DOCUMENT RESUME

ED 051 676 EM 009 074

AUTHOR	Bjerstedt, Ake
TITLE	System Modification and Evaluation in Instructional Programming: The Final Phases of the Program Construction Process.
INSTITUTION	School of Education, Malmo (Sweden). Dept. of Educational and Psychological Research.
REPORT NO	R-33
PUB DATE	May 71
NOTE	130p.; Didakometry series, Vol III
EDRS PRICE	EDRS Price MF-$0.65 HC-$6.58
DESCRIPTORS	Educational Technology, *Evaluation Criteria, *Evaluation Techniques, Glossaries, *Instructional Systems, Manuals, Measurement Techniques, Motivation, *Programed Instruction, *Program Evaluation, Program Improvement, Retention, Systems Approach

ABSTRACT

In this final volume of a three-volume series on the construction of self-instructional systems, the focus is on various aspects of post-construction control and improvement. As the first step in this phase, the process of polishing the material by checking for readability, integration, retention, and motivation is described and some of the aids to be used in this checking phase are suggested. A description of empirical procedures is included to emphasize the need for systematic revision techniques. Some examples are presented of relevant variables and systematic mapping devices which aid in evaluation. Combinations of revision criteria are reviewed and a few special measures to reduce student boredom are suggested. The final task in program construction, compiling a program manual, is detailed with a catalog of the categories of information common to any program manual. A check list is provided to facilitate the assessment and comparison of existing programs. A glossary defines terms used in programed instruction and educational technology. See also volume one (EM 009 072) and volume two (EM 009 073). (JY)

ED051676

SYSTEM MODIFICATION AND EVALUATION IN INSTRUCTIONAL PROGRAMMING: THE FINAL PHASES OF THE PROGRAM CONSTRUCTION PROCESS

Åke Bjerstedt

The construction of a self-instructional system can be seen as a work process with three main phases: (1) System analysis: Preparatory work; (2) System synthesis: Construction of a preliminary system version; and (3) System modification and evaluation: Post-construction control and improvement. - While phases 1 and 2 in the program construction process were dealt with in earlier reports (Didakometry, Nos. 30 and 32), the present survey focuses upon phase 3, discussing various aspects of post-construction control and improvement (such as the polishing and checking phase, the general procedure of successive evaluations, various evaluation criteria and evaluation aids, compiling the program manual). - Terminological and bibliographical appendices are included.

CONTENTS

12. POLISHING AND CHECKING PHASE

When the various instructional units have been constructed, a renewed examination is made of the unit sequence. It is, of course, important to check that nothing important has been forgotten, and tha. the various parts of the program have been put together smoothly. But there are many other factors that the programmer should consider before starting empirical experiments with his material. In this chapter, some of these points of view will be briefly discussed, and various aids (designed to help the programmer in this checking phase) will be described.

We will first group our comments on checking around four problem areas: (1) facilitation of communication (readability), (2) integration, (3) retention, and (4) motivation. We will thereafter devote a special section to various aids for systematic mapping and checking procedures.

12.1 FACILITATION OF COMMUNICATION (READABILITY)

Those responsible for planning and writing the study material do not, perhaps, always notice the difficulties o ambiguities in their formulations. On the other hand, someone who has not been directly involved is aware almost immediately of such things in his capacity as recipient rather than dispenser of information. The final proof of the efficacy of the study program as a communication medium is, naturally, revealed in its effect on representatives of the student group. Ordinarily, however, considerable effort can be saved by carefu examination in advance.

Someone familiar with programming, but not connected with the particular project, reads it through, therefore, with the purpose of simplifying it and pinpointing obscurities. It is usually desirable that two others take part in this preliminary study. An expert on the subject makes a detailed examination of the factual material and also evaluates how well the course material is covering the ground required. The material is then read for a language control by someone who has both a good writing style and experience of the characteristic reading ability of the students in the target group. If the program contains special types of aids (pictures, sound effects, etc.), it is often necessary to call upon other experts to assure the best possible technical use of these aids. In this section we shall briefly present some ideas on one of these aspects of revision, namely, the question of readability or - in more general terms - communicability.

Among the main questions the reviser should ask himself are these: (1) Is the presentation of material as far as possible adapted to the special qualifications of the students (student-adapted communication)? - (2) Is the thought content expressed with an optimal degree of precision?- (3) Does the physical structure facilitate as much as possible a meaningful focusing of attention?

One study (Grace, 1963) indicated that programmed material differed significantly from non-programmed material in respect to several factors which are usually considered to be connected with readability. The programmed study material used, on the whole, shorter words as well as a higher percentage of the more frequently used words in the language, and new words were introduced at a slower rate. Even if the results of this particular study may not be entirely reliable (since we do not know much about the comparability of the two groups of texts used in other respects), the indications are that program writing as such tends to influence the author to a use of words better adapted to the language abilities of the student. This is quite natural, since the author, unlike the writer of textbooks (and through the presentation method itself with its constant demands for an answer), is not allowed to forget that he is speaking to a specialized audience with a claim to attention.

The demand for student-adapted communication does not mean, naturally, that a whole program can be confined within the boundaries of a minimum language area which all students at the beginning of the course can be expected to command. But it means that one must start from that level and must make sure that new terms are not introduced without an explanation understandable to the students (that is to say, an explanation which uses for its definition language which can be understood by the students) and at a pace accomodated to their abilities (i. e., not so rapidly that the students cannot, when tested, demonstrate their ability to absorb the new words). In this connection good use can be made of the original student analyses. For this purpose it may be desirable that the student analyses contain recorded samples of the normal conversation of the test group as well as sample results showing their vacabulary comprehension (passive reading comprehension). Although it may be difficult in the first version always to keep these factors in mind along with ideas for presentation of the content, the final reviser can usually easily observe and correct the more outstanding violations of the demand for adjustment to the language needs of the students (when his attention is directly fixed on this particular problem).

In this connection it might be well to repeat our previous warning against routine application of different readability formulae which deal with average values (for word length, word frequency, etc.). In a more exact and psychological sense, readability can only be studied in relation to an individual. It is an individual matter which, moreover, is in a state of flux. What is "readable" for one student in the last phase of the program would in many cases have been "unreadable" for that same person at the beginning of the program. A routine, average appraisal gives a relatively meaningless result in these cases.

The question of the precision of the expression used must always be viewed in relation both to the students' receptivity to a certain type of information and to the goal which has been set. It is not always a case of striving for maximum precision, but rather an optimum degree of precision in relation to the circumstances and to the objectives decided upon. The risk for excessive accuracy (for example, scientific exactness in a case where it does not fill a need) is considerably less - judging by published programs - than the danger of inadequate precision caused by haste or poor linguistic marksmanship. It is often a question of the hazy use of equivocal abstract words, where the simplest linguistic revision is a transition to a greater concretion (and possible exemplification).

At the present time we do not know enough about the best methods of achieving a well-balanced focusing of attention. Naturally, essential facts should not be allowed to disappear in a cloud of verbal mist. Long didules are often improved by cutting, but a routine pattern cannot be used. Didules should be regarded as information units. It is not a question of cutting an excess number of words, but of weeding out those which are not important to the context: irrelevant and meaningless filler phrases or details which do not contribute to the final objectives. It is possible to shorten long didules by (a) cutting down the amount of detail in the examples given, (b) replacing some verbal explanations with pictures or diagrams, and (c) incorporating some didules in the collateral material in the appendix. In many cases it becomes evident after a closer look at the material that it is advantageous simply to divide a longer didule into several shorter ones.

Most likely, however, an extreme condensation is not the ideal solution. The interaction between student and program can be regarded as a play of alternating tension and relaxation. The relaxation periods need not be long, but they should not disappear entirely. If every single word is loaded with meaning, the tension can be unnecessarily tiring, and the focal moment is not delineated as a clear figure against a background. Individual differences in reading habits for various types of texts should have some bearing on this, and student analyses could offer some insight into the problem. University students who change from a scientific course to liberal arts or vice versa usually experience significant difficulties in readjustment which, among other factors, should have some relation to different "density expectation".

In this play between tension and relaxation, the points of tension should be represented by the cardinal points in the content and not be language difficulties (expect possibly in regard to terms for just these cardinal points in the content). Some idea of the density of content can be derived from an analysis of the percentage of occurrence of "structural words" ("non-significant" words such as articles, conjunctions, prepositions, etc.) and "repeated words" ("significant" words which have already occurred one or more times within a given section). The lower these percentages, the fewer the "rest periods", and the greater the probability that the student will become tired because of the lack of balance in the relation of tension to relaxation (too great tension caused by too little word redundance). The cardinal points and new terms should be allowed to stand out against a more neutral background of familiar material which does not invite tension. Obviously this is a question of general principles which can be difficult to convert to specific rules for concrete examples without extensive trials. We know hardly anything more exact about the optimal values for different ages and different material. In general it can be said that an interesting field is opening up for teaching-related attention research.

A language check should also cover accuracy and consistency (is the sentence structure effective, punctuation intelligible, use of abbreviation uniform, etc.)? This aspect, however, is too self-evident to require any further discussion.

Program writing demands a great deal of the author: He should be able to <u>think</u> clearly on an <u>abstract</u> and <u>complex</u> level (in order to carry out effectively what has been defined by means of the terminal descriptions and clarified through the material analyses), and at the same time he must <u>communicate</u> clearly on a <u>concrete</u> and <u>simple</u> level (to reach his students effectively). We seldom find anyone who has maximum effectivity on both planes, and for this reason teamwork and checking should play an important role as correction factors.

12. 2 INTEGRATION

Writers of programs have often been influenced by theories with a behavioristic tinge, where analysis of sub-components in observable behavior has played a leading role. Psychologists and educators who hold different theories criticize the program writers because they "atomize" learning and because they take altogether too little interest in the "cognitive structure" or "Gestalt" of knowledge, that is, in the integration of the parts into larger meaningful wholes.

These differences of opinion have been clearly expressed at some international conferences devoted to problems of programmed instruction. On the one side we often find the American programmers, many of whom base their reasoning on a stimulus-response theory of pragmatic-behavioristic type, sometimes considered to be fairly self-evident, although to some people somewhat hazy in detail. To them the behavioral terminology is usually the natural means of communication, while expressions like "cognitive experience", "comprehension", or "cognitive structuring" are regarded as suspect remnants out of the philosophical past of psychology. On the other side we often see some psychologists from Central Europe, eager to learn the principles and technical design of programming, but often with a negative attitude to the "atomized" terminology, which to them sounds like the old psychology of association, once and for all - as they tend to think - blown sky high by the Gestalt psychologists.

Such disparities are often interesting, and it should be a definite advantage for the development of programming technique if its principles are gradually studied and worked over by persons with more varying theoretical points of departure. Nevertheless, one has the feeling at times that the differences are more verbal than actual. If the debaters can only be persuaded to climb down from the level of general and verbose theorizing to the level of operational descriptions, they can often agree on facts. The differences reflected are, in other words, in many cases more to be found in "symbol environment" and habits of thought than in more signficant divergences of view about the essential bases of facts.

Thies does not mean, however, that programming in certain forms cannot run the risk of giving an entirely too "atomized" knowledge with too few possibilities for the students to discover greater coherence and to achieve a perspective on that which has been learned. The programmer should under all circumstances be on his guard against these pitfalls (which certainly can be found even in other forms of instruction in our schedule-bound school world).

In order to facilitate "integration" and "structuring" of knowledge and to avoid that what has been learnt is experienced as unrelated pieces of mosaic, the programmer can use some of the following working methods:

(1) Preliminary organization and other supplementary information to explain procedure ("meta-didules"). When introducing a new topic, it is often advisable to build a minor unit sequence into the program, the function of which is to give the student a preliminary idea of the goal and of the scope of the study material. A method of this kind can also be used for smaller task sequences. Ordinarily such organization is of value both for the effectiveness of learning and for the general learning attitude. The student should know where he is going. The programmer can also "speak directly" to the student about the nature of the work under process in other ways, for example, by relating specific learning points to the ultimate goal of the course, pointing out special reasons for combining certain learning units, etc. Principally for what seems to be "puristic-esthetic" reasons, such extraneous comments to the students have sometimes been avoided. Of course, they can break the continuity of the flow of information material (in about the same way as the actor's aside to the audience can interrupt the rapid repartee on the stage). There is, however, good reason to assume that the introduction of such second-order communications ("communication about the communication" or "meta-didules") can have some bearing on the experienced integration of the learning area as well as on the work motivation of the student.

(2) Successive linking of detail (avoidance of discontinuous jumps). One important point in our checking phase is naturally a renewed study of instruction sequence, the individual micro-sequences (designed to teach special concepts or concept relations) as well as the total macro-sequence (with review of both the grouping of the subject areas and possibly the placing and character of branches). It is naturally important to make sure here that no significant item necessary to achievement of the goal is forgotten. But it is also of importance (especially with a view to furthering integration) to determine whether the various parts of the program are joined together smoothly enough and without disconnected jumps. In most types of courses it is both desirable and possible to tie details together to form a meaningful "associative net". This occurs chiefly at the transition from didule to didule and at the transition from

one micro-sequence to another. By making sure that the new tasks contain certain elements from the previous ones ("connecting elements") we get smoother transitions, greater probability of correct answers and better-integrated knowledge. It is, however, difficult to give any general rule of thumb for how this is best accomplished.

(3) Special assignments for furthering integration. Besides emphasis on the successive linking together of single details, it can be fitting to introduce from time to time special, more demanding recapitulation problems. Concepts and operations, which have been introduced in different connections, can then be brought together in more comprehensive assignments.

The possibility of working with program supplements, to which reference is made in the separate didules and to which the student is referred for information and compilations should never be forgotten. Such "supplements" can, for example, consist of maps, diagrams, tables or simply a summing up of verbal information. The assignments which are given in this connection can be directly aimed at integration.

It can often be helpful with shorter test sections (assignment sequences without new information and without aid-stimulation) at regular intervals, perhaps after every 20th of 40th didule. If these are well constructed, they can undoubtedly contribute to knowledge integration. They can, moreover, perform other important functions. They probably increase retention and they give the student a feeling that he really has learned something (which he does not always believe when progress is easy!)

(4) Didule patterns with integration-aim. Not only can the single didule be so formed that it promotes integration of component knowledge derived from earlier didules, but it is often possible and essential to arrange the single didules in an internal sequence so that the integration-aim is aided. Some potential procedures for accomplishing this which would be worth a more systematic testing for effectivity can be mentioned briefly:

a. Counterpoint sequence. Two content units which are to be linked together in the student's behavior repertoire are alternated systematically during the practice phase.

Example: $C_x/C_y/C_x/C_y/C_x/C_y$ etc.

(Here and in the following the slanting line is used to indicate the didule boundary and the letter C the basic subject-matter unit.)

b. Superimposed sequence. Two subject-matter units which are to be combined with each other in the student's repertoire can be organized so that the latter unit is introduced just before that point where discussion of the earlier unit is completed.

Example: $C_x/C_x/C_x/C_x/C_y/C_x/C_y/C_x/C_y/C_y/C_y/C_y$ etc.

c. Accumulated summary. When the material units are part of a larger complex (build a cohesive total structure), but still must be introduced in stages because of the level of difficulty, every new didule can be arranged so that the part-elements which have been introduced earlier can be repeated at the same time as a new element is introduced. That is: If the student's terminal behavior repertoire is to contain the complex C_x & C_y & C_z & C_u, the didule sequence can look like this:

$$C_x/C_x \,\&\, C_y/C_x \,\&\, C_y \,\&\, C_z/C_x \,\&\, C_y \,\&\, C_z \,\&\, C_u.$$

d. Chronological accumulation of temporal behavior sequences can be seen as a special application of the general principle of accumulative summary, that is, the particular case in which a behavior chain based on time order is successively built up in chronological order, as follows:

$$C_1/C_1 \,\&\, C_2/C_1 \,\&\, C_2 \,\&\, C_3/C_1 \,\&\, C_2 \,\&\, C_3 \,\&\, C_4 \text{ etc.}$$

Many behavior series are of typical temporal character, that is, they form a definite chain of actions which occur in a characteristic order. Examples are working with more complicated tools and machines as well as fault localization on different types of apparatus. Instead of first learning every part-subsection separately and then trying to put together the small bits to form a whole ("delayed integration"), it should often be more natural to practise sequence from the beginning in the manner described above.

e. Reversed accumulation of temporal behavior sequences. The temporal sequence should of course always be practised in the natural direction of association. We train, in other words, C_5 & C_6; not C_6 & C_5. On the other hand, it is quite feasible to accumulate the sequences "backwards" instead for "forwards". Instead of the didule series:

C_1
C_1 & C_2
C_1 & C_2 & C_3
C_1 & C_2 & C_3 & C_4,

in which C_4 is thought to represent the natural end link in the behavior chain, the following didule series would be used:

C_4
C_3 & C_4
C_2 & C_3 & C_4
C_1 & C_2 & C_3 & C_4

The advantage of the latter arrangement can be that the student often derives greater motivati- · from being allowed to perform the terminal phase of a behavior seri. .t an earlier stage as well as to see each practice series finished off with the goal clearly attained. Especially in those cases where the terminal phase of the behavior series is one in which the result stands out clearly (gives a definite, even dramatic demonstration of goal achievement) this type of sequence arrangement can be of value. (This method or procedure has been particularly advocated within the frame of the so-called "mathetical approach" represented by Gilbert and his disciples; see, for example, Gilbert 1969; Pennington & Slack, 1961).

