesson vs. conventional television lecture-demonstration.

Sample: 140 junior high school students on a "heat" program, 138 on "nuclear reactions;" experimental and control groups each divided into high-IQ and low-IQ groups.

Programs: Televised lessons on heat and on nuclear reactions.

Two television lessons were presented to students in two forms each—a conventional television lecture-demonstration and an experi-

mental "programed" lesson calling for active responses. In the lesson on heat, the experimental group scored significantly higher than the control group on an immediate post-test, but this difference was due almost entirely to the results in the high-ability group. The low-IQ experimental group was practically identical with the control group on the immediate post-test and was actually lower on the delayed test. The high-IQ group was significantly higher than its control on the delayed as well as the immediate test.

In six comparisons, based on alternate forms of a test administered at three different times after the telecast of the lesson on nuclear reactions, the experimental group was higher than the control on five. Two of the five comparisons were statistically significant, and a third was significant for some of the schools in the sample.

Results tend to support the effectiveness of active response, but to stress the importance of insuring *correct* response, especially with the low-ability group.

GROPPER, GEORGE L., and LUMSDAINE, ARTHUR A. *An Experimental Evaluation of the Contribution of Sequencing, Pre-Testing, and Active Student Response to the Effectiveness of "Programed" TV Instruction. Studies in Televised Instruction, Report No. 3.* Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX, and American Institute for Research, 1961b. 19 p.

Problem: Effectiveness of use of programing tactics in television teaching.

Sample: 150 junior high, 156 seventh- and eighth-grade, and 186 junior high students, respectively, in three studies.

Programs: Televised lessons on Newton's laws of motion, on movies, and on levers.

A report on three studies which some of the tactics of programed learning were used in fixed-pace television teaching. In one experiment, junior high school students who watched a lesson on Newton's laws of motion and who were encouraged to make active responses did not show a significant difference in learning from other students who watched the lesson but were not encouraged to make active responses. Both versions of the television presentation were like a conventional lecture. However, when another television lesson—this one on movies—was "programed," or sequenced, somewhat like a teaching-machine program, students who made active responses scored significantly higher than those who did not. The result is interpreted as meaning that active response makes a difference when the lesson is programed so that students respond *correctly* and actively.

A third experiment used responses of students as a device in trying out and revising television lessons on levers. Revising lesson so that students were better able to respond correctly resulted in significantly more learning.

See also GROPPER and LUMSDAINE, 1961c.

GROPPER, GEORGE L., and LUMSDAINE, ARTHUR A. *An Investigation of the Role of Selected Variables in Programed TV Instruction. Studies in Televised Instruction, Report No. 4.* Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX and American Institute for Research, 1961c. 29 p.

Problem: Effect of different response and confirmation procedures on programed television lessons.

Sample: Junior high school science students, numbering 190, 224, 150, and 220, respectively, in four experiments and divided in each case into experimental and control groups.

Programs: Televised lessons on airplanes, satellites, body chemistry, and how television works.

In each case, the experimental and control groups each watched one of two simultaneous telecasts utilizing two different treatments, which were as follows:

(1) Active student responses and competition with other students in the studio vs. no active responses and no such competition (lesson on airplanes).

(2) Active student responses but only intermittent feedback of results vs. no active responses and no feedback (lesson on satellites).

(3) Writing of responses vs. reading of items with blanks already filled in (lesson on body chemistry).

(4) Active responses following pause and question mark on screen vs. active responses, no pause, and a marker (triangle) on screen to emphasize key words (lesson on how TV works).

Significant differences were not found in any of these four comparisons; compare, however, GROPPER and LUMSDAINE, 1961a and 1961b.

GROPPER, GEORGE L., and LUMSDAINE, ARTHUR A. *Improvement of Televised Instruction Based on Student Responses to Achievement Tests. Studies in Televised Instruction, Report No. 1.* Pittsburgh: Metropolitan Pittsburgh Educational Television Stations WQED-WQEX and American Institute for Research 1961d.

Problem: Value of revising a lesson based on testing student's learning from the original version.

Sample: 6 junior high school classes per experiment, randomly assigned as units to two treatments in each of two experiments.
Programs: Televised lessons on heat and chemistry.

Lessons originally prepared by experienced TV instructors were revised on the basis of data from achievement tests following a preview showing of the original lessons. The original and revised lessons were then telecast simultaneously over two channels to two groups of classes. The classes that watched the revised lessons scored significantly higher on postprogram tests than those that watched the unrevised programs. Applicability to mass media is thus illustrated for procedures of program revision based on student response data, paralleling similar procedures in the case of individually paced learning programs.

See also GROPPER and LUMSDAINE, 1961b.

GUTHRIE, PETER M., and LUMSDAINE, ARTHUR A. *Some Effects of Graduated Partial Cueing on the Learning of Paired Associates.* San Mateo, Calif.: American Institute for Research, 1961.
Problem: Prompting vs. partial cueing vs. confirmation.
Sample: 48 volunteers, aged 16 to 35.
Program: Several lists of 16 city names, paired with their airport code letters.

Comparison of confirmation with partial prompting (stimulus term introduced first, but immediately followed by a weakened exposure of response term as a cue to correct response) showed a rather small, though statistically reliable, degree of superiority for partial cueing over confirmation. Some very good performances by a *fully* prompted group, however, lead authors to suggest that for many occasions, "immediate full prompting may frequently represent an optimum condition, and that the various partial-cueing methods simply provided varying approximations to this immediate, full-prompt condition." Conclusion: conditions under which partial prompt is superior to full prompt are not yet fully understood.

See also ANGELL and LUMSDAINE, 1960 and 1962.

HATCH, R. S. *An Evaluation of the Effectiveness of a Self-Tutoring Approach Applied to Pilot Training.* Wright-Patterson Air Force Base, Ohio: Aero-Space Medical Laboratory, 1959. 19 p.
Problem: Effectiveness of voluntary self-tutoring with a teaching machine.
Sample: Air Force pilots: 43 users of the device, 15 nonusers, 36 without access to it.
Program: Procedures of instrument flying; Pressey approach.

782-239 O—64——5

A Naval Automatic Rater teaching machine was programed with the intention of helping in the ready recall of a large body of inflight job information for Air Force pilots. A voluntary self-tutoring approach was employed, utilizing one type of "game appeal." Two matched groups of Air Force pilots were pretested on their knowledge of instrument flying information. The device was then installed in the crew lounge of one of the groups. No device was available to the other group. After a 2-month period all pilots were tested again. Despite the fact that minimal exposure to the machine occurred, all pilots who played the machine improved significantly on the criterion tests, whereas nonplayers did not improve.

HATCH, RICHARD S., and FLINT, LANNING L. *Programed Learning: A Comparative Evaluation of Student Performance Variables Under Combinations of Conventional and Automated Instruction*. New York: U.S. Industries, Educational Sciences Division, 1962 (mimeo).

Problem: Effectiveness of programed instruction plus conventional classroom teaching, vs. programed instruction alone, vs. conventional teaching alone.

Sample: 46 junior college students.

Program: One semester of basic electronics, intrinsic branching program, 3,478 frames, presented on an Auto Tutor teaching machine.

Group (1) received 3 hours of programed instruction and 3 hours of conventional classroom instruction each week. Group (2) received 3 hours of programed instruction; Group (3), 3 hours of conventional classroom instruction. On a criterion test at the end of the semester, there was no significant difference in the performance of groups (2) and (3), but group (1) performed about one standard deviation higher than either of the others. Academic intelligence appeared to predict achievement from programed instruction about as well as it predicts it from conventional teaching. Individual error rates, criterion performance, and intelligence were significantly intercorrelated.

HERRICK, MERLYN C. *The Effect of Problem-Setting Questions on Rate and Amount of Learning in Programing Teaching Machines*. Bloomington, Ind.: Indiana University, 1962 (mimeo).

Problem: Effect of problem-setting questions in programs.

Sample: 96 students in eighth-grade general science.

Program: Two scrambled-book programs on light and on color.

One group was given the usual expository programs; the other, the same expositions introduced by questions designed to "set" the prob-

lem. No significant differences were found, either in the learning gains or in the number of repetitions needed to complete the programs.

HICKEY, ALBERT E., and ANWYLL, B. JEAN. *Programed Instruction of Package Billing Clerks* (final report to Spiegel, Inc.). Lexington, Mass.: Information Technology Laboratories (Itek), 1961. 30 p. (mimeo).

Problem: Effectiveness of programed instruction in industry.
Sample: 120 employees.
Program: Procedures of package billing; 2,000 frames.

Use of programed instruction resulted in 34 percent reduction in average number of student hours needed to attain criterion level of performance.

HICKEY, ALBERT E.; AUTOR, SANFORD M.; and ROBINSON, EDWARD J. *Programmed Instruction Integrated with Broadcast ETV* (final report to Fund for the Advancement of Education). Boston: Northeastern University, 1962 (mimeo).

Problem: Value of programed instruction as an adjunct to broadcast ETV (Educational Television).
Sample: University extension students registered for Continental Classroom ETV course (n=389); high school students viewing same course (n=106).
Program: Probability and statistics; 4,000 frames.

San Diego high school students who received the program in addition to text and ETV lectures performed significantly better on examinations than those who did not receive program. No significant difference in program vs. non-program groups in sample drawn from several universities, although 78 percent of the university students said the program was "helpful" or "very helpful."

HICKEY, ALBERT E., and LAIDLAW, WILLIAM J. *Programed Instruction in Retail Sales* (final report to U.S. Navy Bureau of Supplies and Accounts, ONR Contract Nonr-3630(00)). Newburyport, Mass.: ENTELEK Incorporated, 1962 (mimeo).

Problem: Effectiveness of programed instruction in U.S. Navy.
Sample: U.S. Navy Supply Officers (n=130).
Program: Retail sales and ship's store management.

Supply Corps School students who used the adjunct program saved 56 percent of usual homework time (17 percent of usual overall study time) in reaching performance criterion. Instructors' lecture hours were reduced 54 percent. Attitude generally favorable.

HICKEY, ALBERT E.; LAIDLAW, WILLIAM J.; TAUB, HARVEY A.; NEWTON, JOHN M.; and TEICHNER, WARREN H. *Requirements for Graphic Teaching Machines* (final report to Commissioner of Education, U.S. Office of Education, Department of Health, Education, and Welfare, Grant No. 7-31-0570-161). Boston: Northeastern University, 1962 (mimeo).

Problem: (a) Relative effectiveness of graphics and symbols in concept formation; (b) classification of graphics in textbooks.

Sample: (a) High school seniors (n=16), college freshmen (n=16); (b) high school textbooks in algebra, biology, and music (n=16).

Program: Boolean algebra.

(a) An experiment is reported which demonstrates that graphics are more effective than symbols in the acquisition of some concepts from Boolean algebra; (b) graphics in high school textbooks can be reliably classified in a matrix of 480 functional stimulus-response categories.

HOLLAND, JAMES G. *Design and Use of a Teaching Machine* (paper presented to American Psychological Association Symposium, 1960).

Cambridge: Harvard University, 1960 (mimeo).

Problem: Experiment 1—effectiveness of items with confirmation vs. items with no confirmation vs. complete statements with no blanks to fill; Experiment 2—effectiveness of asking for significant vs. trivial vs. ambiguous and difficult responses.

Sample: University students—Experiment 1 (n=*n.a.*); Experiment 2 (n=36).

Program: 22 sections of Holland-Skinner psychology program.

In experiment 1, one group had the usual kind of items with one or two words missing; this group wrote answers and received confirmation when the answers were correct. A second group wrote answers to the same items, but received no confirmation. A third group read the same material, cast in the form of complete statements with the blanks filled in; this group, of course, wrote no answers. The third group made more errors than the other groups on a post-test, and there was very little difference between the performances of groups 1 and 2 (significance figures not given). Group 3, not having to write answers, completed the program in less time than the other groups. This is interpreted to mean that a response is an important part of an item.