(5) Adaptation of study material to divergent study behavior motivated by diverging initial repertoire or aims (with gains in integration when used for review). Our ordinary textbooks and handbooks can be used in many ways. Besides being successively learned as course literature, they can be "skimmed" or "read diagonally" by someone who has already mastered the principal contents of the subject (in which case "skimming" provides a quick survey and the possibility of discovering any new points). They can also be used as reference literature by those who have a more specific end in view than that of learning the subject as a whole: by use of the table of contents or the index, it is easy to look up a special part which is relevant to the current object without studying the other material in the book. One criticism of programmed study material has been that the specialized form (for example, the fixed arrangement of "small steps", arranged in "scrambled" order or in horizontal "bands") makes it difficult or impossible to use the study material in a flexible way, and that the reader is bound to using it for only one single reading method - according to the rigid pattern of careful memorization. This means too that those who want to brush up their knowledge after a period of time and once more get a comprehensive view of the material which has been covered have difficulty in doing so - unless they take the time-consuming way of again going through all the material in the prescribed order.

Anyone who tries to read through a finished program for a familiar subject area can easily be persuaded that these objections are valid. It is true enough that the programmer can defend himself with the argument that the program is designed for persons with a certain type of initial repertoire and is meant for a certain type of goal, and that because of this the programmer has no reason to take into consideration the fact that persons with a different initial repertoire (greater knowledge) and/or different goal (covering only certain parts of the field) think that the study material is not practical.

On the other hand, it can reasonably be argued that difficult surveyability, with the rigidity in usage which follows, is often greater than necessary and that the writer could by relatively simple means make the total structure considerably more transparent - to the benefit of, among other things, later reviews for improving retention. Even though, naturally, the primary usage should determine the presentation and arrangement, the secondary usage (as review or reference book, for example) should not be made more difficult than necessary.

One technique which could easily be used to increase the structural surveyability is that of working out detailed tables of contents with references to didules as well as drawing up a list of essential terms and ideas (also with references to didules), and also printing the integrating main sections (preliminary introductions and outlines, for example) in italics, in another color, or emphasizing them by some other means. Ordinarily it is an advantage if the main sections ("chapters") are clearly separated from one another by page. It is usually easiest to set up separate chapter sections in programmed books with vertical progression, but this can very well be done in programmed books of other types.

12. 3 RETENTION

The resistance to obliteration of that which has been learned is naturally affected by many factors. Several of the points on readability and integration which were presented above are also of importance in connection with retention. For example, the well-integrated material has - other things being equal - greater resistance to obliteration than the not so well-integrated. In this section we shall confine ourselves to summarizing some brief rules for repetition:

(1) Avoid too little repetition! Many a programmer includes too little repetition since the construction of repetition assignments is often boring for him and - since he himself is so familiar with the material - may also seem unnecessary. The student, on the other hand, who meets for the first time a number of new ideas and associations, is normally in great need of repetition. Retention of that which has been learned depends, naturally, to no small degree, on the "overlearning" which takes place in the instruction set-up. In this connection, however, the time-lag between the instruction occasion and the terminal situation must be taken into consideration as well as the student's opportunities for natural repetition or practice during the intervening period. Any possible dangers of interfering factors ("competing behavior patterns") must also be noted. As a guide for these judgments, the initial studentanalysis is of value and also the "expanded terminal chart" worked out in conjunction with it (cf. discussions on initial system analysis).

(2) Vary repetition! Routine drill with direct repetition of the same assignments is seldom desirable, partly because going through the material in this manner can easily become boring, partly also because the programmed sequences are worked out for those who have no elementary knowledge and when repeated must be assumed to function less effectively. The practice should instead consist of illuminating and applying the acquired conceptions in new contexts or with different examples. It is often a good idea in this connection to integrate repetition of old material with the introduction of new. In this way repetition monotony is avoided, the area for meaningful associations is increased, and desirable generalizations are facilitated.

(3) Spread repetition! As has been noted above, it is true that a concentration of learning can be effective from the point of view of speed but at least in the case of more comprehensive instruction material and long retention periods a certain spreading out of repetition should be desirable. It is thereby possible, for example, to treat the same idea first in a series of consecutive didules and later to allow the idea to turn up again at certain intervals (perhaps shorter in the beginning and successively with greater infrequency later on). A "distribution plan" could then take the principal form sketched below in Figure 12.2 below.

(4) Individualize repetition as much as possible! The most usual procedure, that of allowing the student to pass from one assignment to the next whether he has answered correctly or incorrectly, without

returning to it later, has certain disadvantages. As mentioned above in our discussion of various "flow models", a "re-directing" to differentiated repetition (of assignments answered incorrectly at an earlier stage) may sometimes be worth recommending. Skinner's disc-machine (also referred to above) functions in just that way, and Holland & Porter (1961) were able to demonstrate that such a method can improve the final result of the student. These methods are especially suitable for learning "subject-matter items" (as, for example, in learning vocabulary in a foreign language), where this differentiated re-direction does not break the continuity. Those types of machines which function according to the "skimming"-principle can be useful here, but simple card collections are often equally effective. In these cases every didule can be noted on an assignment card (with the answer on the reverse side). Those questions which the student answers incorrectly can be placed by him in a special pile. He can then easily go through these once more (or as many times as is necessary until he has performed all assignments correctly).

12.4 MOTIVATION

Reading a long, continuous text without "built-in" work assignments or listening to an extensive oral exposition can be rather monotonous for many students. The result can be increased fluctuation in attentiveness with an accompanying reduction in the effectiveness of the instruction. Some of the increase in effectivity which in many cases has occurred in conjunction with the use of programmed study material seems to have some connection with the less monotonous work dynamics in the didule system (rather than with the sometimes too greatly stressed reinforcement effects of feedback presentation). The student experiences a constant task variation:

(a) he rads the information and the demand for an answer
(b) he works through to an answer
(c) he writes down the answer
(d) he checks his answer and prepares to go on to the next assignment.

The dynamic variation, the play between tension and relaxation (which we discussed in connection with readability), can here perform an important function, and the differences in the activities listed under a, b, c, and d are probably of importance. It is a reasonable hypothesis that points a and b ordinarily signify a rising curve of tension (with the summit at b), while points c and d represent a certain psychic relaxation during the phases with more predominant motor activities (writing and machine manipulation). Both the subjective experience of the dynamic alternation and possible psycho-physiological consequences whould be worth a special research effort.

But even if the programmed study material because of this inherent alternation of tasks is less monotonous than a running text or a continuous oral presentation, it is quite natural that the student can experience a certain monotony also in the programmed study material if, for a longer period of time or in several different connections, he works with the same type of small presentation units (as long as we have only a few and/or short programs there is relatively little danger of this). Several of the measures mentioned earlier for promoting the building of an integrated field of knowledge for the students can also be useful in reducing the dangers of monotony: for example, work with integrating supplemental material in appendix form or going thorugh built-in test sections. Other methods of counter-acting monotony can be to alternate types of answer (certain sections perhaps without demand for an explicit answer), or to incorporate more comprehensive student tasks (directing the students to various source material; a technique termed above "the intermittent model"). There is a risk that, through exaggerated respect for one special method, we fail to try the united effect of different methods. We must keep in mind to avoid this risk.

That the experience of achievement is of great importance for the student's motivation should be indisputable, but in reality the problem is more complicated than some programmers usually describe it. We must take into consideration the great individual variations of earlier achievement experiences and the current aspiration levels as well as some more stable personality characteristics which may be of importance in this connection (for example, what McClelland and others have called "need Achievement"). The fact that such complications do not require much attention in the world of rats and pigeons (and some programmers always seem to think they will find the solutions to problems of the psychology of learning right in the rat laboratory), should not keep us from being aware of them in human learning situations.

We should consider at least to aspects of the problem complex. On the one hand we should note the fact that the learning effectivity of an instruction material is often greater when there is a high frequency of positive feedback. This was one of the chief viewpoints of the early programmer group inspired by Skinner, and it was usually assumed that the error frequency should be low (perhaps under 5 %). On the other hand, the maintenance of motivation in the long run (consistence in work) also seems related to whether demands are high enough in relation to ability. Individuals with high "need Achievement" and good learning technique can, for example, find very boring a long series of all too easy tasks, and this, reasonably enough, may decrease their interest in efficiently working with the material.

In this event we should perhaps try to make use of a "multiple success strategy". This can mean that while we work fairly constantly with a high probability of correct answers to separate tasks, we combine that procedure with "extra success effects" - for example, the possibility of rapid-track progression and/or the satisfaction of being able to answer correctly without clues (individualized offering of clues as in teaching machine Didak 501). Becker (1963) who proposes such a method of procedure compares this strategy with that which is used in certain lotteries or chewing gum machines. For one's "contribution" (a coin) one always receives a certain "prize" (a piece of chewing gum), but there is also the chance of getting a larger "prize" (an attractive object). A play of this kind on two planes of success should - mutatis mutandis - be serviceable in teaching material and might perhaps better do full justice to the varying individual needs.

The reviser cannot, naturally, make too radical changes in the total structure (these should be allowed to wait until the empirical data also are available). But he should make sure that nothing in the study material which can encourage boredom or discourage motivation is allowed to pass. Any possible monotony can ordinarily be eased through small changes in working methods. And if the reviser finds that the primary material provides all too little scope for the ambitious student to feel that he can achieve something on a par with his ability, then the addition of "extra success effects" should not be too difficult.

12. 5 VARIOUS AIDS IN THE CHECKING PHASE

The fact that successive empirical try-outs of the materials are the final and most important steps in the evaluation process does not mean, of course, that we have to rush into premature data collection with hastily produced and unpolished versions of the study material. While an arbitrary reading through of the material may be fairly inefficient, a systematic mapping of certain apparent characteristics of the material, on the other hand, may be very fruitful. We may call this examination of apparent characteristics, the mapping of the "pheno-structure" of didactic sequences (as distinguished from the mapping of the "effect-structure" that might be carried out during the empirical try-outs later on).

In the present section several aids for this mapping process will be described and discussed, including check lists, diagrams of the relations between terminal objectives and the single didactic units, and various kinds of unit-charting protocols. It is felt that the examination is considerably facilitated by summarizing aids of these kinds. The risk of forgetting some important aspects of the examination is reduced, and a long series of questions about the material can be answered much more easily on the basis of a summarizing protocol than on the basis of the non-aided perusal of a long sequence of material, in which the different parts cannot be kept in sight at the same time.

It should be stated from the beginning that many different types of aids are possible, of course, and that each programmer should design his working tools to suit his own needs. The following descriptions should therefore be looked upon only as illustrations of various possibilities.

12. 5. 1 Examination by Means of Check Lists

The simplest way of remembering the important aspects that should be examined is probably to make up a fairly detailed list of questions to be answered for each self-instructional study material and to use this list for checking off systematically the answers for each separate part of the course material. An example of such a list of questions is given in Box 12. 1. (It should perhaps be added that some of the questions in this list are such that only the subsequent empirical try-out can give a definite answer. Nevertheless, it is often possible to make intelligent guesses in advance, and thereby to save some steps in the revision process.)

Box 12. 1 Check list for pre-test examination: An illustration

1. Examination of goal relevance and content:

1. 1 Does the program contain all that it should according to the initial goal statements?

1. 2 Has all irrelevant material been eliminated? (Irrelevant materials are those that are neither included in the goal statements, nor fill any clear-cut educational function as aids in the process of reaching these goals.)

1. 3 Have all those points been eliminated that, according to the target population analysis, the students have already mastered?

1. 4 Does the program avoid giving any single subject-matter aspect a coverage that, according to the goal statements, could be considered unreasonably large or unreasonably small?

1. 5 Has all changeable information been checked, so that the students are given maximally up-to-date information?

1. 6 In the judgment of subject-matter experts, is the content free from factual errors?

2. Examination of the "procedural instruction" of the program:

2. 1 Are the "procedural instructions" of the program to the student (instructions about the ways of handling the material during study etc.) complete, clear and easy to follow?

2. 2 Is the student given a sufficient amount of training in how to give answers or handle the teaching machine in the beginning of the program (as well as in other places where perhaps a new answering technique is introduced)?

3. Examination of the organization and sequence of content:

3. 1 Does the program try systematically to establish a connection between those experiences or concepts that the students already have in their repertoires and the new knowledge that is introduced?

3. 2 Is the program constructed to use - where this is possible and appropriate - the sequential approaches that are usually referred to.by phrases such as "from the simple to the complex", "from the easy to the difficult", "from the concrete to the abstract"?

3.3 Are there sufficient exercises for application and repetition?

3.4 Are the units with material for repetition adequately spaced (in the beginning sufficiently near the first presentation to make correct answers possible and after that with increasing and sufficient intervals to maintain retention over large enough periods of time)?

4. Examination of the stimule-function of the single didules:

4.1 Have irrelevant and distracting details been removed, so that the attention of the student is clearly focused on the central part of the information of the didule?

4.2 Is the communication presented in a sufficiently life-like and concrete way? For instance, have pictures and supplementary demonstration materials been used where such aids can supply the intended information more effectively than verbal description?

4.3 Are the examples given sufficiently varied so that the student does not get a wrong and one-sided picture of the principle or concepts being treated?

4.4 Is the position of information appropriate, so that, for instance, important information has not been placed after the response request?

4.5 Have the prompting techniques used been sufficiently varied?

4.6 Has unnecessary prompting been avoided?

4.7 As a rule, has meaningful prompting (by means of logical induction, parallel exemplification etc.) been chosen where appropriate rather than unnecessary formal prompting of a kind that may distract the student's attention and divert it to unimportant aspects of the communication?

5. Examination of the respule-function of the single didules:

5.1 As a rule, can the questions be answered on the basis of the information that the student has received?

5.2 As a rule, have response requests been avoided that can be answered without understanding of the content of the didule (by formal-linguistic pattern completion etc.)?

5.3 Has the response request been designed so as to give a reasonable guarantee that the student has understood the essential communication of the particular didule? In other words: Are the response requests relevant to the aim of the particular didule (always assuming that each didule has a specifiable aim)?

5.4 Does the program activate varying types of response behavior (writing, drawing, calculating, comparing, etc.)?

5.5 Do the alternatives of a multiple-choice question represent all reasonable, non-trivial sources of misunderstanding? In other words: Is it difficult to think of typical student errors not included among the alternatives listed?

5.6 Are the alternatives in a multiple-choice question so designed that the student's choice is not a pseudo-choice (e.g., so that the incorrect alternatives are not clearly absurd or formally incongruent with the main question)?

5.7 Does the program request the student to give those types of answers that the terminal situation will demand from him (so that, for instance, the program does not merely train the student to recognize correct decisions, if in the terminal situation the student's ability spontaneously to make correct decisions in also going to be tested)?

6. Examination of the integrative function of the program:

6.1 Is the student given an opportunity of applying concepts and principles, first presented stepwise, to problem-solving of a kind that forces him to activate and operate upon several concepts or principles at the same time?

6.2 Are integrating materials in the form of reference tables, reference diagrams, reference maps or reference texts utilized where appropriate?

6.3 Does the program help the student to obtain an adequate initial organization ("Gestalt in advance", "properly structured expectations") as well as an adequate final organization ("final Gestalt", e.g., by means of "properly structured reviews")?

6.4 If a programmed textbook is used, does the student get a table of contents outlining the broad scope and structure of the total contents?

6.5 Are there adequate indices that make it easy for the student to brush up his knowledge afterwards within specific areas of information, if he so desires?

7. Examination of the motivating function of the program:

7.1 Have the programmers tried to provide a sufficient degree of variation in the review items?

7.2 Does the difficulty level of the program appear, throughout the program, to be so adapted to the particular student group under consideration that there is a fair probability that the students will be neither bored nor discouraged?

7.3 Have the examples and illustrations - to the extent that seems possible - been made interesting and relevant to the experiences and needs of the particular student groups?

7.4 Do the possibilities of individualization within the program (available branching arrangements) correspond closely enough to the results of the initial analyses of the target population, so that, for instance, students with considerable previous knowledge of the subject matter are not forced to go through the same material as students without any previous knowledge?