In experiment 2, one group had the normal program with blanks left that called for material related to the critical information in the

item. A second group had a program in which the items were identical with those on the normal program except that the blanks left called for material of a trivial and easy nature; thus the student could easily answer without having noticed the critical information in the item. A third group had the same items except that blanks were left which rendered the items ambiguous and difficult and made for many errors. A fourth group read complete statements with no blanks to fill. Group 1 made the fewest errors on a post-test, and group 3 made the most (no significance figures). This is interpreted to mean that a program should be written so as to keep the error level low, but to make the answer depend on noticing the critical material.

HOLLAND, JAMES G., and PORTER, DOUGLAS. "The Influence of Repetition of Incorrectly Answered Items in a Teaching-Machine Program." *Journal of the Experimental Analysis of Behavior*, 4, 4 (1961), pp. 305–7.

Problem: Effect on learning from a program allowing repetition of items that were first answered incorrectly.

Sample: 14 graduate students.

Program: Holland-Skinner psychology program.

Students were divided into two groups, one of which answered all the items once only and one of which had an opportunity to repeat items missed the first time. Tests consisted of completion items from the program. At every level of difficulty of item, the non-review group made more errors than the review group (significant at 0.01 level). These differences were maintained on a 6-month retention test. The differences in performance between the two groups were almost identical on both the first and second testings.

HOLT, HOWARD O., and HAMMOCK, JOSEPH. "Books as Teaching Machines: Some Data." In *Applied Programed Instruction*, ed. by Stuart Margulies and Lewis D. Eigen. John Wiley and Sons, 1962, p. 50–56.

Problem: The relative effectiveness of programed books and simple teaching machines.

Sample: 63 telephone company technicians.

Program: Basic electricity; 2,600 frames.

No significant difference found between learning from programed text and teaching machines.

HOLT, HOWARD O., and VALENTINE, C. G. *An Exploratory Study of the Use of a Self-Instruction Program in Basic Electricity*

Instruction. Report to Bell Telephone Laboratories, 1961, Murray Hill, N.J. 27 p. plus appendix.

Problem: Construct a self-instructional program covering a large block of instruction (basic electricity), and examine the effectiveness of that program in operational context.

Sample: 64 telephone company technicians: 34 in experimental group and 30 in conventional group.

Program: Basic electricity; self-instructional.

Construction of program was successful as evidenced by comparison of time and proficiency of trainees taught by self-instruction program and those taught by conventional instruction. Time taken to complete the training was the same on the average for both groups. As measured on two final examinations given immediately after training and 6 months later, proficiency of program-taught trainees was significantly greater than that of conventionally taught trainees.

HOUGH, JOHN B. "An Analysis of the Efficiency and Effectiveness of Selected Aspects of Machine Instruction." *Journal of Educational Research*, 55 (1962a), 467–71.

Problem: Teaching machine program vs. lecture, and constructed vs. multiple-choice response vs. combination of constructed and multiple-choice.

Sample: 41 junior and senior college students in education.

Program: Linear program on the secondary school.

There were no significant differences in any of the comparisons, when the criterion was the amount of learning.

HOUGH, JOHN B. "Research Vindication for Teaching Machines." *Phi Delta Kappan*, 42 (1962b), 240–42.

Problem: Effectiveness of programed instruction.

Sample: Education students, 20 experimental, 21 control.

Program: Part of course on "the contemporary secondary school."

Students taught by programed instruction were compared with control group, taught by lecture-discussion method. Two early 50-question tests, and a 150-question schedule used both as pre-test and post-test, revealed no significant difference in learning between the two groups. The programed instruction group, however, saved an average of 47 percent of the time required by the lecture-discussion group.

HOUGH, JOHN B., and REVSIN, BERNARD. "Programed Instruction at the College Level: A Study of Several Factors Influencing Learning." *Phi Delta Kappan*, 44, 6 (1963), 286–91.

Problem: Teaching machine vs. programed textbook, knowledge

of results vs. no knowledge of results, differences between low- and high-scorers.

Sample: 90 university students in education.

Program: 555 frames, linear program, on "Historic Foundations of the Secondary School."

No significant differences were found between any of three experimental groups using, respectively, teaching machines, programed textbooks, or programed texts without knowledge of results. High- and low-scoring students did not differ significantly on verbal ability, attitudes toward programed instruction, attitudes toward the content, or on the Thurstone Temperament Scale.

HOVLAND, CARL I.; LUMSDAINE, ARTHUR A.; and SHEFFIELD, FRED D *Experiments on Mass Communication.* Princeton, N.J.: Princeton University Press, 1949.

Problem: Active response vs. passive viewing of teaching film.

Sample: 742 soldiers.

Program: A teaching film on military phonetic alphabet.

In chapter 9 an experiment is reported comparing active response (audience participation) with passive viewing of review sequences of a teaching film. In the experimental version, the viewers were encouraged to call out the correct equivalents (Able for A, and so forth) when the letters were presented. In the control version, both letters and their equivalents were shown on the screen, and there was no overt responding. The procedure of actively responding was significantly more effective, as shown by both oral and written tests. The measured advantage of the active-response procedure was greater for more difficult than for easier material, for less intelligent than for more intelligent learners, and for less motivated classes than for those which had the incentive of knowing in advance that they would be tested.

See follow-up studies by MICHAEL AND MACCOBY, 1953, and, especially, by MACCOBY, MICHAEL AND LEVINE, 1961.

HUGHES, JOHN L. *A Study of the Effectiveness of Programed Instruction in Basic Customer Engineering Training.* Endicott, N.Y.: International Business Machines Corporation, 1962. 10 p.

Problem: Effectiveness of programed instruction.

Sample: 150 IBM employee-students taught by programed instruction, 57 taught by conventional classroom method.

Program: Basic customer engineering training, 352 frames, linear,

After 6¼ hours of training, both groups were given an achievement test. Programed instruction class averaged 87.4 points on test; conventional class 80.8. Saving in time for students of programed

instruction ranged from 5 to 50 percent. Students who took less time to complete course usually scored higher.

HUGHES, JOHN L. "Effect of Changes in Programed Text Format and Reduction of Classroom Time on the Achievement and Attitude of Industrial Trainees." *Journal of Programed Instruction*, 1, 1 (1962), 43–54.

Problem: Effect of overt vs. covert response, position of confirmation, answers written on text or pad, and classroom time.

Sample: 9 classes of IBM trainees, divided into 4 experimental groups of 33, 28, 44, and 24 adults respectively; also, six classes with 70 students, as control group.

Program: 719 frames on basic computer knowledge; linear.

All experimental classes were allowed only 8 class hours a week; the control group classes, 11. The experimental groups were set up to test the effect of presenting answers on the back of the page or the next recto page; of requiring written answers in the text or on a separate paper; and of requiring written answers as opposed to a covert response. No significant differences in learning were found among the groups, but covert responding reduced the time by about 20 percent.

HUGHES, JOHN L. *Programed Instruction: A Follow-up Study of 7070 Training.* New York: International Business Machines Corporation, 1961. 23 p.

Covert response mode was more efficient than overt.
See HUGHES, 1962.

HUGHES, JOHN L., and MCNAMARA, W. J. "A Comparative Study of Programmed and Conventional Instruction in Industry." *Journal of Applied Psychology*, 45 (1961), 225–31.

Problem: Programed text vs. conventional classroom.

Sample: IBM employees—70 in experimental groups, 40 in control.

Program: 719 frames on use of 7070 computer.

Two classes were taught elements of IBM 7070 use by programed textbook, two control classes by conventional classroom instruction. Programed textbook classes completed work in mean of 8.8 hours, with large individual variation. Conventional classes took 15 hours. Experimental group scored slightly higher. Difference was significant even after partialing out effect of difference in reasoning ability. Students were very favorable to programed work.

IRION, ARTHUR, and BRIGGS, LESLIE J. *Learning Task and Mode of Operation Variables in Use of the Subject Matter Trainer.* Lowry Air Force Base, Colo.: Air Force Personnel and Training Research Center, October 1957, 19 p.

Problem: Amount of learning resulting from different prompting and response modes.

Sample: First experiment, 240 high school students in 12 groups; second experiment, 180 high school students.

Programs: Three "subject matter trainer" (teaching machine) routines: learning names of 20 map symbols, order of 20 steps in processing photographic film, answers to simple navigational problems.

The first experiment (four treatments in each of the three programs) found that the different prompt-response treatments, in descending order of learning effectiveness, were as follows:

(1) Prompted quiz (students were told what the right answers were).
(2) Modified quiz (students were told if their answer was right; if it was not, they were given the right answer).
(3) Practice (students were told only when they gave the right answer).
(4) Single try (students were permitted only one try and were told whether that answer was right or wrong).

A second experiment compared combinations of the first, or most effective, mode with each of the others in turn, but found no significant differences.

For a later version of these data, see BRIGGS, in LUMSDAINE, ed., 1961.

ISRAEL, MATTHEW L. "Variably Blurred Prompting: I. Methodology and Application in the Analysis of Paired-Associate Learning." *Journal of Psychology*, 50 (1960), 43–52.

Problem: Amount of prompting needed at various stages of learning.

Sample: 18 high school and college students, under age 20.

Program: Paired English words and German equivalents.

Varying the physical clarity of the prompt, investigator found that, as student worked through program, successively smaller amounts of prompting were required to attain correct responses.

JENSEN, B. T. "An Independent-Study Laboratory Using Self-Scoring Tests." *Journal of Educational Research*, 43 (1949), 134–37.

Problem: Effectiveness of punchboard as review testing device.
Sample: 24 mature students.
Program: Review tests on psychology.

Students completed a 5-semester-hour course in psychology in 5 weeks in an independent study laboratory, making much use of an extensive set of Pressey punchboard tests covering the whole course. Fifty-four percent of these students made A on the final exam, as compared to 10 percent in the sections taught in the usual manner (IQ not controlled).

KAESS, WALTER, and ZEAMAN, DAVID. "Positive and Negative Knowledge of Results on a Pressey-type Punchboard." *Journal of Experimental Psychology*, 60 (1960), 12–17.
Problem: Effect of punching wrong answers on punchboard.
Sample: 554 college students.
Program: Definitions of psychological terms.

An experiment in learning psychological terms with a simulated Pressey punchboard. Subjects were given one trail on multiple-choice items that presented different numbers of choices—five, four, three, two, and one. Results were evaluated by four trials on five choice items. The number of errors was directly related to the number of alternatives available during training. Reading the correct answers was as effective as *reading plus punching.*

KANNER, JOSEPH H., and MARSHALL, W. P. *The Improvement of Television Instruction by Review Procedures.* (In press) Office of the Chief Signal Officer, Washington, D.C. May 1963.
Problem: Effect of review procedures upon student learning.
Sample: 1,700 Army basic trainees.
Program: Television recordings teaching 58 hours of military subjects by means ranging from lectures to student manipulation of equipment.

Trainees receiving additional instruction through a "review-preview" technique employing television recordings achieved significantly better than those taught by conventional procedures. Low-aptitude students receiving television instruction learned at least as well as high-aptitude students taught by conventional classroom means. Differences were significant at least at .05 level.

KANNER, JOSEPH H., and SULZER, RICHARD L. *"Overt and Covert Rehearsal of Fifty Per Cent vs. One Hundred Per Cent of the Material in Film Learning."* In LUMSDAINE, ed., 1961, 427–41.
Problem: Effectiveness of student participation in different amounts.

Sample: 2,600 Air Force trainees.
Program: Film teaching military phonetic alphabet.

Active review by students—"participation"—was found to be significantly superior to passive review. No significant differences were obtained, however, between oral recitation, written practice, "thinking" answers, and "thinking" plus feedback of results. However, the thinking-plus-feedback was superior to the thinking-without-feedback at the .06 level.

KEISLAR, EVAN R. "The Development of Understanding in Arithmetic by a Teaching Machine." *Journal of Educational Psychology*, 50 (1959), 247–53.
Problem: Effectiveness of a program in teaching understanding (defined here as broad transfer).
Sample: 14 elementary school children.
Program: 110 multiple-choice items on area of rectangles.

Experimental subjects did significantly better than uninstructed matched controls on a test involving completion items and essay-type questions.

KEISLAR, EVAN R., and MCNEIL, JOHN D. "A Comparison of Two Response Modes in an Autoinstructional Program with Children in the Primary Grades." *Journal of Educational Psychology*, 53 (1962), 127–131.
Problem: Overt vs. implicit response.
Sample: 198 primary-grade children in two groups, plus a control group of 57.
Program: 3-week unit in physical sciences; 432 frames.

Both experimental groups (oral and implicit response) learned significantly more than the uninstructed control group, but no significant difference between overt or implicit response groups was found.

KEISLAR, EVAN R., and MCNEIL, JOHN D. "Teaching Scientific Theory to First Grade Pupils by Auto-Instructional Device." *Harvard Educational Review*, 31, 1 (1961), 73–83.
Problem: Effectiveness of programed instruction, administered by an autoinstructional device using both sight and sound, in helping first-grade students to acquire and use certain aspects of scientific theory in explaining physical phenomena.
Sample: 26 first-grade children, divided equally into experimental and control groups.
Program: 432 items on molecular theory as related to evaporation and condensation, divided into 13 daily lessons.

After working through the program, the experimental group demonstrated significantly more learning than the control group (which had received no special instruction in the topic). Success on the post-test was significantly (and negatively) related to the number of errors made during the program. Authors conclude that, while there are great individual differences, first-grade pupils can learn an abstract scientific language and can learn to use this language as part of the chain of intraverbal responses in explaining physical phenomena. Even though the multiple-choice method was used as the sole means of responding, most of the children were able later to use the new vocabulary in explaining problems of evaporation and condensation. The autoinstructional device held the interest of the students for the 3 weeks of the project.

KIMBLE, GREGORY A., and WULFF, J. JEPSON. "'Response Guidance' as a Factor in the Value of Student Participation in Film Instruction." In LUMSDAINE, ed. (1961a), 217–26.

Problem: Guided vs. unguided responses during student participation in a learning task.

Sample: 660 airmen.

Program: Film teaching the use of the slide rule.

Guided responses, making for a minimum number of errors during practice, were found to result in higher proportion of correct test answers and a reduction of common errors.

KIMBLE, GREGORY A., and WULFF, J. JEPSON. "The Effectiveness of Instruction in Reading a Scale as Influenced by the Relative Amounts of Demonstration and Problem Solving Practice." In LUMSDAINE, ed. (1961b), 227–39.

Problem: What is optimum amount of active practice in learning a skill from a programed demonstration?

Sample: In first experiments, six replications with groups of 14 or 15 Air Force trainees; in second, 5 replications with groups of 40 to 60.

Program: Instructional film on using slide rule.

Results indicate that an intermediate amount of active practice (finding correct setting and drawing line at correct point on picture of slide rule) may be superior to less or more practice. In first experiments, practicing 8 of 16 examples was superior to practicing more or less examples. In the second, the trend was likewise toward a middle point optimum, but differences were nonsignificant.

KLAUS, DAVID J., and LUMSDAINE, ARTHUR A. *An Experimental Field Test of the Value of Self-Tutoring Materials in High*

School Physics (an interim report of progress and findings). Pittsburgh: American Institute for Research, 1960.
Problem: Effectiveness of programed materials as supplement to telecasts.
Sample: 450 high school students.
Program: Linear program on elementary physics; about 3,000 frames.

Students in randomly assigned classes were furnished programed materials to supplement the Harvey White physics telecast, which they were following in addition to their regular classroom work. Comparison of these experimental classes with a random selection of otherwise similar control classes which were not given the programed materials was made by means of a comprehensive examination at the end of the ninth unit of materials. The programed materials produced a significant gain in amount of learning from the course.

KNIGHT, M. A. G., and TILLEY, K. W. Report on experiment, in *Teaching Machines and Programmes*, a bulletin prepared by the Department of Psychology, University of Sheffield. Sheffield, England: January 1962.
Problem: Time required to learn to criterion, programed instruction vs. conventional instruction.
Sample: 29 Royal Air Force trainees in experimental group, 30 in control.
Program: Trigonometry.

Using Auto Tutor, experimental group learned to same criterion in about half the time required by control group. (No details given. Full report not available.)

KOPSTEIN, FELIX F., and CAVE, RICHARD T. *Preliminary Cost Comparison of Technical Training by Conventional and Programed Learning Methods.* Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. Technical Documentary Report No. MRL-TDR-62-78, 1962a. 19 p.
Problem: Cost of programed vs. conventional teaching.
Sample. Air Force students in electronics.
Program: 19-week course in Communications Electronics Principles.

The difficulties and uncertainties in estimating comparative costs of automated vs. conventional instruction are candidly stated. Costs are estimated in three different ways, and the costs of auto-

mated instruction compare favorably with those of conventional teaching regardless of the method of estimation. Automated teaching requires a large initial expenditure, but costs per student diminish as the number of students increases. Costs per student of conventional teaching, on the other hand, remain relatively constant. Therefore, there appears to be a point of equal cost, after which automated teaching appears to be cheaper.

KOPSTEIN, FELIX F., and CAVE, RICHARD T. *Preliminary Evaluation of a Prototype Automated Technical Training Course.* Behavioral Sciences Laboratory, 6570th Aerospace Medical Research Laboratories, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio. Technical Documentary Report No. MRL-TDR-62-78, 1962b. 22 p.

Problem: Effectiveness of automated vs. conventional instruction.

Sample: Experimental, control, and blind-control groups of 14 airmen each.

Program: Intrinsic program on Communications Electronics Principles presented on teaching machine.

Experimental group received major instruction from Auto Tutor, but had usual practical problems and homework. Control groups had classroom teaching plus homework plus practical problems. Blind-control group did not know it was participating in project; other control group did know. Study was done twice. Such significant differences as were found favored control group, but these were mainly in the earlier part of the course. Authors conclude, "It seems justified to conclude that students trained with the Auto Tutors progressed at a rate and to a level nearly comparable with the students instructed in the conventional fashion."

KOPSTEIN, FELIX F., and ROSHAL, SOL M. "Methods of Presenting Word Pairs as a Factor in Foreign Vocabulary Learning." (Abstract) *American Psychologist*, 10 (1955), 354.

Program: Prompting vs. confirmation.

Sample: 2,080 basic airmen in basic training.

Program: 8 foreign words with English equivalents.

Found that paired-associate words were learned more quickly when paired terms were presented together rather than consecutively. Advantage for contiguous presentation is greater in early stages of training, less as substantial levels of learning are reached.

KOPSTEIN, FELIX F., and ROSHAL, SOL M. "Verbal Learning Efficiency as Influenced by the Manipulation of Representa-

tional Response Processes: Pictorial-Verbal and Temporal Contiguity Factors." In LUMSDAINE, ed., 1961, 335–50.

Program: Effect on learning paired associates of using pictorial or verbal stimuli for verbal responses, and of presenting stimulus and response terms simultaneously or in sequence.

Sample: 2,080 airmen.

Program: Paired associates; response terms were eight Russian words, stimulus terms were either picture or English equivalents of these words.

The use of pictures as stimulus terms proved to be uniformly more effective than the use of the equivalent words. Neither embedding the terms in contextual sentences nor testing with pictures or with words contributed any significant variance. There was a high degree of transfer from picture training to word testing, and vice versa. A simultaneous presentation of stimulus and response terms was more effective than a staggered presentation. However, the superiority of the simultaneous presentation diminished at about the 50 percent mastery level.

An abstract of the original paper was published in the *American Psychologist*, 10 (1955), 354.

KORMONDY, E. J. *Oberlin Experiments in Self-Instruction in General Zoology: Preliminary Analyses.* Cited in *Automated Teaching: A Review of Theory and Research*, by Robert E. Silverman (Port Washington, N.Y.: U.S. Naval Training Device Center, 1960. 17 p.).

No significant differences, overt vs. covert response modes. Details not available.

KRUMBOLTZ, JOHN D. *The Nature and Importance of the Required Response in Programed Instruction.* Paper presented to American Educational Research Association, Chicago, February 1963. 11 p. (mimeo).

Problem: Effectiveness of key-word response vs. trivial-word response in programed instruction vs. reading of same material in paragraph form (no requested response).

Sample: 53 college undergraduates and 67 graduate students.

Program: 89 frames, linear, designed to teach some fundamentals of educational measurement.

Samples were divided into four groups each: (a) A group which responded with a key word—i.e., an important concept—to each frame, (b) a group which responded with a trivial or minor word, (c) a group which read the identical material written as paragraphs of

textbook prose, and (d) a control group. On immediate and delayed post-test, the key-word and paragraph-format scored about the same, and both were significantly higher than the trivial-word group.

KRUMBOLTZ, JOHN D. and BONAWITZ, BARBARA. "The Effect of Receiving the Confirming Response in Context in Programed Material." *Journal of Educational Research*, 55 (1962), 472-75.

Problem: Effect on learning from programs presenting confirmatory response as complete thought or as isolated, single word or phrase.

Sample: 32 undergraduates in educational psychology.

Program: Test writing, linear, 153 frames.

No significant differences were found in knowledge of terms that were being taught, but group that received confirmation in context of complete sentence was significantly better able to apply principles learned from program.

KRUMBOLTZ, JOHN D., and WEISMAN, RONALD G. "The Effect of Intermittent Confirmation in Programed Instruction." *Journal of Educational Psychology*, 53, 6 (1962), 250-53.

Problem: Effect of intermittent confirmation, in fixed and variable ratios, on learning from programed instruction.

Sample: 121 college students in educational psychology.

Program: 177 frames of a linear program, in programed textbook form, designed to teach prospective teachers how to interpret educational test results.

Four levels of fixed-ratio confirmation (no confirmation, 33 percent confirmation, 67 percent confirmation, and continuous confirmation), and two levels of variable-ratio confirmation (33 percent and 67 percent) were used in different treatment groups. A negative linear relationship was found between number of errors made on the program and the amount of confirmation provided. No significant effects were found on post-test scores from the various proportions of confirmation, and no evidence of differential effect between fixed-ratio and variable-ratio confirmation.

KRUMBOLTZ, JOHN D., and WEISMAN, RONALD G. "The Effect of Overt vs. Covert Responding to Programed Instruction on Immediate and Delayed Retention." *Journal of Educational Psychology*, 53, 2 (1962), 89-92.

Problem: Overt vs. covert response.

Sample: 54 undergraduates in educational psychology.

Program: Fundamentals of educational testing, 177 frames, linear.

Subjects were divided into a group which wrote down each response, a group which mentally composed each answer but did not write it, a group which read the program with the blanks already filled, and a control group which studied a different program of about the same length. On an immediate post-test there was no significant difference among the experimental groups, but on a 2-week-delayed retention test the group which wrote answers scored significantly better than the others. Thus, overt responding seemed to increase delayed retention. The control group was significantly lower on both tests.

LAMBERT, PHILIP; MILLER, DONALD M.; and WILEY, DAVID E. "Experimental Folklore and Experimentation: The Study of Programed Learning in the Wauwatosa Public Schools." *Journal of Educational Research*, 55 (1962), 485–94.

Problem: Overt and covert response modes, and levels of intelligence.

Sample: 552 ninth-grade students.

Program: Sets, relations, and functions; linear; 864 frames.

Intelligence was found to be significantly associated with the amount of information acquired from the program. Covert responses proved to require less time than overt ones and to result in about as much learning (n.s.d.).

LEVIN, GERALD R., and BAKER, BRUCE L. "Item Scrambling in a Self-Instructional Program." Providence, R.I.: Brown University, 1962. 22 pp. (mimeo). Also in *Journal of Educational Psychology*, 54, 3 (1963).

Problem: Effect of scrambling item sequence in an instructional program.

Sample: 36 second-grade pupils.

Program: 180 frames on geometry for the second grade, 1 unit of which (60 frames) was presented either in logical or in random order.

No significant differences were found on acquisition, retention, or transfer between the groups which studied the ordered, and those which studied the randomized, unit.

LEWIS, EARL N. *Experimentation in the Development of More Effective Methods of Teaching Foreign Languages by Making Extensive Uses of Electro-Mechanical Aids*. Baton Rouge, La.: Department of Foreign Languages, Louisiana State University, 1961. 19 p.

Problem: Effectiveness of delegating some of teaching process to automated, programed audiovisual equipment.

732-239 O—64——6

Sample: Experiment I, 46 experimental, 52 control; experiment II, 35 experimental, 25 control; all college students.