8. Examination of the external form of the program: Language

8.1 Do the level of vocabulary and the structure of sentences seem to be sufficiently well adapted to the linguistic habits of the particular student group so as not to be an unnecessary barrier to communication of the main contents?

8.2 Is the meaning of new terms defined or otherwise demonstrated clearly enough as soon as they are introduced?

8.3 Is the language of the program always clear and unambiguous?

8.4 Does the language always have an optimal degree of precision as judged from the goal statements of the program (that is, neither with too low a degree of precision, including sloppy everyday terms where precision is needed, nor with an unnecessarily high degree of precision, such as various scientific distinctions not to be further utilized in the desired terminal behavior)?

8.5 Is the use of numbering, punctuation, abbreviation etc. both correct and consistent?

9. Examination of the external form of the program: Other aspects

9.1 Is the general lay-out of the program both educationally appropriate and economically defensible?

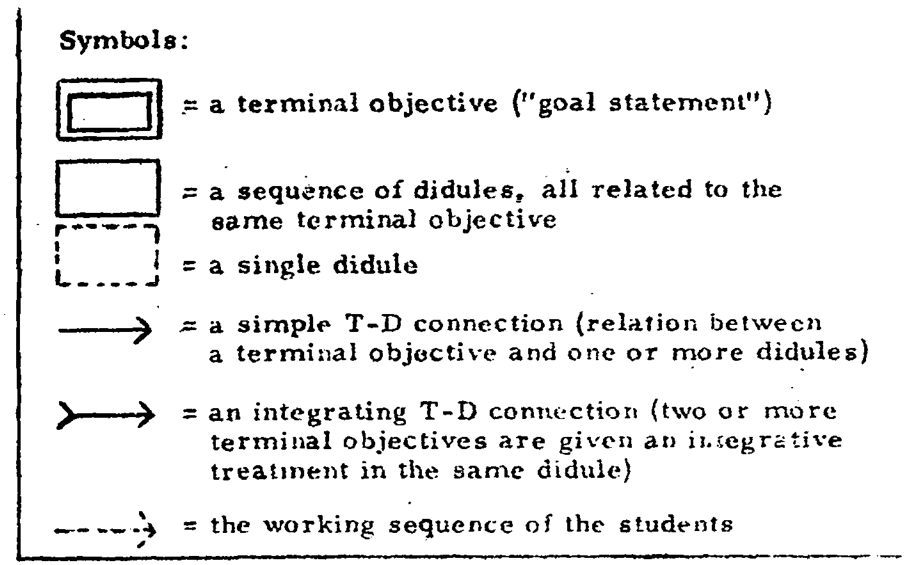
9.2 Is the typography sufficiently clear for the intended student group and also designed to emphasize points that should be emphasized?

9.3 Are illustrations of various kinds, both within the program proper and in possible appendices to be used as reference material, so designed technically that they can be expected to result in the best possible communication effect?

12. 5. 2 Examination by Means of "T-D Diagrams" Depicting the Relations between Terminal Objectives and Single Didactic Units

Even during the writing process proper many programmers find that it makes the work easier to draw simple diagrams of the relations between the terminal objectives as codified in the goal statements on the one hand and the single didactic units (the didules) as appearing in the written program on the other. Very often the program writer arranges both the specific goal statements and the single didule texts on separate cards in a card catalogue. This means that the diagrammatic survey often takes the form of a series of relational diagrams, from which it is easy to see which didule cards (D_1, D_2, etc.) are related to each single terminal card (T_1, T_2, etc.). During the post-writing phase it is easy to examine, with the aid of diagrams of this kind, whether or not certain parts of the total goal structure have been given too much or too little attention in the program. Similarly, it can easily be seen whether or not enough attempts seem to have been made to reach an integrated structure of knowledge. Figure 12. 1 shows an example of such a diagram. The numbers of terminal objectives and didules included in the figure (T- and D-numbers) refer to the program fragment in Box 12. 2.

If we deal with a so-called scrambled book, the diagram should include not only information about the main part of any particular didule (the page with go-ahead signal and new information), but also information about the error-treatment pages (the pages that explain errors and refer the student back to the main pages again). As a rule, in these cases the programmer should also use a separate "pagination control list" for checking off page numbers used. The reason is, of course, that the programmer and the post-writing examiner have to keep an eye on the scrambling system, sometimes fairly complicated, in order to ensure a suitable sequence of pages, that is, a sequence which is instructionally adequate as well as economical with respect to time and space. (No pages should be left empty; the student should not need to turn too many pages each time; tasks that might make the answer easier than intended or that would oterhwise interfere with each other should not be placed side by side on the same two-page fold-up of the book, etc.)



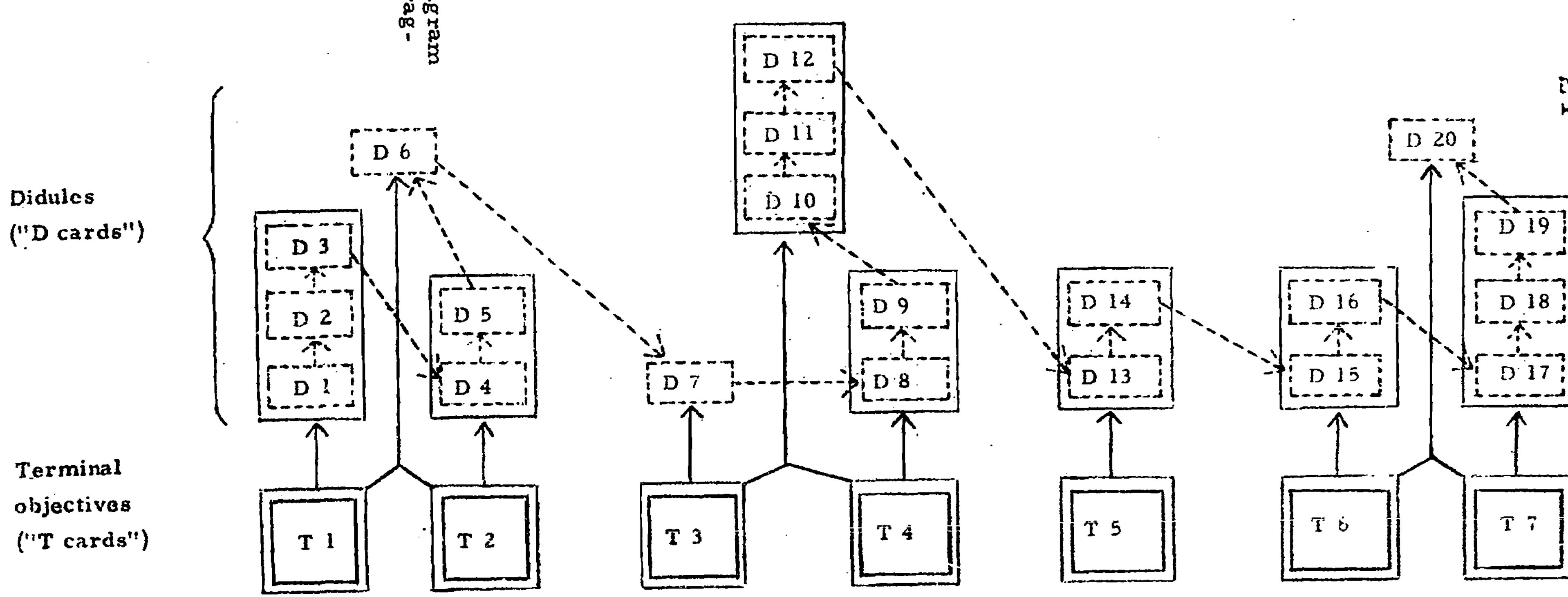


Figure 12.1 An example of a T-D diagram (mapping the program fragment in Box 12.2)

Box 12.2 Terminal descriptions (T cards) and didules (D cards) from a program fragment on programmed instruction (as supplemental material to the T-D diagram in Figure 12.1)

Topic: Introductory course on programmed instruction

Subjects: Teacher students

Note: This fragmentary illustration contains 20 didules (D), covering 7 specific terminal objectives (T) - focusing, in turn, on 3 fairly basic "principles".

T 1. The student should be able to state some acceptable reason for the emphasis on self-pacing in self-instructional materials.
(Principle 1)

T 2. The student should be able to recognize and recall the term "self-pacing".
(Principle 1)

T 3. The student should be able to state some acceptable reason for the emphasis on "overt responses" in self-instructional materials.
(Principle 2)

T 4. The student should be able to recognize and recall the terms "overt response" and "covert-response".
(Principle 2)

T 5. Given a set of instructional items, the student should be able to differentiate between items demanding "overt" and "covert" responses.
(Principle 2)

T 6. The student should be able to state some acceptable reason for the emphasis on low error-rate in self-instructional materials.
(Principle 3)

T 7. The student should be able to state some specific characteristics in self-instructional materials that make low error-rate possible, including, at least, the two notions of "small steps" and "careful sequencing".
(Principle 3)

D 1. In the classroom, the teacher most often works with the total class. If so, all students have to go forward at the same speed - for instance, at the speed of the average child. This rate often tends to make the work (more/less) interesting to the bright and quick-working child.
(Cf T 1) /less/

D 2. What would you guess about the effect on the slow learners of following the average student's speed?
Probably, they (will/will not) grasp enough of the material presented in the time allowed.

(Cf T 1) /will not/

D 3. If the slow learner does not grasp enough of the material presented, he will easily develop a (positive/neutral/negative) attitude towards classroom learning.

(Cf T 1) /negative/

D 4. Students working at their own speed are said to follow the principle of "self-pacing".
Presumably, the slow learner would be able to grasp more fully the material to be studied, if he were allowed to work at his own speed. Consequently, he would probably also develop a more positive attitude toward classroom learning, if the principle of _________ were used.

(Cf T 2) /self-pacing/

D 5. Since the quick learner is less likely to lose interest in the work (less likely to get bored), when allowed to proceed as quickly as he is able to, the principle of __________ also favors the bright and quick pupils.

(Cf T 2) /self-pacing/

D 6. To sum up, then, there is reason to believe that "teacher-pacing" is (more/less) likely to be a favorable learning condition than "________________".

(Cf T 1 & T 2) /less; self-pacing/

D 7. The student who actually works through a statistical calculation is (more/less) likely to learn effectively than a student who merely reads a description of what to do.

(Cf T 3) /more/

D 8. When a student learns, he may make both <u>overt</u> and <u>covert responses</u>. Those activities we can easily observe (like writing or manipulating a machine) we call overt response. Those activities we cannot observe (like thinking) we call ________________ responses.

(Cf T 4) /covert/

D 9. Using our technical terms "overt" and "covert" responses, we may say that the student of statistics who merely reads a discription of what to do, makes ________________ responses to the text, where as the student who works out the calculation in writing in addition shows ____________ responses.

(Cf T 4) /covert; overt)

D 10. A house-wife who tries to learn cooking from a book just by reading may be said to make only ____________ responses to the text. There is reason to believe that this is a (more/less) effective way of learning cooking than actual practice.

(Cf T 3 & T 4) /covert; less/

D 11. (Observable/Non-observable) reactions are, of course, a better guarantee that the student has actively responded to all the important aspects of the materials than are (overt/covert) reactions.

(Cf T 3 & T 4) /Observable; covert/

D 12. To sum up, then, covert responses are (more/less) likely to be effective in learning than are ______________ responses.

(Cf T 3 & T 4) /less; overt

D 13. Look back to the three sentences of item 8 above. In which sentence were you expected to make an overt response? Answer: In the ____________ sentence.

(Cf T 5) /third/

D 14. Look at the four items in Panel I. Put checkmarks in the list below to indicate the kind of response expected in each case from the student.

Item	Covert only	Overt
A		
B		
C		
D		

(Cf T 5) /Note. Panel I is excluded from the present illustration/

D 15. A student who is anxious or uninterested in the work may not be motivated to continue working. A student who makes very many errors may get anxious or uninterested in the work. A student who makes very many errors, therefore, (is/is not) likely to be motivated to keep on working.

(Cf T 6) /is not/

D 16. If we make an error, we tend to remember the error, if it is not corrected. If we give a correct answer, we tend to remember the correct answer. In the first case we have to "unlearn" the error. In the second case we "learn" directly. Consequently, learning might become more efficient, if we could arrange the study material so that the students made (many/few) errors.

(Cf T 6) /few/

D 17.	If the student is going to reach the goal we have set without making many errors, we have to arrange the study material in (small/large) steps. (Cf T 7) /small/	
D 18.	If ITEM X is something the student needs to know before he can learn ITEM A and ITEM B, we should begin his sequence of items with ITEM ________. (Cf T 7) /X/	
D 19.	In order to make sure that the student makes few errors only, it would seem reasonable to recommend the use of ____________ steps and a careful ____________ of items. (Cf T 7) /small; sequence/	
D 20.	To sum up, if we use small steps and a careful sequence of items, we are likely to get (more/fewer) errors. In that case motivation to keep on working will usually be (higher/lower) and the learning will proceed (more/less) efficiently. (Cf T 6 & T 7) /fewer; higher; more/	

12. 5. 3 Examination by Means of Didule Protocols

A didule protocol is a device which makes it easier to make the examination both complete and systematic. Examination by means of general checklists aids the examiner in covering many aspects, but it may not by itself help him to cover every single part of the program in a systematic way. As a rule didule protocol includes, for each didule, a separate examination column, in which - depending on the particular aim - various kinds of questions may be answered when this particular didule is examined. Three general types of such protocols may perhaps be distinguished: subject-matter protocols, methodological protocols, and combined protocols studying both subject-matter distribution and the methodological approaches used.

12. 5. 3. 1 Subject-Matter Protocols

The main aim of the subject-matter protocol is to some extent similar to the aim of the T-D diagrams described above, viz. to facilitate the examination of the distribution of the subject-matter over the didule sequence. A square-ruled paper may be a good starting-point. The numbers of the didules are written in numerical order along the top of the paper. The various subject matter units that are to be examined are listed down the left-hand margin. For instance, these units may be single conceptual units or conceptual relations which have been listed during the pre-writing phase as important key points to be covered in the program. Sometimes it is more natural to use psychological units than logical units, that is, to start out from a series of stimulus-response connections rather than from concepts and conceptual relations. Whatever the particular kind of units used, however, the general examination process will be similar. By means of some kind of check marks in the appropriate cells, we report in which of the didules a particular subject-matter unit is treated (cf. Figure 12. 2).

Among the questions that may be answered with the aid of such a protocol, the following may be mentioned: (a) Is each single basic subject-matter unit represented to a sufficient degree in the program? (b) Are some of the didules totally irrelevant to the main subject-matter to be covered? (c) Is the treatment of the single subject-matter units appropriately distributed (for instance, an intensive series of didules at the first presentation, followed later on by separate repetition items with gradually increasing intervals)? (d) To what extent have attempts been made to integrate the separate units into meaningful structures, and are these attempts made at optimal points in the sequence (for instance, where the separate concepts are sufficiently well treated and practised)?

When answering the last question, for instance, the examiner will study to what extent the same didule column contains check marks referring to several different subject-matter units. It is usually desirable that this is the case in many didules in order to facilitate integrated structures of knowledge. At the same time the examiner of the program studies at what specific points in the didule sequence the new subject-matter units are first presented. As a rule it is confusing to the student, if several new subject-matter units are introduced in the same didule. An integrating didule therefore usually contains either a series of units introduced earlier in the program or one new unit of subject-matter in combination with one or more units introduced earlier.

As mentioned above it is sometimes maintained that integration of knowledge may be facilitated by specific patterns of didules, such as systematic alternation ("counterpoint"), systematic overlapping, chronological accumulation, or reversed accumulation of temporal behavior sequences (cf. section 12.2). Further research is needed on these and other patterns of didules recommended for integration purposes. In this experimentation, it is, of course, valuable to be able to examine, by means of subject-matter protocols, the appearance of the various patterns under study. It is then easy for the examiner to observe inadvertencies and to correct arrangements which have been left in a less than optimal shape during the writing process. (A few examples of typical patterns referred to by the terms used above are given in Figure 12.3.)

12.5.3.2 Methodological Protocols

The purpose of the methodological protocols is to facilitate the examination of various methodological approaches and make it possible to answer systematically questions of the following types:

(a) Has the programmer got stuck in certain working routines with the accompanying risk that the student finds the program monotonous?

(b) Are the educational techniques used of such a kind that, besides facilitating the student's learning of certain subject-matter facts, they also contribute to educationally desirable side-effects (such as good working habits, positive study attitudes, etc.)?