Program: Locally written programed textbook integrated with language laboratory.

No significant differences found in either experiment between control groups, which were taught conventionally, and experimental groups, which spent a larger proportion of their time on the automated program, thus freeing the teacher for other work.

LIPPERT, HENRY T., and STOLUROW, L. M. *Teaching Machines and Programed Instruction in Special Education.* University of Illinois, Training Research Laboratory (undated—probably written in 1963).

Problem: Review of use of programed instruction in special education; report of findings of a study on combined prompting and confirmation in teaching sight vocabulary.

Sample: 25 educable mentally handicapped public school children.

Program: Picturable nouns were taught by the strategy of prompting trials followed by confirmation trials to various learning criteria. A correction procedure was used with the 900-frame program used in an earlier study by Stolurow and Lippert.

Binet Mental Age was not a significant predictor of the learning measures. It was, however, significantly related to retention. The subtest scores of the Illinois Test of Psycholinguistic Abilities were not heavily weighted in the performance on this learning task. This was interpreted to mean that the learning task does not demand high psycholinguistic abilities for success. Some psycholinguistic entry levels were significantly related to successful performance in recalling (but not recognizing) words taught and some were not. Those which were significant were the visual decoding, visual motor association, visual motor sequential, and the auditory vocal automatic abilities. Other psycholinguistic abilities were not heavily weighted or in some cases were negatively related.

The length of the words (from 3 to 9 letters) was not related to ease of learning or to successful retention. The Peabody Picture Vocabulary Test was positively related to ease of learning, and a high score on this test was significantly associated with successful transfer to success after the withdrawal of the prompt under the confirmation teaching strategy. The relative number of prompting and confirmation trials did not seem to be as important in predicting learning as the total number of trials, regardless of the balance.*

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

LITTLE, J. K. "Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology." *Journal of Experimental Education*, 3 (1934), p. 45–49.

Problem: Effectiveness of quick knowledge of results.
Sample: 340 college educational psychology students.
Program: Educational psychology.

Students in educational psychology were taught in three groups: (1) By conventional classroom methods, (2) by ordinary classroom procedures plus very quick machine grading of their quizzes after which they were given a chance to make up deficiencies by additional tests, (3) by ordinary classroom procedures plus a drill machine which instantly apprised them of results and let them practice until they could always get the item right. Conditions (2) and (3) made for significantly better final examination scores than (1). Investigator concluded that greatest benefit from these two methods accrues to students in lower half of distribution.

LUMSDAINE, A. A., ed. *Student Response in Programed Instruction. A Symposium on Experimental Studies of Cue and Response Factors in Group and Individual Learning from Instructional Media*. Sponsored by Headquarters, Air Research and Development Command, U.S. Air Force. Washington, D.C.: National Academy of Sciences, National Research Council, 1961.

Papers from this important collection are listed in this bibliography under the names of individual authors. Lumsdaine's concluding chapter summarizes and brings together common theoretical factors investigated in the studies reported in the preceding 30 chapters.

LUMSDAINE, A. A.; SULZER, RICHARD L.; and KOPSTEIN, FELIX F. "The Effect of Animation Cues and Repetition of Examples on Learning from an Instructional Film." In LUMSDAINE, ed., 1961, p. 241–69.

Problem: (1) Effectiveness of animation in programed instructional film; (2) effectiveness of increasing number of examples.
Sample: 32 classes of Air Force basic trainees, assigned as units to experimental treatments.
Program: teaching film on micrometer use.

Overlay animation devices (moving arrows, superimposed labels, etc.) in a film on the micrometer contributed significantly to learning; increasing the number of examples from 3 to 6 to 10 also contributed significantly, but the *rate* of improvement diminished as the number of examples was increased, suggesting a saturation point in the use of additional examples. Giving a short pretest on micrometer reading

(without knowledge of results) also added significantly to learning of this skill. These results were consistent across a wide range of conditions in a $2\times2\times2$ factorial design. The saturation point for repetition was reached *sooner* for the less intelligent learners than for the more intelligent, suggesting that programing techniques other than mere repetition are needed to improve the performance of the former group.

MACCOBY, NATHAN; MICHAEL, DONALD N.; and LEVINE, SEYMOUR. "Further Studies of Student Participation Procedures in Film Instruction: Review and Preview Covert Practice, and Motivational Interactions." In LUMSDAINE, ed., 1961, p. 295-325.

Problem: Causes of student participation effects.

Sample: 993 Air Force trainees.

Programs: A film on "Patterns for Survival" and one on "Airplanes Change the World Map."

Among the results of these two rich experiments were the following:

(1) When review questions were asked in participation sessions and the learners merely thought about (covertly practiced) the answers, significantly more learning resulted if and only if the instructor supplied the correct answers after the students had responded.

(2) Evidence was found that the additional learning produced by participation was of the "meaningful" rather than the "rote" type.

(3) Participation gains are overwhelmingly due to the added practice generated by the procedure and not to increased motivation when the general level of motivation is already very high. When the general level of motivation is low, a relatively small part of the participation effect is due to the increased motivation generated by the procedure.

MACCOBY, N., and SHEFFIELD, F. D. "Theory and Experimental Research on the Teaching of Complex Sequential Procedures by Alternate Demonstration and Practice." In Symposium on Air Force human engineering, personnel, and training research. Ed. by G. Finch and F. Cameron. Washington, D.C.: National Academy of Sciences, National Research Council, 1958.

Problem: Size of step and control of pacing, as they affect learning from a programed film.

Sample: Experiment I, 4 groups of 40 undergraduates individually seen; experiment II, 40 classes ($n=900$) of undergraduates, plus 4 groups of high school students.

Program: Filmed demonstration of an assembly task.

For superior students, a self-pacing procedure in which the student himself regulated the length of sequence before practice worked best. Worst treatment of all was the demonstration of the whole procedure

before practice. It was theorized that small demonstration-practice units would result in fewest errors, but that longer units would be necessary to provide integrated learning. In a series of interrelated experiments the predictions were confirmed. The optimum procedure turned out to be a transition one in which the beginning steps were short and the subsequent ones progressively longer until the entire demonstration was run through prior to practice.

MAGER, ROBERT F. "On the Sequencing of Instructional Content." *Psychological Reports* (1961), p. 405–13.

Problem: Sequence of learning topics and exercises as determined by learner, rather than by instructor.

Sample: 6 adults.

Topic: Electronics.

Each of six adults was given an opportunity individually to learn about electronics in whatever sequence of topics he wished to follow. The instructur behaved as a response mechanism, answering questions and giving information as asked by the learner. The question was whether a learner-generated sequence would be similar to an instructor-generated sequence, and whether there would be any commonality among sequences generated by independent learners. Considerable commonality was found, even though no specific objectives within the broad field of electronics were set for the learners. However, the learner-generated sequences bore little resemblance to that used in most introductory electronics curricula. In general, the learner proceeded from a simple whole to a more complex whole, whereas traditional sequences proceed from parts to whole, i.e., from components to system. The learners' motivation increased with the apparent control they were allowed to exercise over the learning experience.

See also others papers in which Mager is senior author (notably Explorations in student controlled instruction, Mager and Clark, 1963; and Mager and McCann, Learner-controlled instruction, no date, both offset) describing further trials of student-controlled learning sequences. These trials resulted in reducing training time as much as 65 percent, and resulted in well-equipped and confident students. The content selected for study and the sequence varied from student to student, but in no case coincided with the sequence used previously.

MARGOLIUS, GARRY J., and SHEFFIELD, FRED D. "Optimum Methods of Combining Practice with Filmed Demonstration in Teaching Complex Response Sequences: Serial Learning of a Mechanical Assembly Task." In LUMSDAINE, ed., 1961, p. 33–53.

Problem: Most effective combinations of practice with demonstration; size of step.

Sample: Experiment 1, 40 male college undergraduates; Experiment 2, 30 male college undergraduates.

Program: Film on how to assemble an automobile ignition distributor.

Experiment 1 compared (*a*) practice after each relatively short segment of the film, the segments having been proved in pretest to be assimilated on the first trial by about 75 percent of the type of subjects used, (*b*) practice after longer segments, (*c*) practice only after the entire film had been shown, (*d*) practice at first after short segments as in (*a*), then after gradually lengthening segments. Treatments (*a*) and (*d*) were superior to the others in teaching accurate performance, and all others were superior to (*c*).

Experiment 2 compared two forms of gradually increasing steps with a third form, in which the student was allowed to set his own length of step, i.e., to stop the film whenever he felt he was ready to practice what he had seen. The form of gradually lengthening steps used in experiment 1–*d* proved significantly superior to the self-pacing method in rate of improvement, and superior, though not at the .05 level, in overall learning.

The advantages of perfect practice, which are gained with short demonstration-practice periods, can be combined with the advantages of whole demonstration-practice for maximum task integration. This can be achieved best by a skillful combination of these methods. Self-pacing, when done by bright students, approaches this optimum.

MARGOLIUS, GARRY J.; SHEFFIELD, FRED D.; and MACCOBY, NATHAN. Repetitive versus Consecutive Demonstration and Practice in the Learning of a Serial Mechanical-Assembly Task." *In* LUMSDAINE, ed., 1961, p. 87–100.

Problem: (Considering d as demonstration; p as practice; and 1, 2, and 3 as consecutive steps) d_1p_1, d_2p_2, d_3p_3, d_1p_1, d_2p_2, etc. vs. d_1p_1, d_1p_1, d_2p_2, d_2p_2.

Sample: 20 male college undergraduates.

Program: Experiment 1, a film teaching assembly of an automobile distributor; Experiment 2, a film teaching assembly of a servomotor.

Experiment 1 showed superiority, both in practice and final test, of repeating a unit until learned. Experiment 2 showed advantage for repetitive group only in second practice, and the advantage was lost in the final test. Authors attribute different results partly to difference in task, suggesting that a task which is intrinsically well-organized and which is presented in this well-organized pattern is probably learned readily in a consecutive manner with little intra-serial interference between parts of task.

MARGOLIUS, GARRY J., SHEFFIELD, FRED D., and MACCOBY, NATHAN. "Timing of Demonstration and Overt Practice as a Function of Task Organization." In LUMSDAINE, ed., 1961, p. 101–05.

Problem: Whether practice is more efficient at the end of "natural units" of a task, or elsewhere.

Sample: 20 university undergraduates.

Program: Film intended to teach assembly of a servomotor.

No difference was found in amount of learning resulting from practice interspersed at end of each subassembly in the task, and practice interspersed without regard to these natural division points. The authors suggest that the natural unit organization procedure might prove to be effective in other less well organized tasks.

MCDONALD, FREDERICK J., and ALLEN, DWIGHT. "An Investigation of Presentation Response, and Correction Factors in Programed Instruction." *Journal of Educational Research*, 55 (1962), p. 502–07.

Problem: Effect of different arrangements of programed material.

Sample: 122 high school students.

Program: Teaching an unfamiliar game.

Program forms were these: (*a*) Information, example, response from learner, correct response, explanation for correct response; (*b*) same, except for absence of explanation; (*c*) information and example only; (*d*) same as (*a*), except for absence of example; (*e*) information only. No significant differences were found in learning from these forms. Authors suggest that result may be in part a function of length.

MCGUIRE, WILLIAM J. "Audience Participation and Audiovisual Instruction: Overt-Covert Responding and Rate of Presentation." In LUMSDAINE, ed., 1961, p. 417–26.

Problem: Overt vs. covert vs. no participation, at different rates of presentation.

Sample: 48 subjects, males, inmates of correctional institution, aged 17 to 26.

Program: Slides intended to teach names of nine mechanical parts.

Subject was asked to participate overtly, covertly, or not at all after each teaching unit. Overt participation resulted in more learning than no participation at a slow rate of presentation, less at a fast rate. Covert participation was superior to no participation at both slow and fast rates, and superior to overt at the fast rate.

McGuire, William J. "Effects of Serial Position and Proximity to 'Reward' Within a Demonstration Film." In Lumsdaine, ed., 1961, p. 209–16.

Problem: Usefulness of presenting vicarious "rewards" during learning task (e.g., rewarding an actor for carrying out task properly), and effect of serial practice during learning.