Figure 12. 2 An example of a subject-matter protocol (didule protocol, type A)

Subject-matter units	Didules																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C_1		X	X	X			X					X							X	
C_2					X	X	X			X				X						X
C_3								X	X	X				X						X
C_4											X	X	X	X						X
C_5															X	X	X	X	X	X
Unrelated	X																			

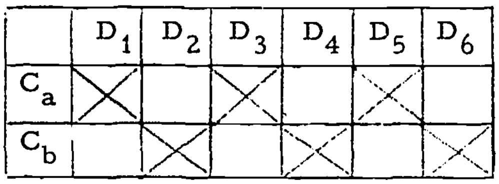
Reading key: Subject-matter unit C_1 is treated in didules 2, 3, 4 and later on also in didules 7, 12 and 19. Subject-matter unit C_2 is treated in didules 5, 6, 7 and later on also in didules 10, 14 and 20. Etc.

Comments: The present example shows a subject-matter distribution often considered to be favorable. Only the introductory didule is unrelated to the subject-matter to be covered. Each new subject-matter unit is introduced through a number of didules in sequence and is later on repeated, first after a short interval, then after increasing intervals. The repetitions are at the same time arranged so as to include integration with other subject-matter units.

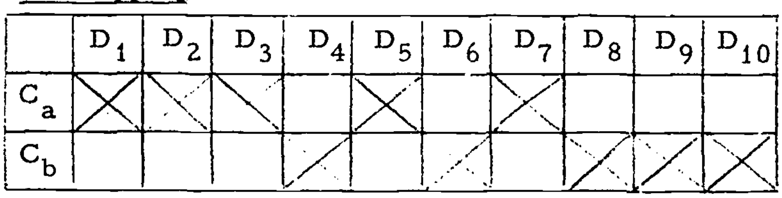
Figure 12.3 Examples of special types of didule sequences as reflected in a subject-matter protocol

Note: In this figure, numbers have been given to subject-matter units that make up a connected temporal chain, whereas letters have been given to subject-matter units that are not temporally related to each other. A small arrow below the didule number shows in which order different parts appear within a single didule.

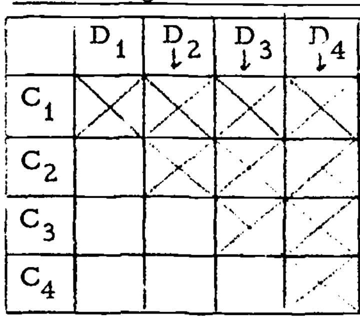
a. Systematic alternation ("counterpoint")



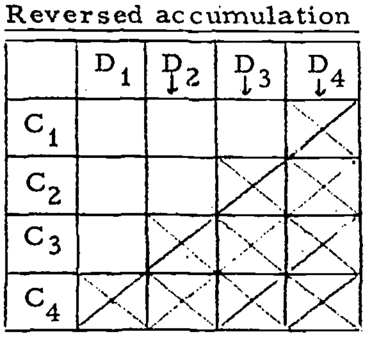
b. Overlapping



c. Chronological accumulation



d. Reversed accumulation



The protocol sheets will be similar to the ones used for subject-matter protocols, that is, square-ruled papers with didule numbers along the top. However, questions about the educational approach or key-terms for different methodological categories (instead of subject-matter units) are now written down the left-hand margin. If the question is answered by "yes" for a certain didule, the examiner puts a check mark in the appropriate cell (where the didule column and the row of the question intersect). Similarly, he inserts a check mark whenever the didule under study can be classified under the methodological category listed in the margin.

A very simple question that may be included for study in a protocol of this type is the following: Where has the student to go in order find a solution to the task set for him in the didule? Among the various possible sub-categories representing different groups of conceivable answers to this question, we may, for instance, mention the following:

(1) The solution can be found directly in the text of the same didule ("internal-explicit" source).

(2) The solution can be indirectly deduced (through a process of drawing conclusions etc.) from the new information that is given in the same didule ("internal-implicit" source).

(3) The solution can be found in the information presented earlier in the program ("external-proximal" source).

(4) The solution cannot be found within the program, but can be constructed in some way, usually with the aid of the student's background experience and/or his ability to use logical reasoning (or mathematical operations) ("external-distal" source).

At first sight, a subdivision of this type may seem to represent an unnecessary and pedantic eagerness for box-sorting. However, the observation of "response sources" is a very important part of the didule examination, and many existing programs would probably have looked quite different and been more effective, if the programmer had been made aware early enough of one-sided techniques utilized. The special terminological labels used ("internal, explicit" etc.) should not, of course, be taken too seriously, but a brief and handy term may be useful in writing out these protocols as well as in discussing these phenomena.

Some of the categories mentioned above can be further subdivided, if desirable. The third category, for instance, may be divided into maximally proximal cases (when the answer is to be found in the immediately preceding didule) and other cases (when the answer is found earlier in the program). Separate "appendices"(or "panels") as response sources can, when appropriate, be listed as a separate category, etc.

Such simple mapping gives the examiner quick information about the possible monotony of a sequence, showing, for instance, whether or not the programmer has used exclusively "internal-explicit" sources (which seems very often to be the case in some published programs). Each technique may serve an important function. The internal-explicit approach focuses the attention of the student on the key points in the immediate text and seems to be of particular value when new terms and concepts are introduced. The external-proximal method contributes to the gradual training of retention and may also be used in the process of integration. In order to ensure valuable educational side-effects (good working habits etc.), the programmer will probably often find it most profitable to use either an internal-implicit or an external-distal technique.

Other questions that can be studied in a similar way, are: What type of activity is expected from the student (for instance, copying, guessing, analogous exemplification, induction from own experiences, logical conclusions)? What kind of prompting is used? How is the answer related to earlier answers and to the key-points of the didule? What is the linguistic structure of the didule? (For instance, how high is the "density" of information-loaded words?) Etc. etc. (For further illustrations, cf. Figure 12.4.)

In certain cases it might be desirable to calculate some index score covering the "educational variation" (the distribution of didules with respect to a series of sub-categories related to a certain methodological question), or some index score covering the proportion of didules including certain educational features probably resulting in desirable side effects (when some of the sub-categories belonging to a certain methodological question are thought to be educationally more desirable than the other sub-groups). Such calculations should not, of course, be used too routinely, nor with too great a confidence in the exactness of the scores derives. Nevertheless, they can serve an important function as general warning signals, telling the examiner, in fact, "The program ought not to look like this. Revision urgently needed!" The methodological protocol can then show more exactly which particular points should be revised.

Figure 12. 4 An example of a methodological protocol (didule protocol, type B)

	Didules									
Methodology:	1	2	3	4	5	6	7	8	9	10
1. Response sources:										
1. 1 internal-explicit	X	X	X	X	X		X	X	X	
1. 2 internal-implicit										
1. 3 external-proximal						X				X
1. 4 external-distal										
2. Types of prompting:										
2. 1 word-form prompts	X	X	X	X	X		X	X	X	
2. 2 formal emphasis	X	X	X	X	X		X	X	X	
2. 3 syntactic controls	X	X	X	X	X		X	X	X	
2. 4 parallel exemplification										
2. 5 induction from experience										
2. 6 logical conclusion										

Reading key: In didule 1 the student can obtain the information needed for his answer directly from what is said within this didule ("internal-explicit" response source), and the answer is facilitated by strong formal prompting (word-form prompts, emphasizing the key words, and syntactic controls). Etc.

Comments: The example shows a one-sided use of "internal-explicit" response sources as well as formal prompting. In two cases, however, there is no prompting at all, but the student is expected to recall material from earlier didules (no. 6 and no. 10). In no case is use made of parallel exemplification, induction from experience, or logical conclusion. If the one-sidedness illustrated here continues over a larger material, there is the risk of both monotony and less adequate side-effects (lack of training in appropriate study techniques).

12. 5. 3. 3 Combined Protocols

Combined protocols register both subject-matter distribution and particular methodological approaches in some kind of combination. In many cases it is not enough to see whether the subject-matter is suitably divided over the didules, and whether the methods used are sufficiently varied. We also want to make sure that the various methods enter at the most appropriate points in the subject-matter sequence. Sometimes we consider certain methodological approaches more suitable when a new concept is introduced than when the concept is already relatively well incorporated in the behavior repertoire of the student. The types of prompting techniques that we are perhaps reluctant to accept when we deal with a terminology that is well known to the student (for instance, prompting of a wholly formal-mechanistic type), we may sometimes find acceptable, or even natural, when we introduce a new term. In addition, we sometimes want to make sure that we do not connect a certain subject-matter area too closely with a certain method in the student's behavioral repertoire. Not only do we want to have an overall methodological variation, but now and then we also have to check that we use a methodologically varied technique in the treatment of each single subject-matter area.

In some respects, the combined protocols are very similar to the subject-matter protocols: a square-ruled arrangement with didule numbers along the top and subject-matter units listed down the left-hand margin. However, instead of the simple check-marking process, we now use a series of different symbols for various methodological categories (cf. Figure 12. 5).

12. 5. 4 A Final Note

It should be mentioned that protocols and diagrams of these types may often be good training instruments for persons wanting to train their ability to observe educationally important differences between various published programs. Hence, in courses about self-instructional materials and programming, protocols of these and similar types seem to be well suited to make the practical training of the study group more effective and efficient. However, the main aim of these protocols is, of course, to be an aid in the process of constructing a program making the programmer or an independent examiner aware of possible faults in the program even before the first empirical try-out. As stated above, it is hoped that this will reduce the work and cost involved in the final phase of successive revision.

Figure 12.5 An example of a combined protocol (didule protocol, type C)

Subject-matter units	Didules														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_1		○W	○M	□				□			□				□
C_2					⊔I	◡L	□	□			□				□
C_3									◡E	◡E	□				□
C_4												⊔I	◡L	□	□
Unrel.	☒														

Reading key:

Symbols for response sources:

○ = internal-explicit
◡ = internal-implicit
□ = external-proximal
⊔ = external-distal

Symbols for prompting types:

W = word-form prompts
M = formal emphasis
I = induction from experience
L = logical conclusion
E = parallel exemplification

The protocol is read in the following way: The subject matter unit C_1 is treated in the didules 2, 3, 4 and later on also in numbers 8, 11 and 15. In didules 2 and 3 the student obtains his answer within the same didule (first with the aid of word-form prompts, then with the aid of formal emphasis). In the other cases with C_1 the student has to rely on memory. Etc.

Comment:
The didule sequence seems to be well constructed. The micro sequences go from fairly easy tasks (with prompting) to more difficult ones (without prompting), non-formal prompting is frequently used, and the sequences contain problems that contribute to the integration of knowledge as final elements (number 8, 11 and 15).

As our scientific knowledge of the self-instructional learning process is gradually increased, the questions studied by means of these check lists, diagrams and protocols will be more specific, and the evaluation will be more securely founded on hard facts. This will probably mean that the careful study of "pheno-structures" by means of aids of these kinds will gradually become more important and more economical. The empirical try-outs with subsequent revisions will probably always remain an important final control process, but the amount of work involved in it will decrease, and at the same time it will be possible and natural to pay increased attention to the post-writing, pre-testing examination.

13. THE PROCEDURE OF SUCCESSIVE EVALUATIONS

13.1 INTRODUCTORY VIEWPOINTS

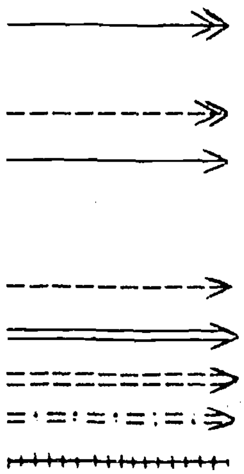
A comparison between different types of instruction (for example, lecture, conventional classroom instruction, group work, private tutoring, and programmed instruction) using a systematic examination of various characteristics (for instance, presence or absence of disturbance effects, the degree of data storage, and the degree of feedback from student system to teacher system) seems to give several points of departure for a discussion of educational effectiveness (cf. Figure 13.1). Especially notable seems to be that the programmed instruction is the only situation which combines (1) continous feedback from student system to educator system with (2) complete data storage.

Among other things this means that the response residuals in the used-up programs give good starting-points for revisions (cf. also Figure 13.2). Usually, therefore, the construction of self-instructional material is a pragmatic procedure in which one considers and acts upon the feedback information. That which functions well is retained, but that which does not function satisfactorily is taken away or changed. Another way of stating the basic characteristic of this way of acting in the system control phase: A program is not really a program until it works. Or even more drastically: One should not be surprised that programmed instruction works. It works per definitionem!

The examination of the actual interaction between student and program is therefore of basic importance. This means, on a small scale, a Copernican revolution in the way of looking at student achievement. Traditionally many people look at student achievement in the first place as a measure of student ability, and those students who do not seem to achieve much are sorted out from the system of the school. In the evaluation of self-instructional materials, on the other hand, student achievement is considered primarily as a measure of the effectiveness of the teaching materials, and instructional units which do not seem to lead to good student achievement are sorted out or replaced. (Of course, probably nobody believes that all students are able to learn everything. But a program evaluation with successive revisions based on feedback information means a serious attempt to adapt the method to the individual, thereby trying to reach more individuals in a more effective way.) Programmed instruction has been one of the major influences behind the increased interest in what is now often called mastery learning (cf. Bloom, 1971).

Figure 13.1 The role of feedback and data storage in various types of instructional systems

Symbol explanations:



= the function of subject-matter presentation (normal educator function)

= partial educator function

= continuous feedback from educand (student) to educator (teacher or machine and/or instructional material with instructor function)

= partial feedback

= continuous data storage [1)]

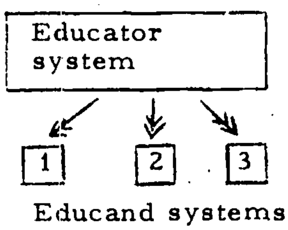
= partial data storage [1)]

= rudimentary data storage [1)]

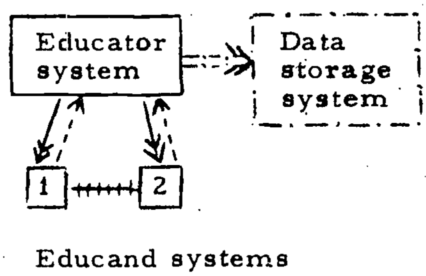
= disturbance effect

1) Quantitative values for what is hereby to be considered "continuous", "partial", or "rudimentary" can hardly be given, but most judges will still be able to agree on the classification of the typical cases over this rough scale.

A. "Lecture"
(No feedback; no data storage via the educator system; no obvious disturbance effects)

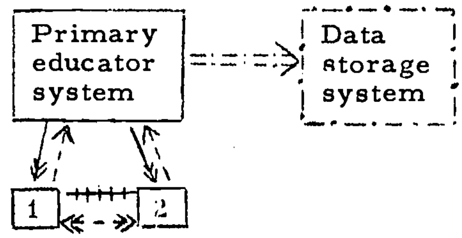


B. "Classroom instruction"
(Partial feedback from educand to educator; rudimentary data storage via the educator system; disturbance effects between the educand systems)



C. "Group work"

(Partial feedback from educand to educator; rudimentary [1] data storage via the educator system; disturbance effects, but also partial educator effects between educand systems)

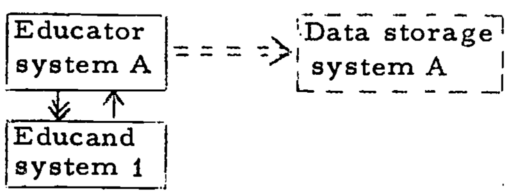


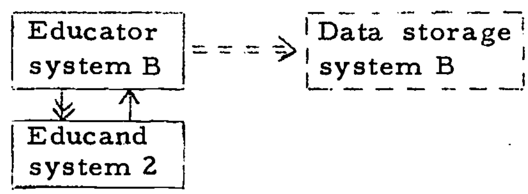
Educand systems = secondary educator systems

1) In certain forms of group work with successive work reports, the term "partial data storage" may be more adequate.

D. "Private tutoring"

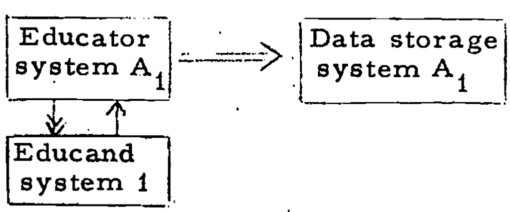
(Continuous feedback from educand to educator; educator systems highly flexible and different from each other in an unsystematic way; partial data storage; no disturbance effects)





E. "Programmed instruction"

(Continuous feedback from educand to educator; educator systems often with no or low flexibility, either totally identical or identical in case of identical educand reactions; total data storage; no disturbance effects)



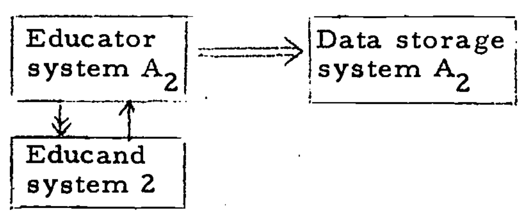
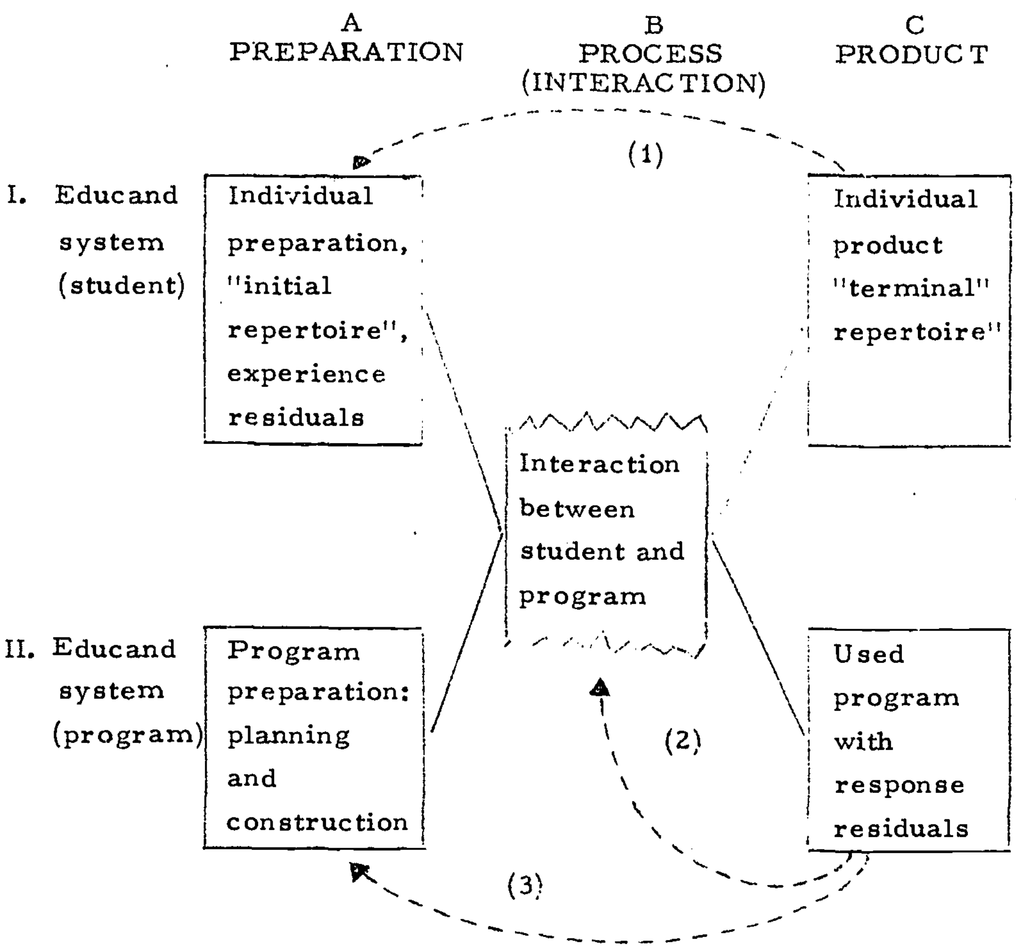


Figure 13.2 The three main phases in the interaction between student and program and three types of feedback effects



Three feedback effects:

(1) "individual effect": addition to or change of the experience residuals of the individual, which means that the initial repertoire is different when confronted with a new educator system
(2) "feedback to the program": in adaptive educator systems the information on the individual's responses influence the program's contribution to the next phase of the interaction
(3) "feedback to the programmer": the response residuals in the used-up program give indications of where and/or how to revise, that is lead to reconstruction in the preparation phase.

13.2 THE MAIN STAGES OF THE EMPIRICAL EVALUATION

Revision work and the final testing can of course be carried out in many different ways. The main principle is an alternating sequence of empirical tests and revisions. The following are some of the most usual work stages.

13.2.1 First Phase: Explorative Testing of Program Sections on Individual Students

The first testing stage is usually considered one of the most exciting moments in programming. The program is tried out on individual students. It is one of those moments when the programmer's self-esteem can receive the most unpleasant jolts, namely, if it is shown that working sequences, on which he has devoted a great deal of intensive work and which he himself believes to be well night perfect, do not in fact work at all. In order not to be kept in suspense for too long many programmers consider it appropriate not to keep back this explorative and individual try-out until the whole program is ready. Instead, it may prove good strategy to test out small sections during the actual work of composition, before going any further.

The students selected for this first try-out should, from a knowledge and intelligence point of view, be fairly typical of the group to whom the program will subsequently be directed, but preferably they should not be too docile and timid. Students who are not afraid to criticize and proclaim their opinion are those who will prove most valuable in this first test stage. The intention here is not that the material should necessarily be gone through under "field conditions". The most important aim is to get as detailed an idea as possible of the student's way of reacting to the details of the program. For this reason the try-out will preferably take place with one student at a time. As a preliminary, one ought - possibly with special tests - to make certain that the student does not know the material in advance. One should also make clear to the student that the try-out applies to the study material not to him. Such a frank declaration usually creates favorable conditions for free and untrammelled communication without unnecessary anxiety. A quiet testing room and plenty of time are also recommended.

The first try-out is often made orally with student comments for each didule and with a detailed follow-up interview. However much the programmer values his own program, he must take care not to be on the defensive during these interviews. His task is not to defend himself, but to collect reaction data. Should we begin to argue, we risk the student becoming unwilling to express his points of view. If we are dealing with adult students, the comments can sometimes be given in writing. The student, for instance, is given the instructional units in the shape of a collection of cards, and he writes his responses on a response-sheet, on which there is also space for marginal side-comments (e.g., a sheet with two columns, one for responses and one for notes). Moreover, the student can possibly fill in a special questionnaire. As soon as the instruction has been tested with one or more students, the first revisions can be made. Ambiguities are removed, better examples are chosen, illustrations are made clearer, extra units are added, and so forth. At this stage the number of students is less important than the detailed analysis of how well the communication "penetrates". A small number of carefull; studied students is often a greater value than a larger number whose reactions are more superficially studied.

13.2.2 Second Phase: Successive Testing of the Entire Program on Larger Groups of Students

When we consider that most of the obvious obstacles have been cleared away, the testing of the total program is begun under normal educational conditions (and with presentation aids intended for use in the field). A sufficiently large group is used so that a certain amount of fairly reliable information can be obtained concerning, for instance, the percentage of mistakes in the individual instructional units. Pre- and posttests of knowledge are carried out in order to assess the total learning effect. The program is revised on the basis of analysis of mistakes within the program and the terminal effects. Individual units are discarded, additional tasks are inserted, the sequence of units is altered. Perhaps it is shown that extra branchings are necessary. After revision a new try-out on a new group of students is made. The pro procedure can be repeated several times. In principle, the group try-outs should be continued, until those values of error data and, above all, of terminal performance data have been attained that are considered desirable and agree with the objectives of the course. In order to forestall confusion and, in addition, in order not to give a false impression of having a fully adequate material, the program versions, which are used in these preliminary try-outs, should have suitably unambiguous headings such as "Experimental edition 2", "Experimental edition 3" or something similar. (A more detailed discussion of evaluation criteria will be given below.)

13.2.3 Third Phase: T·rminal Field Testing

When revisions are no longer considered necessary, the time is ripe for the final field testing. In this the demands on the representativeness of the student groups and on the "normality" of the field conditions are more severe. Exhaustive data on pre-knowledge, intelligence, time spent and final results are collected. Pre-knowledge and final results are often measured best by parallel tests. In view of the student being able to learn something also from the pre-test, and taking into consideration that the student can learn a great deal from the study material without on that account attaining the same level of knowledge as a student taught by another material, it is often desirable to use the pre- and post-tests, for the sake of comparison, also in a control group (with a corresponding student group, but taught in a different manner). Another reason to use a control group for the pre-test is that this test can have a certain motivating effect, in that the student gains advanc. information about his own future possibilities after the end of the course. (A comparison between the first and the third testing stages is given in Box 13.1; an exampl e of the planning schedule is shown in Box 13.2.)

After this field test no alteration of the program should occur. Data of this test's conditions and results will, namely, constitute an essential part of the basic information about the program's character for teachers and administrators. This information should normally be published in the form of a special "program manual", which can be said to be the program producer's "declaration of ingredients" for potential program consumers.

Box 13.1 A comparison between the first and third main stage in the successive try-out work

	First Stage	Third Stage
Aims	Explorative tests of the ability of the program to communicate and teach	Collection of data for final "declaration of ingredients" for program consumers
Time	Several times during developmental work	Not until entire program is in terminal state
Main responsibility	Program constructor	Expert of investigation Methods in behavioral sciences
Test persons	A small number of persons typical of the target population	A larger number of persons, an unbiased sample of the target population
Test conditions	Fairly informal with good prospects for relaxed communication and intensive observation	Strictly reproduceable conditions, demands for detailed information re student characteristics, demands for the situation to correspond to the conditions under which the program will be used
Follow-up measures	Revision of program	Tabulation and analysis of data re program's effect; no revision

Box 13.2 Planning and realization schedule: An example

Note: Preparation, construction and follow-up work demand exhaustive planning. Here a planning schedule can be of great use. It should- in addition to the designations for the different work stages - include the names of those responsible for the different work stages as well as the date when the stage's work should be finished (PT = planned time), and when it actually was finished (AT = actual time). The latter details can give a sounder ground-work for the planning of the next programming project.

Project	Project leader		
Task:	Responsible:	PT	AT
1. Main direction of goal analysis is established			
2. Goal analysis, including detailed terminal descriptions and terminal tests			
3. Student group identified and analyzed			
4. Enlarged terminal chart and modified tests			
5. Situation analysis			
6. Subject-matter analysis: logical structuring			
7. Subject-matter analysis: psychological supplementation			
8. Choice of presentation media			
9. Points 1-8 are coordinated and tabulated in the form of detailed directives to the program author, supplemented with T-catalogue, C-catalogue, etc.			

Task:	Responsible:	PT	AT
10. Examination of already existing didactic material. Possible decision on project alteration.			
11. Preparatory sequential decisions: Choice of "flood model"			
12. Preparatory subject-matter structuring, possibly in the form of a didactoplan			
13. Version 1 is completed			
14. Version 1 is re-examined			
15. Version 1 is revised			
16. Version 2 is subjected to preliminary try-out with individual observations			
17. Version 2 is revised			
18. Decision on further preliminary try-outs			
19. Version x is group-tested			
20. Version x is revised			
21. Decision on further group tests for revision purposes			

Task:	Responsible:	PT	AT
22. Editorial final examination			
23. Printing (or the like) of final version			
24. Field testing			
25. Writing of program manual with data from field testing			

14. EVALUATION CRITERIA AND EVALUATION AIDS: SOME EXAMPLES OF RELEVANT VARIABLES AND SYSTEMATIC MAPPING DEVICES

14. 1 INTRODUCTORY POINTS OF VIEW: DIFFERENT TYPES OF EDUCATIONAL MEASUREMENT

In general, educational measurements can be made of conditions before, during, and/or after a particular instructional sequence, and they can aim at an evaluation both of the individual and of the instructional system (as is outlined in Figure 14. 2).

In the past, work on educational measurement has above all dealt with problems concerning cell 3 in our figure: how one constructs measurement instruments that give as reliable and valid a picture as possible of the final competence of the individual, mostly in comparison to some "norm group". This is, of course, an important aspect. However, other measurement aspects have by comparison been underdeveloped. First, it should be of interest not only to assess the status of the individual but also to evaluate the instructional system he is confronted with. How good are our systems (constructed to bring forth an optimal sequence of behavioral changes in accordance with given goal descriptions among the largest possible number of relevant students)? Second, it can well be said that so far we have been too often satisfied to obtain evaluations afterwards only (cell 3 and 6), in which case we primarily get comparisons between individuals or comparisons between systems, while a combination of measurements during and after the instruction (2+3 or 5+6) would often give more "useful" information. By useful information we then mean such data that directly serve to localize weak links in the instructional system: single parts that should be improved.

Programmed instruction is one of the fields of recent educational interest that have served to turn the interest of educational measurement specialists in new directions and broaden the focus. When successive revisions of the material are made, one is guided both by the behavior of the students in the instructional situation (to what extent the student succeeds and how he reacts to single didules) and by their terminal achievements (their ability to reach the goals as measured by terminal tests).

	Before instruction	During instruction	After instruction
Focusing on a single individual	(1) Establishment of initial level	(2) Stepwise testing for successive feedback, e.g., to guide choice of instructional alternatives	(3) Establishment of terminal level; report on "competence"
Focusing on the "system"	(4) Studying relationship between actual student level and difficulty level of the system	(5) Localization of defects in function; control of relevance of branching system	(6) Assessment of total effectiveness and efficiency of system in relation to goal catalogue

Figure 14.1 General examples of various types of educational measurements

14. 2 TERMINAL VARIABLES: FINAL ACHIEVEMENTS

The terminal achievements are naturally enough the most important indicators. If we get good final results, then we need not bother so much about the previous behavior during the instructional situation. However, should we obtain a less satisfactory final achievement, then this is an important indicator that something is not in order, but this piece of information usually gives the programmer rather little specific guidance on where and how improvements should be made. In this case, then, study of the didule behavior (correct and incorrect reactions on single instructional units) is far more informative.

The character of the final tests is obviously very important. If the test is unsatisfactory (unreliable or not representative as far as the goal of the course is concerned), then we are severely misguieded when we try to assess the effectivity of our study material. The final tests should therefore always be worked out in close cooperation with experts in psychological and educational testing and with reasonable attention paid to usual measurement criteria (including demonstrations of reliability and validity).

Since there are many books on these topics, we will not enter into any description of these general rules here, but refer any reader who feels that he does not know enough about these matters to common text-books on educational measurements.

Besides these general criteria, we would like to emphasize three desirable characteristics of our measurement instruments. They should:

(a) be so designed and so related to the specific instructional system that the results obtained in using them will have clear and meaningful consequences for further planning;

(b) have a broad coverage, rather than be one-sidedly perfect;

(c) be clearly goal-related, rather than content-related or norm-related.

The first condition mentioned was already touched upon above in our brief discussion on Figure 14. 1, and we will return to concrete examples of this below. The second point may seem self-evident, but nevertheless is sometimes too easily forgotten. One should keep in mind the total goal area of a course, so that one does not rest satisfied with too narrow a test instrument. The temptation to do this often derives from the fact that certain behaviors are more easily tested than others. If, for example, retention over longer periods of time or a certain kind of transfer effect are important parts of the goal area, one should try to assess these things, even though they may be difficult to measure.

Our measurement techniques are comparatively advanced as long as we deal with the cognitive field. Nevertheless, there is a risk that we rest satisfied with tests dealing primarily with simple recall of separate items of knowledge (since such tests are so easy to construct), in spite of the fact that we really want to cover more complicated abilities (such as application, analysis, or synthesis of knowledge items). In order to help us remember to cover those aspects that we intend to cover it may often be worth while to use test construction matrix of the kind presented in Figure 14. 2. In the headings of the columns the various cognitive functions we want to deal with are entered (for example, using the terms of the well-known taxonomy of the Bloom group; Bloom, 1956), while the rows represent the various subject-matter areas. Such a simple cell system may be a good aid for the test constructor in not forgetting any important type of test item. (Cf., for example, Payne, 1968.) This does not mean, of course, that it is always easy to construct items which correspond to all the cells of such a matrix. We may find it especially difficult when we are interested in including not only cognitive, but also non-cognitive functions. In spite of attempts to construct taxonomies in non-cognitive fields also (see, e. g. Krathwohl, 1964), we still have very few good test instruments available in these areas.

In our third and last point on desirable characteristics we touched upon the frame of reference of the evaluation or the critera used for comparison when we make our assessment. We may distinguish three main categories of such assessments:

(a) Goal-oriented assessments. The assessment of the instructional result in then related to a specific series of goal descriptions. ("Out of the available test items, designed to cover the demands of our goal descriptions, Student S_1 has been able to pass 90 %.")

(b) Process-oriented (or content-related) assessments. The assessment of the instructional result is related to some specification of what has been presented during a course. ("Out of the available test items, designed to be a representative sample of the content of the course - as this is seen from a particular book or summary of lectures, Student S_2 has been able to pass 89 %.")

(c) Product-group oriented (or norm-related) assessments. The assessment of the instructional result is related to information about results of other individuals. ("On the tests used Student S_3 has obtained

	Behavioral dimension:						
Content areas:	Knowledge	Interpretation	Application	Analysis	Synthesis	Evaluation	Sum
Subject-matter area 1							
Subject-matter area 2							
Subject-matter area 3							
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Subject-matter area n							
Sum							

Figure 14.2 Test construction matrix as an aid in avoiding one-sidedness in item choice

a result which places him in the best 25 %. ")

All three methods can give valuable information in some situations - and sometimes, in addition, rather similar information. In other cases, however, the similarities are less marked, and then it is reasonable to ask which of the methods is most meaningful for one's particular aim. It might first be noted that the third type of assessment is most clearly different from the other two in being more relative in character. Thus, we may talk about "absolute" or "criterion-related" assessment on the one hand (and to this group we may then refer both the a and the b type) vs. "relative" or "norm-related" assessments on the other (the c type; cf. for example the dichotomy in Glaser, 1963).