Sample: 32 male college undergraduates.

Program: 10-minute film teaching a motor skill.

Instructions given early in film were remembered best; those toward end, next best; and those in the middle, least well. Vicarious rewards shown in the film also benefited the learning of items immediately before or after the reward.

McNeil, John D. "Programed Instruction as a Research Tool in Reading: An Annotated Case." *Journal of Programed Instruction*, 1, 1 (1962), p. 37–42.

Problem: Effectiveness of oral responding in program designed to teach reading.

Sample: 188 kindergarten children.

Program: Elementary reading skills, 700 frames, linear.

Oral responding (saying, rather than merely looking at, the word) resulted in significantly greater learning. The oral response was particularly effective for children with lower IQ's, and resulted in more children wanting to read. For some reason, males learned significantly more than females from the programed instruction.

A more complete account of this experiment is scheduled for publication in the *British Journal of Education Psychology* under the title "Value of the Oral Response in Beginning Reading."

McNeil, John D. *Superior Reading Achievement of Boys Through Programed Instruction versus Inferior Progress Under Female Teachers*. Los Angeles: University of California (undated—probably written in 1962).

Problem: Sex differences in the reading performance of pupils taught by programed instruction, and taught by female teachers.

Subjects: 132 kindergarten children (72 boys, 60 girls), 91 of whom (49 boys, 44 girls) were later studied as first-graders with female teachers.

Program: 17 daily lessons, each consisting of approximately 35 frames, designed to teach students to recognize 40 words.

Boys scored significantly higher than girls under programed instruction, but significantly lower than girls in the classroom under

female teachers. Author concludes that female teachers may fail to adjust themselves or their school procedures as well to the traits of boys as to those of girls, and that a study of the features of auto-instruction might help in developing teaching behavior more appropriate for boys.

McNeil, John D., and Keislar, Evan R. "Questions versus Statements as Stimuli to Children's Learning." *AV Communication Review*, 10 (1962), p. 85–8.

Problem: Relative effectiveness of questions vs. statements as stimulus elements in programed instruction.

Sample: 134 first- and third-grade children.

Program: Linear program on molecular theory, 432 frames.

This experiment was intended to test the hypothesis that because covert responses would occur more readily to questions than to statements, a program made of questions would elicit more learning than one made of statements. No significant difference was found in the learning resulting from the two kinds of program.

Melaragno, Ralph J. "Effects of Negative Reinforcement in an Automated Teaching Setting." *Psychological Reports*, 7 (1960) p. 381–84.

Problem: Effect of negative reinforcement on learning from programed instruction.

Sample: 28 junior college students.

Program: Translation of logical symbols, 50 multiple-choice items.

Subjects were divided into three groups, one having an all-positive reinforcement sequence, one a spaced negative-reinforcement sequence, and one a massed negative-reinforcement sequence. For negative reinforcement, five ambiguous items were inserted in program. These items had no correct answer. On post-test, found that massed negative reinforcement depressed the learning from the program, but that spaced negative reinforcement did not depress learning significantly below all-positive reinforcement.

Meyer, Susan R. "Report on the Initial Test of a Junior High School Vocabulary Program" in *Teaching Machines and Programed Learning*, ed. by Lumsdaine and Glaser. Washington, D.C.: Department of Audiovisual Instruction, National Education Association, 1960, p. 229–46.

Problem: Effectiveness of immediate knowledge of results; and of correcting errors.

Sample: 55 eighth-graders in 4 groups.
Program: English vocabulary, linear program, 580 frames.

Groups that received immediate confirmation of results learned a greater amount (.03) than group that received knowledge of results later, when teacher had graded the answers. Significant differences in learning were not found to result from students' reviewing the program and correcting their errors. Greatest predictor of score on post-test was error score on program items.

MICHAEL, DONALD N., and MACCOBY, NATHAN N. "Factors Influencing Verbal Learning From Films Under Conditions of Audience Participation." *Journal of Experimental Psychology*, 46 (1953), p. 411–18.

Problem: The relative contributions of practice and motivation to the audience participation effect under varying conditions.
Sample: 1,029 high school students in 12 groups, 4 of them controls.
Program: Film on civil defense against atomic attack, with questions inserted at regular intervals to provide for viewer response.

Half of the test items on the film were practiced at breaks, and all items were tested at the end. All audience participation gains were in the practiced items; none were on the nonpracticed items. This gain in practiced items held for all conditions of practice—overt-covert, with or without provision of Knowledge of the Correct Response (KCR). However, KCR was far superior to no-KCR under all conditions of practice. The investigators also attempted, by announcing to half the classes that they would be tested, to study the relative contributions of varying levels of external motivation to the beneficial effect of audience participation. The attempt failed, probably because motivation was already extremely high.

MOORE, J. WILLIAM, and SMITH, WENDELL I. "A Comparison of Several Types of 'Immediate Reinforcement.'" Preliminary report of study supported by Air Force Office of Scientific Research. In *Programed Learning*, ed. by Smith and Moore. New York: Van Nostrand, 1962, p. 192–201.

Problem: Effect of different kinds and amounts of reinforcement on learning.
Sample: 220 university students in 10 groups.
Program: 1,152 frames of Skinner-Holland linear program.

Different experimental groups were given (a) no knowledge of results, (b) correct answer after each response, (c) a flashing light

after the correct response, (d) one penny for each correct response, (e) a program with response blanks filled in in advance. No significant differences in learning were found among the groups.

MOORE, J. WILLIAM, and SMITH, WENDELL I. "Knowledge of Results in Self-Teaching Spelling." *Psychological Reports*, 9 (1961), p. 717–26.

Problem: Effect of knowledge of results and controlled pacing on learning from programed instruction.
Sample: 62 and 35 sixth-graders, respectively, in the two studies.
Program: Linear program on spelling, 846 frames.

Found no significant difference between criterion learning of (a) students who received knowledge of results and those who did not, (b) students who worked at their own speed and those who worked on a specified number of units each week.

NAUMANN, THEODOR, F. "A Laboratory Experience in Programed Learning for Students in Educational Psychology." *Journal of Programed Instruction*, 1, 1 (1962), p. 9–18.

Problem: Student reactions to and performance on program.
Sample: 44 college students in educational psychology.
Program: First 16 sets of Holland-Skinner psychology (linear) program.

Students worked through first 16 sets of program, reacted in a generally favorable way—more favorably than the Oberlin students reported by Van Atta, slightly less favorably than the Harvard students reported by Holland. Attitude data for these three studies are presented in this article. Author concluded that program was usable by students in educational psychology, but that some sets needed revision designed to lower the error rate, and also that previously published time data on the program were lower than the experience here reported.

OAKES, WILLIAM F. "Use of Teaching Machines as a Study Aid in an Introductory Psychology Course." *Psychological Reports*, 7 (1960), p. 297–303.

Problem: Effect on test performance of availability vs. unavailability of teaching machines for use as a study aid.
Sample: 72 students in university evening classes.
Program: Psychology, linear program, 450 frames.

All students were from the same class, and *all* were taught by conventional methods. Half were permitted to use the machines to study on their own time. No significant differences were found between

students who did, and those who did not, have access to programed materials.

PETERSON, J. C. "The Value of Guidance in Reading for Information." *Transactions of the Kansas Academy of Science*, 34 (1931), p. 291–96.
Problem: Effectiveness of programed instruction.
Sample: Information on sample not avilable.
Program: Psychology.

A small multiple-choice drill machine giving immediate knowledge of results resulted in significantly more gain in knowledge by a psychology class that used the machine than one which did not use it.

POPHAM, W. JAMES. *The Influence of Novelty Effect upon Teaching Machine Learning*. San Francisco: The San Francisco State College, 1962 (mimeo).
Problem: Extent of novelty effect resulting from a student's first use of programed instruction.
Sample: Groups of 11 and 12 sixth-grade students matched according to scores on California Test of Mental Maturity and SRA Arithmetic Test.
Program: Plane geometry, 1,000 frames.

One group used programed instruction first semester; both groups used it second semester. There was no discernible novelty effect in the second semester, inasmuch as no significant difference was found between the groups.

PORTER, DOUGLAS. *An Application of Reinforcement Principles to Classroom Teaching*. Cambridge, Mass.: Laboratory for Research in Instruction, Graduate School of Education, Harvard University, 1961 (mimeo).
Problem: Teaching machine vs. conventional classroom instruction.
Sample: 45 second-grade, 63 fourth-grade, and 37 sixth-grade students.
Program: Linear program on spelling.

Students learned as much from teaching machine as from conventional teaching, and in about one-third the time. Greatest gains in spelling were achieved by students in the lower-IQ half of the teaching machine groups, the least gains by individuals in the comparable groups taught by the teacher. In upper-IQ groups the machine produced a nonsignificant advantage. Investigator points out that the smaller gains in higher-IQ groups were primarily due to poor matching

between subject matter of program and subject matter level of students, and that IQ correlations with spelling gains were heightened by teaching machine treatment when subject matter matching was good. A drop in achievement was observed over the entire period of the experiment, but it was no greater on the teaching machine than on the teacher-taught lessons. Attempts to teach and measure transfer of training to unfamiliar words, in fourth- and sixth-grade students, produced negative results, which investigator attributes to content of teaching rather than to the use of programed instruction.

PRESSEY, SIDNEY L. *A Puncture of the Huge "Programing" Boom?* Tucson, Ariz. (undated—probably written in 1963) (ditto).

Problem: Comparison of learning from linear program, prose text, and text plus Pressey questions.

Sample: 3 classes (about 120—exact number not stated) of university students in beginning course in education.

Program: First 54 frames of Skinner-Holland program on psychology.

The first section of the Holland-Skinner program was rewritten in six prose paragraphs. Classes of the beginning course in education were given, respectively, (a) the 54 frames of the program, (b) the 6 prose paragraphs carrying the same content, (c) the 6 paragraphs plus 6 objective Pressey-type questions. A fourth class served as a control group. All the classes were then tested on two summary questions. All the experimental groups did significantly better than the control group. Groups (a) and (b) did equally well on the test, but group (b) required about one-tenth the time to work through the learning materials. Group (c) did better than the others but not significantly so.

PRESSEY, S. L. "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Subjective Tests and Concomitant Self-Instruction." *Journal of Psychology*, 29 (1950), p. 417–47.

Problem: Validation of punchboard as testing and teaching tool.

Sample: Various numbers of university students.

Programs: Tests on Russian vocabulary, English vocabulary, and psychology.

Conclusions: (a) Punchboard discriminates better among students than does regular test; range of scores is significantly greater; (b) with students who use punchboard, errors drop significantly more on retest; (c) same trend obtains when retest is of recall rather than

recognition type, but results are not so clear-cut; (d) when questions from first test are repeated in later test, punchboard users do better with them than students who took ordinary test first; (e) there is some evidence that punchboard makes for more gains in learning with brief practice tests; (f) significant gains result when punchboard is used as major part of course; these gains are slightly, but apparently not significantly, greater than those resulting from similar use of regular tests.

PRICE, JAMES E. *A Comparison of Automated Teaching Programs with Conventional Teaching Methods as Applied to Mentally Retarded Students.* Tuscaloosa, Ala.: The Partlow State School and Hospital, 1962 (mimeo).

Problem: Constructed response vs. multiple-choice response vs. conventional classroom teaching for mentally retarded students.

Sample: 36 mental retardates, divided into 3 groups which were equated for chronological and mental ages, and arithmetic achievement level.

Program: Addition and subtraction; 3,635 frames; linear.

One group received conventional classroom instruction; the second, a program that required constructed responses; the third, a program requiring multiple-choice responses. All groups showed significant improvement during the addition part of the program, and there was no advantage for any one group. In the subtraction part of the program, the group using the multiple-choice program made significant improvement, whereas the other two groups did not. The conventionally taught group spent 2 semesters, or 130 class periods, studying arithmetic, while the pupils in the teaching machine group averaged 86 class periods.

REED, JERRY E., and HAYMAN, JOHN L., JR. "An Experiment Involving Use of English 2600, An Automated Instruction Text." *Journal of Educational Research*, 55 (1962), p. 476–84.

Problem: Program vs. conventional teaching.