Earlier educational measurement was usually dominated by norm-related measurements. Then the main aim was to differentiate among individuals, for example, in order to make individual predictions of future study success. Test results are thus mainly seen as prognostic instruments (we predict from cell 3 - cf. Figure 14.1 - at time 1 to cell 3 at time 2 etc.). But for many other purposes, especially when we are interested in the effectiveness of a particular instructional material, this type of information is not specific enough. It is not very info ative to know that a certain person has passed through a s..dy material A and done somewhat better than the average student. It would often be much more interesting to know <u>what</u> he is able to do.

If this is what we want, the two other criterion-related as..e...en , type a and b, come into focus. The difference between them is sometimes not very large, and the distinction is often overlooked. If the particular course has started from and closely kept to a given goal description, then the difference should usually not be of importance, (if the main goal is not quite outside the scope of the subject-matter; there may be courses, where the subject-matter is more or less irrelevant, but where the goal is to teach the students a particular study technique, for example; this may not explicity be seen from the particular course material used). The reason why it is important to distinguish between type a and type b is, however, primarily that many courses lack a clear relation between a specified goal description and a closely related course content. It may be easy and tempting to choose a textbook of a course and"make a test on it" without pondering upon the relation of the book to a given goal description or a goal conception behind it. Then, however, we will lose our possibility of detecting defects in the goal-relatedness of the course.

In general then, goal-oriented assessments are in many cases to be preferred to the other two types: they are definitively more informative than the norm-related ones, and they are often more relevant than the process-oriented ones. When we deal with programmed instruction, it is quite obvious that we should usually prefer to work with goal-oriented assessments, very closely connected with the terminal analyses started from. This statement may seem rather self-evident, but it has certain consequences that may not be so obvious and even seem a little unusual to a person used to the routines of educational measurements.

The fact is that some of the measurement routines that have been built up are closely related to the aims of norm-related measurements, and that these are not equally meaningful in the construction of goal-related instruments. For instance, when we deal with norm-related tests, it has often been natural to increase the discriminatory power of the items by screening away such test items that show an especially high or especially low frequency of correct solutions. If one has the general aim of obtaining a clear dispersion in one's group, this is a natural procedure. In a goal-related test, however, one's main demand is that a test item should be clearly relevant to a specific sub-goal of the goal catalogue, and that the test item should differentiate between those who, in a terminal situation, are able to reach this particular sub-goal and those who are not. Let us suppose that we have a test item which fulfils the demands just mentioned. Should such an item get a very high frequency of correct solutions, then this cannot be an arguemnt to remove it: the high frequency shows only that the instruction has been effective in this respect. Should the item get a low frequency of correct solutions, on the other hand, then this fact cannot be a reason for screening it out either: this may rather be an important indicator that the instruction on this point has to be improved. The discriminatory power of the items in the particular product group is thus not an adequate basis for screening out items, when we deal with goal-oriented tests; it is more important that the items discriminate between groups about which one knows, from other sources of information, that they have/ do not have abilities which correspond to the goal description (cf. Glaser, 1963 and Marton, 1969).

In general, then, we do not choose test items with general predictive symptom value, but items which are closely related to specific goal descriptions; and we do not necessarily weed out items which most students are able to solve after going through the course.

14. 3 MEDIAL VARIABLES: BEHAVIOR IN THE INSTRUCTIONAL SITUATION

The behavior in the instructional situation is studied in detail during the explorative, individual try-out phase and from certain points of view (for example, with respect to correct vs. incorrect response behavior on individual instructional units) also during the group try-outs and in the final field study.

14. 3. 1 Medial variables of achievement character: Analysis of correct and incorrect responses during the work with the program

Among the most common aspects of the examination of student behavior in working with programmed materials, is error analysis. This is very reasonable both from a common sense point of view and from the viewpoint of Skinnerian theory. The risk is that the programmer stresses the counting of errors more than the analysis of errors.

Instructional micro-units ("didules") that have caused many errors should not be discarded without further investigation (in some cases and for certain purposes a somewhat higher percentage than usual may be acceptable and educationally meaningful), but these error-causing units should always be carefully examined, so that the programmer can come to as realistic conclusions as possible about the reasons for the increased error frequency. Some of the most common reasons are, of course, that the response demand has been ambiguous, that the language has been poorly adapted to the particular student group, or that the background behavior, built up earlier in the program, has not been trained to a sufficiently high level of stability and precision to ensure a correct solution in the specific task under study. In all these cases a revision seems to be the appropriate next step for the programmer.

Even though the study of errors is important, we should never forget that the error reaction alone doe not necessarily tell us very much about the effectiveness of the specific didule where the error occurred. One of the reasons is, of course, that the correct answers are usually given as feedback. We have no simple one-to-one correspondence between the dichotomy "correct versus incorrect response" on the one hand, and the dichotomy "understanding of the task versus lack of understanding" on the other hand.

During the first individual try-out phase, at least, it may therefore be advisable (Gordon, 1963) to have the students mark off the units that they do not quite understand (whether they have answered the questions involved correctly or not). . A simple question mark in the margin of the text or on a special sheet of comments would usually be adequate. This gives us four categories of student reactions: (1) "error with remaining doubt", (2) "error without remaining doubt", (3) "no error, but doubt", and (4) "neither error nor doubt". It is obvious that didules of the first type have to be revised. On the other hand, the second type may sometimes be educationally fruitful. (Once in a while the most effective instruction may result from the student's being allowed to make an incorrect response and having it clearly classified as wrong; this may be the case when we are dealing with a "spontaneous error tendency", caused by strong interference from earlier response patterns.) The main advantage of this analysis probably lies in the fact that we can identify those units where the student have answered correctly, but where they nevertheless register lack of understanding (the third category above). In these cases the programmer may, for instance, have used too strong prompting techniques, a fault that a simple error analysis unfortunately cannot demonstrate. Such units should usually be revised, too.

This kind of analysis is useful not only for linear programs, but also for the error-treatment program of the scrambled book type. The lack of correspondence between understanding and correct answers may be quite frequent in some of these programs, if no extra checks are made. It is quite possible that the student chooses the correct answer immediately in a multiple-choice situation, without understanding the correct answer (the other alternatives may be too obviously wrong, for instance). There is also no guarantee that the final correct answer given by the student is connected with understanding. It may simple be the only alternative left after the student has passed through all the error branches (without having understood enough of the error treatment given).

If we try our program on one student at a time and simply extend the number of steps at each point where errors occur, there is a considerable risk that we end up with a very large (and inefficient) program. On the other hand, if we feel content with overall error percentage (making no revisions at all if these percentages fall within certain preestablished limits), we do not fully utilize the information collected. The first approach places too much weight on the behavior of single individuals which may or may not be atypical (and is too much influenced by chance), whereas the second approach gives us too little information about individual items. Instead, a systematic mapping of error distributions and error types would usually be preferable, and as a rule this mapping should also include subject reports on understanding. A natural starting point is then a simple survey table, including both errors and reports of lack of understanding. A square-ruled background grid with students as row headings and didule numbers as column headings would be a simple, but adequate design (cf. Figure 14. 3). This survey table should be supplemented with a frequency list of deviant responses.

These aids are good starting-points when the programmer has to answer questions of the following types: (1) Do the errors concentrate on certain specific didules ("column comparison")? (2) Is there any good educational reason for this, or is it simply an indication that these particular units or some immediately preceding units should be revised? (3) Do the errors concentrate on specific individuals ("row comparison")? (4) Are these errors of such a systematic character that some preparatory additional course (remedial branching after a filtering test) would probably reduce their number considerably? (5) Do the same error types occur repeatedly for some individuals? (?) Could this best be avoided by more complete information in the main program or by utilizing special branching for remedial instruction?

For a summary discussion on the relation between item difficulty and learning effect we may refer to Jacobs (1963). The main thesis of Jacobs is that many variables influence the degree of difficulty without influencing the learning effect and vice versa, and that it may therefore be misleading to concentrate the study onesidedly on item difficulty (in terms of error frequency) as the only evaluation criterion. Our examples above point in the same direction.

14. 3. 2 Medial Variables of Experience Character: Analyses of Attitude Development and the Degree of Subjective Difficulty

In addition to such evaluation variables that deal with the final learning result (the terminal behavior) and such that deal with correct vs. incorrect behavior during the study process, it is often also desirable to examine the development of attitudes in the instruction situation (or, to use another word, the general "climate of work"). Successive observations of the study behavior during the course as well as direct attitude questioning at certain points of time are natural methodological approaches. Among questions in which one would usually be interested in such studies are: How does the student experience the self-instructional work situation? Is this felt to be more or less stimulating than the

Units / Students	U_1	U_2	U_3	U_4	U_5	U_6	. . .	U_n
S_1			E ?		?		. . .	
S_2	E ?	E	E		?		. . .	E
S_3			E ?				. . .	
S_4			E		?		. . .	
S_5			E ?				. . .	
S_6							. . .	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S_n			E		?		. . .	

Reading key:

E = error
? = lack of understanding

Figure 14.3 Typical structure of a simple survey table ("Error-and-doubt distribution")

In this simple hypothetical example we find a clear error concentration on unit 3 on the one hand (for almost all subjects) and for Subject 2 on the other hand (for almost all units). We also find a relatively large number of question marks for unit 5 in spite of the fact that this unit has not attracted any errors at all. These kinds of information are obviously better starting-points for revision than overall error percentages for total programs. For more detailed discussion, cf. the text.

traditional teaching the student is used to? And if it should be experienced as more stimulating: Is this only a temporary phenomenon, some kind of novelty effect, or is it more stable, observable during a longer period? What attitudes to the particular subject-matter do we find under the particular conditions of programmed instruction? Can we find any influence on the students' attitudes towards studying in general?

In the debate on programmed instruction, particularly during the early years, many participants seemed to have clearcut and general answers to such questions. Unfortunately, however, they were often more emphatic and certain than unanimous. Probably such questions can only to a very limited extent be given general answers. The answer for a given situation will depend both on the particular study course under work (there are stimulating and dull programs just as there are stimulating and dull textbooks) and on the other experiences of the particular student group (the "comparison" criteria). It should be no hindrance to our study of attitudes, however, that the answers only seldom can be generalized. On the contrary: it will be so much more important to evaluate the specific course material, as experienced by a particular target group, not only from the point of view of achievement but also from the point of view of attitude development.

Side by side with the attitude variables proper, it may sometimes be of value to study other "experience variables", especially the "subjective difficulty level" of the items. When constructing tests one commonly define the difficulty of an item in terms of frequency of correct, solution, but difficulty defined in this way is obviously not meaningful in the individual case. Let us suppose that the calculating tasks (1) 2 x 21 = ; (2) 3 x 34 = ; (3) 11 x 13 = ; and (4) 145 x 112 =; for the four students A, B, C and D result in the following distribution of correct and incorrect solutions:

	1	2	3	4
A	C	C	C	C
B	C	C	C	E
C	C	C	E	E
D	C	E	E	E

C = correct solution
E = error

We note that Task 1 is solved correctly by all. Hence, it has the p-value of 1.00 and is then, using common test-construction terminology, an "easy" task. But is it equally easy for student D as for student A ? If we define difficulty level in terms of frequency of solution, the answer is 'yes' - or rather: then the question is meaningless. But very probably the degree of "subjective difficulty" is greater for student D than for student A. We might get some idea about whether or not we are correct in making this guess, by examining the amount of time for solution and/or by means of direct student ratings on difficulty, and it should be of interest to relate such variables to both achievement variables and common attitude variables. (Our discussion of this point follows ideas in Jacobs, 1963).

14. 3. 3 Special Aids: Didactograms

If the program is non-linear (for instance, using parallel tracks, remedial loops, by-pass arrangements, etc.), the simple survey table is sometimes not the best possible starting-point for a study of error frequencies on various items. There are two main reasons for this: (1) Error frequencies should usually be considered in close connection with their position within the branching system, and (2) we also need information on frequency of use for the various tracks or branches in the system. Hence, in these cases, it is often useful to draw special diagrams.

The basic outline of such a diagram will be stable for a particular program version, irrespective of the student reactions, and will show the sequence of the didules, including all available branching possibilities. As starting-point for these diagrams it is often suitable to use the "didactoplan" constructed earlier during the preliminary decisions on sequence (cf. above). On this background diagram we may fill in information about the individual's actual working route as well as his working results on the single units (correct and incorrect responses symbolized, for instance, by means of plus and minus signs). Such a program outline, filled in with student behavior registrations, might be called a "didactogram", and it gives us a relatively surveyable map of the interaction between the instructional system of the program and an individual (cf. Figure 14. 4).

Data from several individual diagrams can be brought together in a "group diagram", on which information about the percentage of correct solutions for each single response demand, as well as information about the number of individuals using various possible tracks within the branching system, can be registered (cf. Figure 14. 5).

Such a diagram can be used by the programmer for many different kinds of decisions. To take only one example; if may often be shown that some of the branching possibilities have not been used at all. If so, it is natural task of revision to get rid of these tracks or give them some other form. Such an analysis is important, for instance, when we work with multiple-choice arrangements and subsequent error instruction connected with the various possible error choices. Perhaps some of the error choices never attract the attention of the students, and the special error instruction included within the program is then dead wood. Unfortunately there is seldom any evidence that the productive authors of scrambled books have examined their often very bulky error instruction in this way.

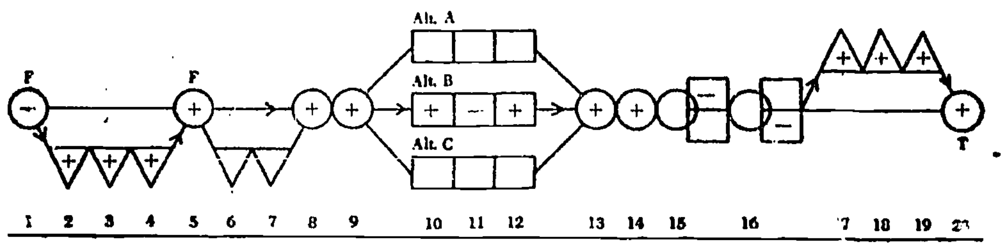


Figure 14.4 A diagram of branching use and response correctness for a single student ("individual didactogram"): An example

Background symbols:

(The background structure is the same in all diagrams representing the same course version)

F = filter function

T = terminal testing function

○ = a didule that all students should go through

▽ = a didule in a special, additional course with remedial function

△ = a didule in a special, additional course containing optional "enrichment"

□ = a didule in a parallel branching system

○□ = a didule with connected error treatment (two error alternatives)

Working symbols (individual data):

→ = the route followed
\+ = correct solution
– = incorrect solution

The diagram is read in the following way:

The subject gives a wrong answer to the introductory filter unit (no. 1) and is then directed to go through the special remedial items 2-4, the response demands of which he answers correctly. Since he also gives a correct response to unit 5, he does not need any further remedial treatment, but is allowed to proceed directly to no. 8. Among the possible parallel tracks (covering units 10-12) he chooses the one called Alternative B. He makes one error choice in each of the multiple-choice tasks (15 and 16). He chooses to go through the optional enrichment sequence (17-19). After that he solves the terminal test item correctly.

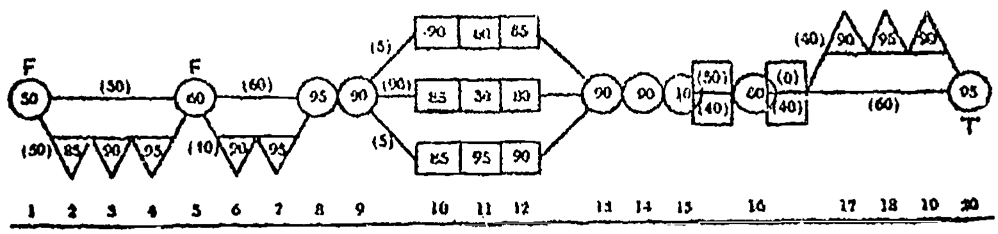


Figure 14. 5 A group diagram of branching use and response correctness ("group didactogram"): An example

Background symbols: See Figure 14. 4

Working symbols: Figures without parenthesis give the percentage of correct responses ("frequency of correctness"). Figures with parenthesis give the percentage of students who have used a specific passage ("frequency of use").

The diagram is read in the following way: 50 per cent. of the students answer unit 1 correctly, and these students are directed to go forward directly to number 5. The rest of the students go through the remedial items 2-4 (in which the resulting correctness percentages are respectively 85, 90 and 95 per cent.) On item 5, correct answers are given by 60 per cent of the students; these students are allowed to go directly to number 8. The remaining 40 per cent. proceed through numbers 6 and 7, etc.

Comments: The correctness percentage is fairly high for most of the units. An exception is number 11 in Alternative B; this unit should therefore be given special attention. The units with filter function (i and 5) have a moderate correctness percentage, which usually can be considered to be adequate (if all students should answer a filter unit correctly, the appended remedial sequ nce would be unnecessary). The error-treatment units 15 and 16 give uneven values (D 15 has few correct answers, and D 16 has uneven distribution between error alternatives) and should therefore be especially examined. Finally, it is, of course, a good sign that the terminal unit with testing function (20) has a high correctness percentage. - Frequency of use: The optional choice of parallel sequences gives an uneven distribution (almost all students choose Alternative B). Here revision is probably indicated. The distribution in the choice of enrichment courses, however, does not seem to give any reason for revision.

14.4 ANALYZING RELATIONSHIPS BETWEEN STUDENT BEHAVIOR WITHIN AND OUTSIDE THE PROGRAM

The final criterion of the effectiveness of a program is, of course, the student's terminal behavior. As was stated above: If this reaches acceptable levels within reasonable time periods, we need not be so concerned about the details of the behavior within the program. However, when the terminal behavior is not fully adequate or when the time used seems too long, an analysis of the student's behavior within the program is needed to show us the place where revision is most indicated. About these general ideas there seems now to be a fair degree of agreement among programmers. However, two deficiencies stand out in many reports on program development. Sometimes the programmers are so exclusively interested in the error percentages within the program that all other considerations, even those concerning the terminal behavior, seem to be forgotten. In other cases attention is paid both to final tests and to the error frequencies within the program, but the programmer does not care to study any particular relationships between student behaviors within and outside the program, even when these relationships should give him important additional information (such as, for instance, would often be the case in branching programs). In the following we will give a few brief illustrations of relationships between student behaviors within the program and outside the program that should be of interest to the programmer.

14.4.1 Information Density and Information Speed of a Program

As has been pointed out (Björkman, 1963) the information value of a unit of information can be defined as the difference between the probability of correct answer before and after the communication. Since the single units (didules) of a program often do not contain complete terminal reactions, however, but have other purposes as steps on the way towards such terminal reactions, it is often unrewarding to study the information value (in this sense) of the single units. (This value may be maximal in the stated operational meaning without any gurantee that we have come closer to any of the terminal reactions.) On the other hand, it should be quite meaningful to study the special tasks, sometimes called "criterion problems", which test the student's terminal behavior after he has gone through an instructional sequence of related units. Even when formally built into the program proper, these criterion problems are equivalent to test items and may be considered to represent "behavior outside the program" in the sense intended here. If we examine the ability of the students to solve these criterion problems before going through the instructional sequence (error percentage A) as well as their ability to solve them in their natural context, that is, after going through the sequence (error percentage B), we shall obtain more meaningful measurements of "uncertainty reduction" (error percentage A - error percentage B).

Let us assume that, for each new concepts we introduce we have at least one instructional sequence, and at least one criterion problem. The uncertainty reduction of the cri. ion problem divided by the number of units (didules) included in the sequence then gives us a measure of the "information density" of that particular sequence. The information density of the total program may then be defined, in principle, as an average of the density of the separate micro sequences:

$$\frac{\frac{A_1-B_1}{D_1}+\frac{A_2-B_2}{D_2}+\frac{A_3-B_3}{D_3}+\cdots+\frac{A_n-B_n}{D_n}}{n}$$

A_i = the error percentage of the student group under examination on criterion problem i before the relevant instructional material has been studied.

B_i = the error percentage of the student group under examination on criterion problem i after going thorugh the instructional material.

D_i = number of units (didules) in the micro sequence that leads up to criterion problem i (and the purpose of which is to teach the student the concept C_i).

C_i = basic concept (represented in the program by the i^{th} criterion problem and the i^{th} instructional sequence).

n = number of basic concepts included in the program.

Note. Obviously a formula of this type only gives one aspect of the complex phenomenon of information density as intuitively conceptualized. For instance, it does not consider the fact that various conceptual units can make widely differing demands on the "intermediary processes". If these demands could be expressed quantitatively (and average judgments from a group of raters might perhaps be used in this connection - at least for a first appr oximation), a weighting of each quotient $(A_i-B_i)/D_i$ with respect to the demands on the intermediary processes might "improve" the formula, that is, mean that one more possible criterion of step size has been taken into consideration.

If we exchange the information about the number of units (didules) in the micro sequences (D_1, D_2, D_3, ... D_n) against the time the micro sequences take (T_1, T_2, T_3, ... T_n), we would obtain a measure of "information speed" instead of a measure of "information density". Since one of the most basic criteria of educational effectiveness is how much information the student absorb per unit of time (not how much information has been poured out from the mouth of the teacher or otherwise been potentially available for absorption!), such a formula would give us one starting-point in our assessment of the educational effectiveness of the program. In any case it ought to be of interest to be able to compare this measure with other data about a program. It is hoped that comparisons between different operational program characteristics will gradually give us a clearer picture of what is essential and what is less essential in program construction.

14.4.2 Choice-Point Relationships (Survey Tables of the Distribution of Achievement Patterns in Filter Units)

The measure of "information speed" just mentioned combines information about time used (a "medial" variable) with information about probabilities of correct responses before and after the study ("initial" and "terminal" variables). Another combined evaluation criterion will be of interest when examining branching programs where the students, on the basis of their achievement at a choice-point (in a filter unit), are led different ways through the program. Simple error frequency data for these filter units are, of course, insufficient as a basis for the revision, since we are not here concerned merely with minimizing the number of errors. Instead, in these cases the aim is to lead the student, by means of the diagnostic indications given by the filter unit, to the particular study material that is most appropriate to him. This means that the student's achievement on initial test (pre-test), filter unit, and terminal test have to be examined together. It is obvious that certain combinations ("achievement patterns") are acceptable, whereas others show faults in the program and, hence, indicate where revisions are needed. (A demonstration of this was given by Markle, 1963. Some of Markle's ideas will be incorporated in the following illustration.)

As an aid in this examination the programmer can utilize a survey table of the general kind demonstrated in Figure 14.6. The eight possible "achievement patterns" are placed as row headings, and the filter units the program as column headings. The distribution of the students over the various achievement patterns is then built up within the table by entering tally marks.

Let us take a brief look at some typical achievement patterns when the program contains filter units of the general type that sort the students into two sub-categories at each choice-point, one of which passes through only part of the total number of units included in the total programs. This formulation is a very general one, describing a formal branching structure, and it therefore covers at least two main situations that the educator will often consider to be quite different. In the first of these, the main group of students goes through all units, while a smaller group of students is found at the choice-point to know already some of the material that will be treated and is therefore directed to skip some part of the program ("shortening of learning sequence for high-achievement group"). In the other situation, the main group of students goes through the shorter sequence, while a smaller group of students is found at the choice-points not to know enough about materials covered in some earlier program and is therefore directed to go through an additional "remedial loop" ("lengthening of learning sequence for low-achievement group"). The pattern symbols used in the following are explained in Figure 14. 6. Only some of the patterns included in the figure will be commented upon here, and, for the sake of brevity, only faults of the program proper are mentioned (that is, we assume that pre- and post-tests are reliable and valid):

A. I_r FD_r T_r

In this case the student answers correctly (r) both in the filter unit (FD) within the program and in the corresponding parts of initial test (I) and terminal test (T). Students with this pattern will skip the special part of the program immediately following the particular filter unit. In other words, they go on the quicker track (whether this is considered to be the main track or a track for selected minority of high-achievers). No criticism can, of course, be directed against the fact that this pattern appears. One indication for revision may come up, however. If the number of students with this pattern approach 100 per cent in a representative sample of the target population, the additional part of the program in superfluous.

C. I_r FD_w T_r

If the tests are assumed to be valid (an assumption we made above for brevity of discussion), this particular pattern indicates that we have

Achievement patterns:	FD_1	FD_2	FD_3		FD_i
A. $I_r\ FD_r\ T_r$				. .	
B. $I_r\ FD_r\ T_w$				. . .	
C. $I_r\ FD_w\ T_r$				. . .	
D. $I_r\ FD_w\ T_w$				. . .	
E. $I_w\ FD_r\ T_r$				. . .	
F. $I_w\ FD_r\ T_w$				. . .	
G. $I_w\ FD_w\ T_r$				. . .	
H. $I_w\ FD_w\ T_w$				. . .	

Reading key:

I = initial test
FD = filter unit (filter didule)
T = terminal test
r = right
w = wrong

Figure 14. 6 Possible table arrangement in analyzing the distribution of achievement patterns in filter units ("choice-point relations table")

had an inadequate filter unit which directed the student through an extra part of study material that he did not need. The response demands in the filter unit may have been formulated in such a way as to lead the student astray, or the filter unit may simply be too difficult for its purpose. Hence, in this case the revision should focus on the filter unit.

F. I_w FD_r T_w

This combination can be an important indicator of insensibility on the part of the filter unit. The student who does not know the pertinent facts from the beginning is not correctly sorted by the filter unit and, hence, does not get the additional study he needs. This may happen quite often, for instance, if the filter unit uses multiple-choice techniques with less adequate alternatives (or if, in the filter unit, the student is tested on discrimination behavior only, whereas he is tested on construction behavior in the terminal tests). As was the case in pattern C, the revision should be directed at the filter unit, but in pattern F the revision has the opposite purpose of <u>in</u>creasing the difficulty of the filter problem.

G. I_w FD_w T_r

This combination is usually a good sign. The student shows error in the initial test and in the filter unit, but he can solve the terminal problem after the additional training supplied by the extra part of the program. One possibility of revision should, however, be kept in mind. The programmer should ask himself whether or not the "additional" program part ought to be incorporated within the main part of the program. If the number of students with G patterns approaches 100 per cent. in a representative sample, such an incorporation is clearly indicated.

The tabular and diagrammatic arrangements demonstrated in the present section should be considered only as examples. Many other types are possible and suitable to aid the programmer in his attempts to deal systematically with the data from the try-out phase, and each programmer should, of course, design his aiding devices to fit his particular needs in each specific situation. Our main purposes have been (a) to emphasize the relative complexity of the evaluation task and the risk of using one-sided criteria in the try-out process, and (b) to show some of the possible ways in which the programmer can equip himself with simple aids in order to make his evaluations more systematic, complete, and efficient. It is felt that a more widespread use of such systematic and comprehensive evaluation schemes would also considerably enhance the accumulation of basic knowledge in the programmed learning field.

15. EXAMPLES OF PROCEDURES IN REVISION OF PROGRAMS

We have already discussed a series of steps in revision, arising from different test results, when we were dealing with the various evaluation criteria above. In the present section we shall briefly review different combinations of criteria which to a greater or lesser degree call for revision and also mention a few special measures which can occasionally be indicated.

15.1 SURVEY OF PRELIMINARY INDICATIONS FOR REVISION FROM CERTAIN COMBINATIONS OF CRITERIA

Figure 15.1 gives a simplified survey of the extent to which different combinations of criteria indicate revision. Such a survey must obviously be fairly schematic. In practice, it is not always easy to determine what are "high" or "low" error rates or decide, for example, how a "limited revision" should be designed. But as a theoretical outline such a schematic survey can perhaps nevertheless give a lead and counterbalance the one-sided consideration of a single criterion of which some programming theoreticians are guilty.

Let us comment briefly on some of the criteria combinations. (For abbreviations used, see the explanations in Figure 15.1.)

All combinations with a high terminal error rate (TE_h) obviously definitely indicate revision In these cases the goal set has not been attained. The program is, to use very obvious terminology, simply not "finished". If, in addition, we consider the student's attitudes, we get yet another indicator, and the difference made in the survey between "maximal" and "strong" indication for revision is directly dependent on the difference between programs as experienced by students with negative and positive attitudes respectively.

A program which not only misses its performance goal but also created an unpleasant study situation does not justify its existence. In this case it makes no difference whether the error rate is high (as in the first criteria combination) or low (as in the second). Programmers who discuss revision only on the basis of the students' errors during the course of the program risk accepting a program of the type $ME_lA_nTE_h$, which is here shown as indicating a maximal need of revision.

Everyone would agree that criteria combination 8 ($ME_lA_pTE_l$)is advantageous, and that a program with these characteristics scarcely needs revising. (For the sake of simplicity, we are disregarding certain other variables, e.g. teaching time, which is in many cases important, but which is less so compared to the final performance and working climate.) We have also accepted criteria combination 7 ($ME_hA_pTE_l$). The error rate during the working of the program is admittedly high, but this has resulted neither in bad performance nor a negative attitude. The difficulties can have resulted from a low degree of support stimulation or a teaching method which failed to explain errors satisfactorily, but these difficulties are not such as to have an overwhelmingly deterrent effect. It is possible that this combination is comparatively rare, but when it does occur, it should hardly be rejected, despite the ME_h factor.

Finally, combinations 5 and 6 are characterized by a negative attitude together with a good final result. In one case, ($ME_hA_nTE_l$), the negative attitude is probably connected with the error-reaction frequency. Here it is possible that a limited revision to reduce the error rate would aldo produce a better attitude climate. In the other case, it s more probable that the negative attitude is related to the fact that the program is experienced as monotonous and uninspiring. In this situation, an attempt should be made to increase motivation and reduce monotony (cf. our discussion on motivation above).

In a particular situation, the occurrence of different combinations depends not only on the character of the program, but also on the general working climate of the classroom or school and on the student as an individual. Some programs probably give rise to widely different criteria combinations for different individuals which may be of interest, for example, from the point of view of personality research and social psychology. From the point of view of educational technology, however, the general aim must be to construct the program or system in such a way that as many of the target students as possible come up with combinations 7 or 8.

Figure 15. 1 Outline of principle indications for revision in certain criteria combinations

Criteria combination:	Preliminary indication:
1. $ME_hA_nTE_h$	Maximal revision indication
2. $ME_lA_nTE_h$	Maximal revision indication
3. $ME_hA_pTE_h$	Revision strongly indicated
4. $ME_lA_pTE_h$	Revision strongly indicated
5. $ME_hA_nTE_l$	Limited revision indicated
6. $ME_lA_nTE_l$	Limited revision indicated
7. $ME_hA_pTE_l$	Revision not indicated
8. $ME_lA_pTE_l$	Revision not indicated

Symbols:

ME = medial error rate (% of errors during the instructional program)

TE = terminal error rate (% of errors in final test)

A = attitude

h, l = high, low

p, n = positive, negative

15. 2 SPECIAL PROCEDURES

15. 2. 1 The Question of the Student's Capacity for Independent Work and the Possible Alteration of the Boundary between Educand-System and Educator-System

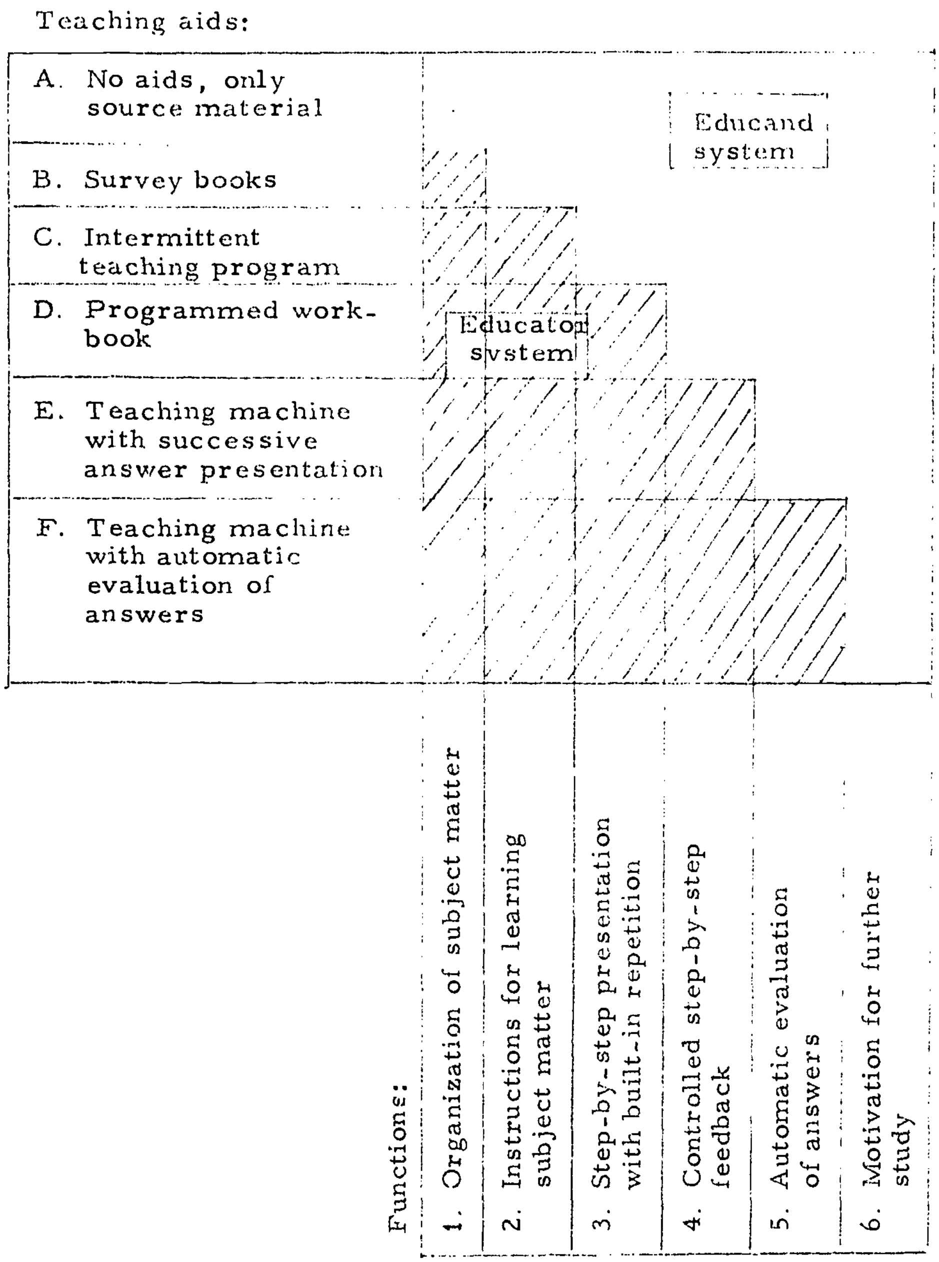
As is shown in Figure 15. 2, the change from one type of teaching situation to another often involves a shift of the boundary between the educator system and the educand system. An extended educator system may be valuable for the rapid learning of a certain subject-matter. On the other hand, when it is a question of developing the capacity to cope independently with subject-matter which is hard to learn, a comprehensive educator system would often be a hindrance. If the try-out shows that the learning of certain subject-matter takes more time than is desirable, one should consider changing over from a less detailed learning process to a more rigidly controlled one. If the aim is to induce the ability to treat subject-matter independently, and if the try-out seems to indicate that this goal is not being reached, a shift of the boundary between the educator system and the educand system in the opposite direction would be indicated.