Sample: 250 tenth-grade students.

Program: English grammar and usage; linear program; approximately 2600 frames.

No significant differences were found overall between experimental and control groups, but high-ability students did significantly better with the program than with conventional teaching whereas low-ability students did significantly better with conventional teaching. Considerable differences were found in time required by students of different ability levels to finish program.

RIGNEY, JOSEPH W., and BUDNOFF, IRVING J. *The Relative Efficiency of Different Combinations of Prompting and Confirmation for Learning a Boolean Algebra Program.* Prepared for Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, Los Angeles: Department of Psychology, University of Southern California, 1962 (mimeo).

Problem: Relative effect of prompting (P), confirmation (C), and three combinations (PP, PC, and CC) on learning from a program.

Sample: 105 high school students.

Program: 107 frames on Boolean algebra.

It was hypothesized that (1) the PC combination would be superior to the CC combination for acquisition of associations, and to the PP combination for retention of associations; (2) the PC and CC combinations would produce higher transfer scores than the PP combination; and (3) the brightest students would be less affected by these combinations of experimental variables than the other students.

None of these hypotheses was confirmed by the main effects of the experiment. Tests of the simple effects revealed evidence contradictory to hypothesis (1): treatment CC resulted in lower error scores on acquisition tests than did PC for upper and lower intelligence groups, with the middle group showing the reserve. All these simple effects decreased over retention intervals. Acquisition scores of the brightest students (hypothesis 3) were differentially affected by the experimental treatments.

ROE, ARNOLD. "A Comparison of Branching Methods for Programed Learning." *Journal of Educational Research*, 55 (1962), 407–16.

Problem: Simple branching methods vs. linear programing.

Sample: 189 freshman engineering students.

Program: Introduction to certain probability concepts, linear program, 107 frames.

Significant learning took place from program, but there were no significant differences between any of the simple branching methods and the linear program when measured by amount of learning. Forward branching saves time over backward branching. Logical sequence of program results in more learning than random sequence.

ROE, ARNOLD; MASSEY, MILDRED; WELTMAN, GERSHON; and LEEDS, DAVID. *Automated Teaching Methods Using Linear Programs.* Los Angeles: UCLA Department of Engineering, 1960. 57 pp. Same data also in Roe, Arnold, Automated teaching methods using linear programs. *Journal of Psychology*, 40 (1962), 198–201.

Problem: Constructed vs. selected response, overt vs. covert response, use of teaching machines in private and in group, programed vs. nonprogramed lectures.
Sample: 186 freshman engineering students.
Program: Elementary probability, 230 frames, linear.

No significant difference was found between student performance on any of the programed methods—machine, lecture, or book; and all the programed methods were significantly better than the nonprogramed lecture. Learning time, however, is significantly more for the machine methods, less for the lecture.

ROE, KIKI VLACHOULI. *Scrambled vs. Ordered Sequence in Auto-Instructional Programs.* Report 48. Los Angeles: Department of Engineering, University of California, 1961. 10 pp. (mimeo). Same data also in Roe, K. V., Case, H. W., Roe, A. "Scrambled vs. ordered sequence in auto-instructional programs." *Journal of Educational Psychology*, 53 (1962), 101–04.

Problem: Effect on learning of random vs. logically sequential ordering of items.
Sample: 36 university freshmen, in two groups.
Program: 71 frames from a program on elementary probability.

For a short program (equivalent to a 1-hour lecture), the mean criterion test score of students using the random sequence was not significantly different from that of students using the ordered sequence.

ROSHAL, SOL M. "Film-Mediated Learning with Varying Presentation of the Task: Viewing Angle, Portrayal of Demonstration, Motion, and Student Participation." In LUMSDAINE, ed., 1961, 155–75.

Problem: Effectiveness of presenting cues during learning that correspond as closely as possible to cues to be used in later performance.
Sample: 4,200 naval recruits.
Program: Eight different versions of a film designed to teach how to tie certain knots.

Results indicate that learning is more efficient if the presentation approaches a representation of the learner himself performing the desired task. This includes "subjective" camera angle and accurate representation, consecutively, of the motions and changes involved in the act.

ROTHKOPF, ERNST Z. "Automated Teaching Devices and a Comparison of Two Variations of the Method of Adjusted Learning." *Psychological Reports*, 8 (1961), 163–69.

Problem: Effect of different amounts of review in programed instruction.

Sample: 100 airmen, divided into 2 experimental and 2 control groups.

Program: Eight paired-associate items.

No significant differences were found in learning between two versions of the method of adjusted learning. Both versions resulted in subjects' responding to each item correctly exactly twice before being tested. In one version an instructive item was dropped out after *one* correct response, and when all items had met this criterion the entire procedure was repeated. In the second version each item was dropped out only after *two* correct responses.

ROTHKOPF, ERNST Z. "Programed Self-Instructional Booklets, Mnemonic Phrases, and Unguided Study in the Acquisition of Equivalences." *Journal of Programed Instruction*, 1, 1 (1962), 19–28.

Problem: Relative effectiveness of pairs, mnemonic phrases, and linear programs in learning and retention of simple equivalences.

Sample: College students in groups of 32, 31, and 33.

Program: Resistor color code.

Three treatments were used in an effort to teach the color code used for identifying electronic resistors: (a) A list of simple color-number equivalents, (b) a similar list with the addition of a short mnemonic phrase for each pair, and (c) a 60-item booklet in linear program form, also making use of the mnemonic phrases. No difference was found as a result of the experimental treatments when subjects were tested within 2 days after the exposure. On retention tests, administered up to 120 days after the exposure, however, the programed booklet resulted in better performance than the simple list of pairs. During the first 60 days, the mnemonic list produced about the same result as the simple list, but in later tests it was significantly superior to the simple list and about as effective as the program. Subjects who reported that they *used* mnemonic aids retained substantially more than others.

ROTHKOPF, ERNST Z. "Some Observations on Predicting Instructional Effectiveness by Simple Inspection." To be published in *Journal of Programed Instruction*, 1963.

733-269 O—64——7

Problem: Ability to predict effectiveness of self-instructional program by reading it.
Sample: 12 high school teachers or principals.
Program: Seven versions of a program designed to teach relationships between pairs of nouns in fictional anthropological subject matter.

Actual effectiveness of the seven treatments had been previously determined experimentally (Rothkopf and Coke, 1963). The rank correlations between actual effectiveness and the prediction of the 12 educators was −.75.

ROTHKOPF, ERNST Z., and COKE, ESTHER U. "Repetition Interval and Rehearsal Method in Learning Equivalences from Written Sentences." To be published in *Journal of Verbal Learning and Verbal Behavior*, 1963.

Problem: Relative effectiveness of three kinds of rehearsal modes, and of different numbers of items interposed between initial presentation of an item and its subsequent rehearsal.
Sample: 544 clerical personnel.
Program: Different forms of a program designed to teach relationship between pairs of nouns in fictional anthropological subject matter.

The three rehearsal modes were (1) stimulus anticipation—anticipating during practice the term which would be used as a stimulus in a subsequent test trial, (2) response anticipation—anticipating during practice the term which would be the required response in a subsequent test trial, and (3) simultaneous presentation of the two terms side by side, with mere instructions to the subjects to read them. Rehearsal took place either immediately after the initial presentation of an item (zero interval), or after a randomly determined number of other items had been interposed. It was found that zero interval was appreciably less effective than randomly determined interval. Rehearsal through active anticipation with zero repetition interval was the worst practice condition examined. Simultaneous presentation used throughout the experiment proved to be as good as or better than active anticipation, although active anticipation is a more effective mode of rehearsal than simultaneous presentation for subjects who use *both* modes in practice. Response anticipation is more effective than stimulus anticipation when subjects operate under only one rehearsal mode.

SCHAEFER, HALMUTH H. "E. A. Poe as a Reinforcer." *Psychological Reports*, 8, 1961, 398.

See GLASER AND TABER, 1961.

SEVERIN, D. G. "Appraisal of Special Tests and Procedures Used with Self-Scoring Instructional Testing Devices." *Abstracts of Doctoral Dissertations*, 66 (1955), 323–30.

Problem: Effectiveness of punchboard in contributing to learning.
Sample: 2,000 college students.
Program: "A number of different topics."

Advantages of punchboard as testing device are pointed out. Trials of it with college students on a number of different topics proved that it also contributed to learning. In comparison of punchboard study with ordinary text study, punchboard came out ahead more often, but differences were small. Concluded that if learning rather than testing is the object, two-choice multiple-choice items are sufficient.

SHAY, CARLETON B. "Relationship of Intelligence to Step Size on a Teaching Machine Program." *Journal of Educational Psychology*, 52 (1961), 98–103.

Problem: Relation of step size and IQ to amount of learning from programed instruction.
Sample: 90 fourth-grade pupils.
Program: Use of Roman numerals; programs of 103, 150, 199 steps in decreasing step size.

Inferring step size from number of errors made on program, Shay found no relationship between intelligence and step size in amount of pupil learning. Step size was measured in terms of "difficulty of giving the correct answer," i.e., number of errors made on given items.

SHEFFIELD, FRED D.; MARGOLIUS, GARRY J.; and HOEHN, ARTHUR J. "Experiments on Perceptual Mediation in the Learning of Organizable Sequences." In LUMSDAINE, ed., 1961, 107–16.

Problem: Advantages, if any, of short pictured review of mechanical assembly at end of each sub-unit of assembly task.
Sample: 60 university students.
Program: Film teaching assembly of a servomotor.

Pictured review of this kind as compared to no review, was found to enhance significantly the effectiveness of learning.

SHETTEL, HARRIS H., and LINDLEY, RICHARD H. *An Experimental Comparison of Two Types of Self-Instructional Programs for a SAGE System Paired-Associate Task*. Pittsburgh, Pa.: American Institute for Research, 1961. 21 p.

Problem: Effectiveness of different programed methods vs. different uses of flashcards in teaching the phonetic alphabet.
Sample: 60 college students, in 6 groups.

Program and conditions: (*a*) 90-frame constructed-response program, overt response, training time not limited; (*b*) 30 frames selected from longer program, same conditions; (*c*) same as *b*, but with covert responding; (*d*) flashcards, covert responding, training time limited; (*e*) same as *d*, but training time unlimited; (*f*) same as *d*, but with "drop-out" feature—all designed to teach the military phonetic alphabet.

Least number of errors on two post-tests were made by subjects who studied the long program (*a*) or any of the flashcards (*d*, *e*, *f*). Since the flashcard groups led to high performance in about 14 minutes as compared to 40 for the long program, author concludes that in this type of paired-associate task there is no advantage in programed self-instructional methods.

SHUNNY, JOHN. *A Report on the Naval Reserve Russian Course.* Albuquerque, N. Mex.: Naval Training Center, 1961. 9 p.
Problem: Effectiveness of a vocabulary program, plus grammar and tape recording study.
Sample: 11 Naval Reserve officers.
Program: Russian linear program.

Officers devoted 70 hours during 10 days to full-time study of the Russian language. As each one finished the program, he was put on the Semeonoff text half-time and on the spoken Russian recordings half-time. Criterion was to be able to write simple Russian sentences and translate Russian passages into English. The supervisor sums up results by saying that these students learned about as much Russian in these 70 hours as they would have learned in about one and one-half semesters of a college-level course.

SIDOWSKI, JOSEPH B., KOPSTEIN, FELIX F.; and SHILLESTEAD, ISABEL J. "Prompting and Confirmation Variables in Verbal Learning." *Psychological Reports*, 8 (1961). 401–06.
Problem: Prompting vs. confirmation; overt vs. covert response.
Sample: 120 college students, in 6 groups.
Program: 15 Russian-English paired vocabulary items.

Taking into account both time and amount of learning, covert response proved more efficient than overt; and prompting (pairs presented together) more efficient than confirmation (words presented consecutively).

SILBERMAN, HARRY; COULSON, JOHN; GUNN, EVELYN; and MELARAGNO, RALPH. *Development and Evaluation of Self-Instructional Materials for Underachieving and Overachieving Students.*

Santa Monica, Calif.: System Development Corporation, 1962.
Problem: Rote form of program (which moved step by step from beginning to end of theorem) vs. conceptual form (which started from what had to be proved and worked backward) for students of different achievement groups.
Sample: 258 high school students in second-semester geometry.
Program: Linear program on plane geometry (rote version 526 items; conceptual version, 550).

Students were divided into three achievement-level groups on basis of standard tests. Each achievement group was divided into two treatment groups, one of which received the rote program; the other, the conceptual form of the program. The high-achievement group did better than the medium-achievement group, which did better than the low-achievement-group, with aptitude level and pretest scores controlled. There was no significant difference related to form of program. Significant interactions were obtained when data from each school participating in the study were separately analyzed. In one school, the overachievers did better on the rote form while the other groups did better on the conceptual form. In a second school, overachievers did better on the conceptual form while normal-achievers and underachievers did better on the rote form. A third school showed a significant difference in favor of the rote method on the application portion of the criterion test.

SILBERMAN, HARRY F.; MELARAGNO, RALPH J.; and COULSON, JOHN E. "Confirmation and Prompting with Connected Discourse Material." *Psychological Reports*, 9 (1961), 235-38.
Problem: Prompting vs. confirmation.
Sample: 44 junior college students.
Program: Linear program on simple and compound statements, connectives, arguments (61 frames).

Compared achievement of three groups receiving three different treatments: (1) Items with confirmation; fixed sequence, (2) items with prompting, but no confirmation; fixed sequence, (3) paragraph statements without confirmation; student-controlled sequence. No significant differences were found in amount of learning, but no-response conditions (2 and 3) took less time.

SILBERMAN, HARRY F.; MELARAGNO, RALPH J.; COULSON, JOHN E.; and ESTAVAN, DONALD. "Fixed Sequence vs. Branching Auto-Instructional Methods." *Journal of Educational Psychology*, 52 (1961), 166-72.

Problem: Effectiveness of fixed sequence vs. voluntary review and voluntary sequencing; fixed sequence vs. branching.
Sample: Experiment I, 3 groups of 17 high school students each; experiment II, 36 high school students.
Program: Experiment I, 61 multiple-choice items on logic; experiment II, 411 multiple-choice items on logic, using a digital computer as a teaching machine.

Experiment I compared three conditions: (1) fixed sequence with confirmation procedure, (2) same items as in (1), but with subjects being allowed to review at will, backing up one card at a time, (3) items in statement form, with subjects being allowed to choose their own way of studying the cards. Treatment (3) proved significantly better than (1). Between (1) and (2) there was no significant difference.

Experiment II compared (1) branching determined by errors and (2) fixed sequence. No significant differences were found.

SILVERMAN, ROBERT E. *The Use of Context Cues in Teaching Machines.* Port Washington, N.Y.: U.S. Naval Training Device Center, June 1961. 69 pp.
Problem: Effectiveness of different response modes.
Sample: Information on sample not available.
Program: Information on program not available.

Compared groups using programs with (1) written response, (2) written and vocalized response, (3) vocalized response, (4) covert responses. Found greatest learning for the covert response group.

SILVERMAN, ROBERT E., and ALTER, MILLICENT. "Note on the Response in Teaching Machine Programs." *Psychological Reports*, 7 (1960), 496.
Problem: Overt vs. covert response.
Sample: 60 college students.
Program: Basic electricity; linear; 87 frames.

Students who simply read items learned significantly more than students who read items and responded actively. Preliminary report of same data described later in SILVERMAN and ALTER (1961), and ALTER and SILVERMAN (1962).

SILVERMAN, ROBERT E., and ALTER, MILLICENT. *Response Mode, Pacing, and Motivational Effects in Teaching Machines.* Port Washington, N.Y.: U.S. Naval Training Device Center. Tech. Rep. 507-03, 1961. 69 p. (Five experiments were reported in this document. For experiments I, IIA, and IIB, see ALTER and SILVERMAN, 1962.)

Experiment III:

Problem: To compare three degrees of pacing: self-pacing, automatic pacing that allows ample time, and automatic pacing that allows excessive time.

Sample: 45 college students.

Program: 87 frames on basic electricity; linear; constructed response.

Results: No significant differences in achievement test, but group that was allowed excessive time reported boredom.

Experiment IV:

Problem: To investigate novelty effect of teaching machines.

Sample: 30 college students and 30 high school students.

Programs:

90 frames on psychology; 90 frames on binary numbers. Both were linear, constructed-response programs.

One group used a programed text; a second used a crude spool-type device; a third used an elaborate electromechanical device. Each subject took one of the programs and after an average interval of 1 week, returned to take the other program on the same device.

Results: No significant achievement differences as a function of the device used or as a function of repeated exposure.

SMITH, EDGAR A., and QUACKENBUSH, JACK. "Devereux Teaching Aids Employed in Presenting Elementary Mathematics in a Special Education Setting." *Psychological Reports*, 7 (1960), 333-36.

Problem: Effectiveness of automated instruction with slow learners.

Sample: 23 slow learners.

Program: Multiple-choice items on arithmetic: 1,440 items initially, later 9,216.

Program was in form of a workbook resting on a machine with two rows of pushbuttons. When subject pushed button adjacent to what he considered correct response, buzzer sounded if the response was correct. Instruction continued for a year, resulting in average gain of 0.51 grade levels as measured by California achievement test. Corresponding gains by similar students in previous year had been 0.19 grade levels in arithmetic, and an average of 0.25 of a grade in other academic areas.

See also SMITH, EDGAR A., MCKEAN, MORGAN W., and KUNSMAN, BARBARA, "Specialized Use of Teaching Machines in the Classroom," Devereux Schools Forum, 1, 2 (1961). This report on three interrelated studies using Esperanto as subject matter is obtainable from

MORGAN W. MCKEAN, The Devereux Foundation, Devon, Pa. Primary attention is on retention and IQ.

SMITH, H. NORMAN. "The Teaching of Elementary Statistics by the Conventional Classroom Method versus the Method of Programed Instruction." *Journal of Educational Research*, 55 (1962), 417–20.

Problem: Scrambled book program vs. conventional teaching.
Sample: 128 freshman cadets, U.S. Air Force Academy.
Program: Scrambled book treatment of statistics.

No significant difference in performance between experimental and control group, although the program required less time. Test of interest showed no significant difference between the two groups.

SMITH, LEONE M. *Programed Learning in Elementary School: An Experimental Study of Relationships Between Mental Abilities and Performance.* Urbana, Ill.: University of Illinois, 1962 (mimeo).

Problem: Relationship between different levels of ability in performance under conventional methods of instruction and self-instructional methods.
Sample: 8 treatment groups totaling 195 fifth-grade students.
Program: Same subject matter as that presented in an arithmetic textbook used by teacher-taught groups; 757 frames.

Four learning conditions were identified, and these provided the set of instructional procedures. The conventional classroom teaching methods comprised one learning condition; sets of programed self-instruction material represented three other learning conditions. A comparison of the two strategies of instruction showed (1) different patterns of correlations between ability and post-test scores of learning at different stages of learning and according to the kind of content, (2) both strategies produced correlations between the same number of ability test scores and post-test scores of learning, (3) operationally different instructional strategies did not produce positive correlations of significant magnitude when the same ability was correlated with post-test scores of learning.*

SMITH, WENDELL I., and MOORE, J. WILLIAM. *Programed Materials in Mathematics for Superior Students in Rural Schools.* Lewisburg, Pa.: Bucknell University (undated—probably written in 1962). 91 p.

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

Experiment I:

Problem: Effectiveness of programed instruction, by machine or programed textbook, plus a weekly seminar vs. textbook plus weekly seminar.

Sample: 100 high school juniors and seniors.

Program: Sets, relations, and functions; linear; 1,036 frames.

No significant differences in achievement related to use of teaching machine, programed text, or conventional textbook.

Experiment II:

Problem: Effectiveness of teaching machine program, programed textbook, and conventional textbook *without* supplementary classwork.

Sample: Same as in experiment I.

Program: Groups and fields; linear; 1,426 frames.

Machine group scored significantly higher than conventional text group, both on immediate and delayed post-test.

Experiment III:

Problem: Effectiveness of teaching machine program vs. programed textbook vs. reading and copying correct responses.

Sample: Same as in experiments I and II.

Program: Introduction to probability; linear; 970 frames.

No significant differences were found.

Reporting on studies of student attitudes during the experimental periods, the authors say that "students believe they learn as well with programed material as they do in a 'traditional' course; they like programed material, particularly as a supplement to other instructional materials and teacher-centered instruction; they prefer a programed textbook to a machine. They become bored with sustained exposure to learning programs. Let no one forget that students also become bored with sustained exposure to lectures, to films, to textbooks."

SMITH, WENDELL, and MOORE, J. WILLIAM. *Size-of-step and achievement in Programed Spelling.* Lewisburg, Pa.: Bucknell University, 1961. 15 p.

Problem: Effect of different sizes of step.

Sample: 96 fifth-grade students.

Program: Linear program on spelling.

Found no difference in weekly test grades or difference between gains from pre- to post-test, associated with programs requiring about 1,128, 830, and 546 steps, respectively to cover same ground—and no difference when pictorial cues were used in any of the three programs. Group with highest error rate had lowest mean gain.

SMITH, WENDELL I., and MOORE, J. WILLIAM. "Size of Step and Cueing." *Psychological Reports*, 10 (1962), 287–94. (Also reprinted in SMITH and MOORE. *Programed Learning*. New York: Van Nostrand, 1962, p. 202–06.)

Problem: Learning as affected by interaction of size of step and cueing.

Sample: 95 fifth-grade children, in 6 experimental conditions.

Program: Spelling of 166 words, in programed versions of 1,128, 830, and 546 frames each.

Three versions of program were used, each one with and without pictorial cues. Small-step version of program had 4 to 9 frames per word, medium-step version had 3 to 7, and large-step version had 3 to 6. No significant differences were found between gains from any of the treatments.

SORENSEN, PHILIP H.; ANDERSON, JEANNE; and ROSS, SHEILA. *An Experimental Test of Programed Instruction in Postgraduate Dental Education*. Menlo Park, Calif.: Stanford Research Institute, 1963.

Problem: Relative effectiveness, utility, and acceptability of two types of self-instructional materials for postgraduate professional dentists.

Sample: 57 practicing dentists.

Programs: Two topics—"periodontal membrane" and "cementum"—each treated in two styles of presentation. Programs were brief, branching-style in scrambled book form in the body of the program and linear-style with constructed responses in review sections. Competing modes were straight expository narratives developed from the original programs with questions from the programs converted to declarative statements. Narrative versions required no overt response and provided no explicit feedback. Topic content of programs and narratives identical.

Knowledge tests showed no difference in gains between program groups and narrative groups for either "periodontal membrane" content or "cementum" content. Both program and narrative groups showed significant gains between preinstruction tests and post-instruction tests. Narrative groups required about 35 percent less time than program groups to complete instruction in both topics. Narrative groups were more likely to think they had learned "a great deal" than were program groups. Findings suggest that learner characteristics (e.g., prior knowledge of subject matter and interest in subject matter) and instructional material characteristics (e.g., systemization

of content and mode of response required) interact to influence instructional effectiveness. For example, the requirement for active or overt responses was hypothesized as (1) helpful to persons with substantial interest in the subject but little prior familiarity with it but (2) detrimental (e.g., actively irritating) to persons with some prior knowledge of the subject but little interest in it.

STOLUROW, LAWRENCE M. *A Comparative Study of Methods of Programing Materials for Efficient Learning in Self-Instructional Devices.* Urbana, Ill.: University of Illinois, 1961. 14 p. (mimeo).

Problem: Prompting vs. confirmation.
Sample: Retarded children; number not available.
Program: Vocabulary.

With a shorter practice time, prompting was superior to confirmation; with a longer practice period, confirmation made for more retention. Prompting required fewer trials to achieve desired learning.

STOLUROW, L. M., and LIPPERT, H. *Prompting, Confirmation and Vanishing in the Teaching of a Sight Vocabulary.* University of Illinois, Training Research Laboratory, (undated—probably written in 1962) (mimeo).

Problem: Effect of varying numbers of prompting trials followed by varying numbers of confirmation trials on the learning and retention of sight words.
Sample: 25 educable mentally handicapped public-school children with a Binet mental-age range of 4 years, 3 months, to 6 years, 5 months.
Program: Vocabulary program consisting of 25 nouns, 5 in each of 5 concepts (clothes, animals, toys, foods, and furniture); presented in combinations of 1, 3, 6, 9, 12 criterion trials where prompting was the first sequence and confirmation was the second. 900 frames.

Median number of errors and trials to criterion was zero under each sequence, indicating this was a program with a low error rate. Retention was maintained at about the same level throughout the retention testing period of 1, 7, and 30 days after learning. The effect of vanishing the prompt by introducing the confirmation sequence after various numbers of prompting trials was complex and not entirely consistent. The point biserial correlations between mental age and recall at 1 day and at 30 days showed an increase as the total number of criterion trials increased, indicating that mental age was accounting for an appreciable proportion of the variance in the recall scores.

The preschool mental-age group, i.e., children with a mental age below 5 years, 6 months, recalled a median of 3 out of 25 words taught, whereas the kindergarten mental-age group, i.e., children with a mental age above 5 years, 6 months, recalled a median of 11 of the 25 words.*

STOLUROW, L. M., and WALKER, C. C. "A Comparison of Overt and Covert Response in Programed Learning." *Journal of Educational Research*, 55 (1962), p. 421–32.

Problem: Effectiveness of overt vs. covert response.
Sample: 51 university students in beginning psychology.
Program: Descriptive statistics.

No significant difference was found in learning from writing responses or merely thinking them. Covert responses required less time than overt.

SUPPES, PATRICK, and GINSBERG, ROSE. "Application of a Stimulus Sampling Model to Children's Concept Formation With and Without Overt Correction Response." *Journal of Experimental Psychology*, 63 (1962), p. 330–36.

Problem: Effectiveness of overt correction response.
Sample: Forty-eight 5- and 6-year-olds.
Program: Learning concept of 4-ness and 5-ness in binary number system.

Two groups were compared: (1) A group in which subjects were required to make an overt correct response after every incorrect one, (2) a group in which subjects were merely told whether their response was correct. Group (1) performed significantly better.

Two methods of analysis were used: a paired-associate analysis, in which the stimuli were treated as if they were independent items, and a pure property analysis, in which all stimuli describing a single concept were treated as if they were identical items. A simple sampling model for paired-associate learning was applied to the data, and a good fit was obtained for the paired-associate analysis. The pure property analysis was applied to group (1), and did not approach the adequacy of the paired-associate analysis.

SUPPES, PATRICK, and GINSBERG, ROSE. "Experimental Studies of Mathematical Concept Formation in Young Children." *Science Education*, 46, 3 (1962), 230–40.

Problem: Concept learning from programed materials, and application of a paired-associate learning model to the data.
Sample: Experiment I, forty-eight 5- and 6-year-olds; experi-

*This abstract was prepared by the Training Research Laboratory, University of Illinois.

ment II, 96 first-grade children; experiment III, 32 first-grade, 32 kindergarten children; experiment IV, 36 kindergarten children; experiment V, 60 kindergarten children.

Program: Sections of a program on sets and numbers, with variations in stimulus displays and response conditions.

Experiment I is presented in greater detail in Suppes and Ginsberg, 1962b. Experiments II, III, and IV involve stimulus variations; and Experiment V, response variations.

Conclusions: (1) Learning is more efficient if the child who makes an error is required to make the correct response in the presence of the stimulus to be learned. (2) Incidental learning does not appear to be an effective method of acquisition for young children. (A group of children who responded to a color discrimination gave no sign subsequently that they had learned the underlying concept.) (3) A condition which focuses the child's attention upon the stimuli to be learned enhances learning. (4) Transfer of a concept is more effective if the learning situation has required the subject to recognize the presence or absence of a concept in a number of stimulus displays than if the learning situation has only required matching from a number of possible responses. A multiple-answer situation offering three responses is more effective than one offering only two possible responses. (5) A young child's learning tends to be very specific. (In one of the experiments, prior training on one concept did not improve learning on a related concept.)

A good fit was found between obtained frequencies in the concept learning and frequencies predicted by a model that was derived from a paired-associate analysis of learning.

TABER, JULIAN I., and GLASER, ROBERT. "An Exploratory Evaluation of a Discriminative Transfer Learning Program Using Literal Prompts." *Journal of Educational Research*, 55, 9 (1962), p. 508–12.

See GLASER and TABER, 1961.

TEACHING MACHINES, INC. *Some Industrial Uses of Programed Instruction.* New York: Teaching Materials Corporation (undated—probably written in 1962). 6 p. (mimeo).

Problem: Effectiveness of programed instruction as used by industry.

Sample: Sandia sample was 208 adult employees who completed at least one course, out of 279 who had started.

Program: Sandia used electronics (n?), algebra (1,931 frames), and Russian (1,990 frames).

Generally satisfactory results are reported for programs used by

ACF Industries, but no figures. Sandia Corporation reported on programed courses offered to their employees. Percentage of completions for program-using employees was higher than that for class-taking employees in same subjects. Students of algebra programs did as well as class students of algebra. No pre- and post-test scores were reported for other subjects. It cost $57.15 per completion to teach Russian in classes, $20.19 per completion to teach it by program; $20.50 per completion to teach algebra in classes, $16.79 to teach it by programs. Costs of instructors' salaries were approximately the same for both types of teaching.

See Utter, Robert F., "Programed self-instructional materials in background courses". In Dolmath, T. B., Marting, E., and Finley, R. G., eds., *Revolution in training: programed instruction in industry.* New York: American Management Association, 1962.

VAN ATTA, LOCHE. "Behavior in Small Steps." *Contemporary Psychology*, 6, 10 (1961), p. 378–81.

Problem: Student attitudes toward programed instruction.
Sample: 90 college students at Oberlin.
Program: Holland-Skinner program on human behavior.

In a review of the Holland-Skinner program, investigator reports data on reactions of Oberlin students to the program. Although their reactions were generally favorable, the students objected to the amount of repetition and the short steps in the program.

VICORY, ARTHUR C. *Preliminary Research Report of the Corrigan Telecommunication System.* 23 p. (Place of writing or publication unidentified.) 1963. (Ditto.)

Problem: Effectiveness of overt response with immediate knowledge of results vs. response without knowledge of results vs. no overt response.
Sample: 87 college students.
Program: 45-minute presentation on the development of the number system, with 31 questions inserted for student response.

Experimental group A responded to each of the questions by pressing one of a number of buttons, and received immediate knowledge of results. Group B responded by marking on IBM answer sheets. Group C neither received questions nor made responses. Group A performed significantly better on a post-test than did either of the other groups. Another part of the experiment tested whether the program presents material comparable to the classroom course.

VICORY, ARTHUR C., and CORRIGAN, ROBERT E. *Learning Math Concepts through Required Responses and Feedback.* San Jose,

Calif.: Instructional Television, San Jose State College, 1963, 24 p. (mimeo).

Problem: Effects on learning math concepts with response feedback apparatus.

Sample: 66 college undergraduates.

Program: Development of the number systems—15 concepts.

The feedback apparatus used in this experiment is intended to be used with television. In this experiment, however, the students were given oral instructions with visual examples in a classroom situation. They listened, observed the examples, and responded to multiple-choice questions that were asked at specific portions of the programed instruction by pressing a correct key on the multiple-choice response apparatus. A green light signaled a correct response; a red light, an incorrect response. The student was instructed to press keys until the green light flashed.

Three subsamples were employed to test the learning effect of immediate feedback by the response apparatus. (1) One subsample was asked questions and was instructed to respond with apparatus during the program (feedback). (2) The second subsample was asked questions, but was instructed to respond on IBM sheets (no feedback). (3) The third subsample was given the same program without questions or response apparatus. A criterion test was given to all three subsamples 1 week later.

It was hypothesized that performance on the criterion test would be significantly better for subsample (1) than for either subsamples (2) or (3), and, further, that criterion test performance would not differ significantly between subsamples (2) and (3). Both hypotheses were supported.

WEISS, WALTER; MACCOBY, NATHAN; and SHEFFIELD, FRED D. "Combining Practice with Demonstration in Teaching Complex Sequences: Serial Learning of a Geometric-Construction Task." In LUMSDAINE, ed., 1961, p. 55–76.

Problem: Length of demonstration between practice periods.

Sample: Experiment I, 900 high school students; experiment II, 28 junior college and 37 senior college students.

Program: Film on how to construct a geometric pentagon.

Experiment I verified the assumption that short steps (short periods of demonstration followed by practice) would result in perfect trial performance by 75 percent or more of students. Experiment II compared (a) these short steps, with (b) gradually increasing steps, with (c) self-selected size of step, with (d) practice only after the entire film had been completed. Performance-rate scores (number

correct per unit of time) showed treatment (d) significantly inferior to all the others, and treatment (b) significantly superior to treatment (a). Treatment (c) turned out, by student choice, to be a series of gradually increasing steps like (b). Performance gradually deteriorated among students practicing short steps only. This method of short steps appears to give good results only during training, but not during the final test, when there is no immediate support for training. The implication is that the optimum method of utilizing overt practice may not be the best method for integrated unprompted performance of the total task.

In comparing these results with those of MARGOLIUS and SHEFFIELD (1961), it should be noted that the present experiment uses a shorter program on a different topic, and that the task did not lend itself so readily to partition.

WENDT, PAUL R. and RUST, GROSVENER. "Pictorial and Performance Frames in Branching Programed Instruction." *Journal of Educational Research*, 55 (1962), p. 430–32.

Problem: Teaching program vs. lecture, usefulness of branching and of pictorial frames.

Sample: 12 sections of freshman English at a university.

Program: Library use.

Found that students could learn as much about library use from program as from lectures. For brighter students, branching saves some time over linear program. Pictorial frames proved effective in this particular learning task.

WOLFE, MARTIN S. *Effects of Expository Instruction in Mathematics on Students Accustomed to Discovery Methods* (doctoral dissertation). Urbana, Ill.: University of Illinois, 1963 (mimeo).

Problem: Effect of change to an expository method of teaching on students accustomed to "discovery" methods of teaching.

Sample: 300 ninth- and tenth-grade students in second year of University of Illinois Committee on School Mathematics program.

Program: Linear programs on modern mathematics, one made on an expository model (in which the initial set in each teaching sequence was a verbal statement of the concept of generalization to be learned); the other, on a discovery model (in which each concept or generalization was developed inductively before being verbalized).

All students had been taught by discovery method the previous year. During the test period, half the students studied the discovery program, half the expository one. By means of mathematics achieve-

ment scores, each group was divided into low-, average-, and high-ability subgroups. No significant differences were found between treatment groups or between corresponding subgroups in different treatments. Author concludes that there is no evidence in this experiment that students taught by discovery methods early in school receive unsatisfactory preparation for later teaching in which discovery is not emphasized; and also that there is no reason to fear the combining of discovery and expository methods, for example, in a textbook.

WULFF, J. JEPSON, and KRAELING, DORIS. "Familiarization Procedures Used as Adjuncts to Assembly-Task Training with a Demonstration Film." In LUMSDAINE, ed., 1961, p. 141–53.

Problem: Whether there is an advantage familiarizing a learner with the relevant features of stimulus objects to be used in a learning task, and, in particular, whether it is more effective to do this before the entire task or just before the part of the task in which the stimulus objects are to be used.

Sample: 33 prison inmates.

Program: Film teaching assembly of automobile ignition distributor.

Familiarization was found to result in fewer selection errors, and the group given familiarization training just before the appropriate part of the task made fewer errors than the group given familiarization at the beginning.

ZUCKERMAN, CARL B.; MARSHALL, GEORGE R.; and GROESBERG, SANFORD. *Research in the Automation of Teaching.* Port Washington, N.Y.: U.S. Naval Training Device Center, February 1961. 160 p.

Problem: Effectiveness of different response modes and of ordered vs. random sequence of items.

Sample: Information on sample not available.

Program: 60 items on electricity.

No significant differences were found in gain scores between groups which had studied constructed-response, multiple-choice, and true-false programs. Likewise, no significant differences were found between groups studying an ordered and a random program.

732-239 O—64——8

Author Index to Bibliography