15. 2. 2 The Question of How the Student Accustoms Himself to Active Cognitive Handling of Subject-Matter and Possible Limitation of the Number of External Response Reactions

External response reactions are considered valuable as a guarantee that the student actively handles the subject-matter presented to him. Is there a possible risk that the student accustoms himself to the special working technique in such a way that he avoids active cognitive working when there is no explicit response demand? Can the technique sometimes involve an unnecessary loss of time (for those with already established habits of study)?

If the try-out indicates a positive answer to one or both of these questions, one should possibly try a technique which is a compromise between the program's continual demand for external response reactions and the ordinary text-book's lack of such a demand. The program can, for example, be so constructed that a response which generally need only be mental (implicit) is demanded of the student. From time to time (at irregular intervals), however, a demand for an external statement occurs. Cook (1962 a), who suggested this type of technique, thinks that by this means the individual might be forced to keep his covert responses alert and so save time.

Figure 15.2 Outline sketch of the varying boundary between educator system and educand system



A schematic diagram of the reaction between student and program when using such a technique (with partial demands for explicit statement of responses) is given in Figure 15. 3.

15. 2 3 Special Measures against "Boredom"

We discussed above some general steps which could be taken to reduce monotony and increase motivation. Should investigation show that the students find the program monotonous and boring, the revision should aim at an even greater use of these tactics to give variety. In this section we shall give a few brief hints on other possibilities.

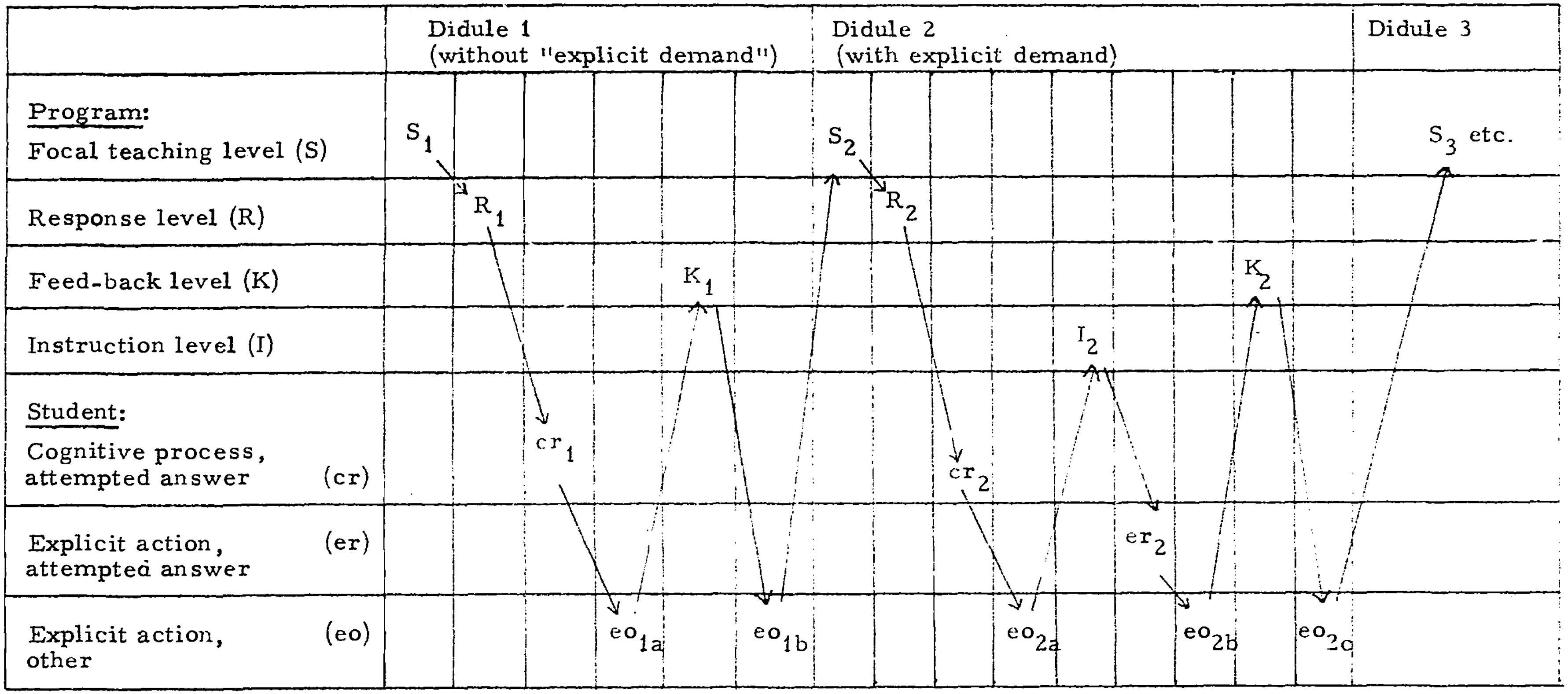
(1) Adaptation to the students' previous knowledge has perhaps not been worked out with sufficient care. As illustrated in Figure 15. 4, the student's attainment of the skills relevant to a certain goal is seen as a continous process, which has already gone on, to some extent, before he is faced with the particular study material. If he is aware that he is being taught about things he already knows, he will very often experience boredom. In this case a more thorough student analysis should form the background for a revision of the program.

It may become apparent that the previous knowledge of the target population has been seriously underestimated, and then the program should be shortened - either by cutting out certain sections in the more elementary parts of the program or by "increasing the size of the steps" (reducing the number of application- and practice-tasks). But it can also happen that the underestimation applies only to part of the target population and in that case consideration should be given to changing from a linear to a branching program.

The most effective teaching should, in theory, be that which starts out from an individual analysis of the initial repertoire, i. e. which begins by testing and analyzing the individual's previous knowledge and tendency to errors and then teaching only those points which are shown to be lacking - with emphasis on individual difficulties. The fact that the programmer for practical reasons finds it difficult to follow such a procedure completely does not prevent him approaching this technique, since the number of individual variations important for teaching is relatively limited in most groups of students. (One works then with a number of "streams" in which the students are placed after an initial filter test.)

It may be thought that student analysis is troublesome and timeconsuming, and it is often neglected in preparing the contents of the curriculum. However, the amount of work that one imagines one saves by

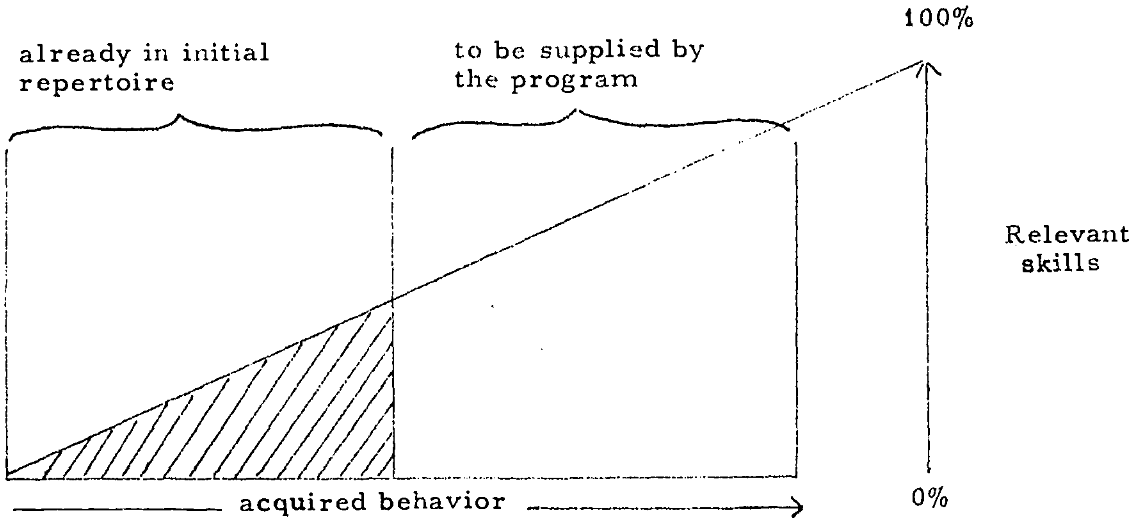
Figure 15.3 Schematic diagram of interaction between student and program demanding only partial explicit response reaction



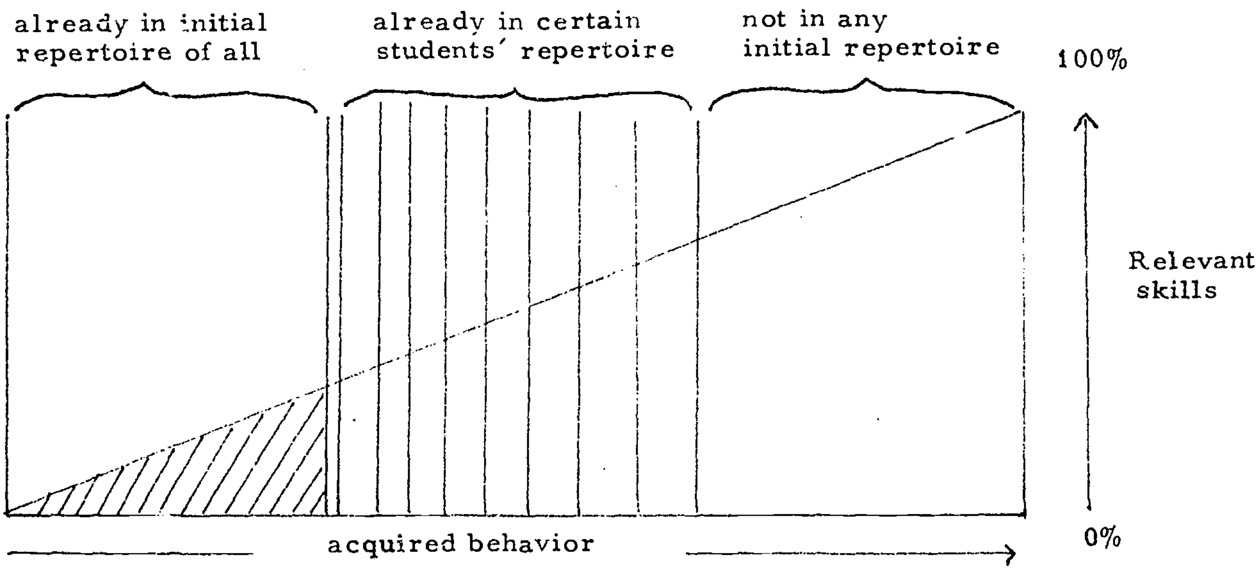
Note: The other actions, shown as eo, here indicate manipulation of the machine. In Didule 1 the student thinks of the answer (cr_1), turns the machine forward one frame (eo_{1a}) and gets the correct answer (K_1). In Didule 2, however, when he turns the machine forward (eo_{2a}), he gets instead the instruction to write down the answer (I_2). Only after he has done this (er_2) is he given, after a further manipulation of the machine, the correct answer (K_2).

Figure 15. 4 Theoretical illustration of relation between initial repertoire and contribution of the program

For the individual student S:



In the total group of students Z:



omitting this procedure will be lost over and over again, partly in lowered motivation among the students, with the consequent lower effort they put into their work, partly in an unnecessary expenditure of time and study-material when one teaches what the students already know.

(2) Alteration of tempo of presentation. As far as the tempo of presentation is concerned there are three main possibilities: student-controlled speed, machine-controlled speed and the combination student-and-machine-controlled speed. Student-controlled speed is the most usual, and it has the obvious advantage that each student has as much time as he needs to think. However, one can imagine (especially when it is not a question of concept-learning or problem-solving activity, but of a motor skill) that the student-controlled tempo involves a risk that the student will stay too long at a low level of skill. This is likely to lead to boredom, since clearly increasing demands are not continually being made on the student. Should a try-out show this to be the case, a situation should be considered in which the student is spurred on to greater efforts by a machine-controlled increase in speed, or which uses a student-machine system where the average speed of the machine is regulated by the student's skill (cf. SAKI).

(3) Variation of didule components. We normally work with a fixed series of didule components and student responses: information or "stimule" (S), response-demand or "respule" (R), student response or "response" (r), feed-back of correct answer or "corrule" (K). This in itself, as we pointed out above, gives a good dynamic variety (S-R-r-K, S-R-r-K etc.), with greater variation than the usual text's monotonous: S-S-S-S-S etc. However, if we work only with series of the type S-R-r-K, even this can naturally gradually become monotonous and boring. There is a particularly great risk that this will happen when, in addition, there is a great number of consecutive didules which contain very similar S, which is often necessary in the training of skills (e.g. in languages), where the student, by persistent practice, has to get as close as possible to the ideal performance (e.g. perfect pronunciation). If investigation shows that boredom is experienced by the students, one should consider interrupting the working sequence at suitable places with differently arranged component sequences. If, for example, in training of skills, we occasionally intensify the r-K combination by "extended" sequences, this can serve a useful purpose both for effective work (the ideal performance is attained more easily perhaps when the whole task need not be repeated) and for the students' need of "something different". This type of sequence - e.g. S-R-r-K-r-K-r-K- is of course only one of many possible variations.

16. COMPILING THE PROGRAM MANUAL

Program consumers need certain information from the program producer to help them choose their study material. As we have already pointed out, this information should be readily available in a program manual. The writing of this manual should be the final task of the program constructor.

It is naturally not a great deal of use to describe in detail what such a manual should contain. The most suitable choice will vary greatly from program to program. But certain categories of information are usually necessary in one form or another:

(A) Description of goal and students

The aim of the program and the use planned for it should be clearly stated. An attempt to attract customers among a wide and varying public by using vague and all-embracing headings is easily seen through and is not a good advertisement for the vendor of the program. The previous knowledge needed is stated an preferably specified by means of a pretest. Similarly, the purpose of the course is clearly defined by the final test. (For pre- and post-tests, see E below.)

(B) Character of development work

(a) Those responsible for the programming and testing should be named.

(b) The goal analysis on which the program is based should be presented.

(c) A summary of the other stages of the preliminary work can give valuable information. (To what extent has a thorough student analysis been carried out? Etc.)

(d) The revision and try-out processes should be given in outline.

(C) The structural characteristics of the program

should be briefly described and illustrated with extracts from the program. It is possible that in time a standard set of descriptive categories may be used as a basis for such a description. Fry (1963, p. 199 ff.) exemplifies a series of possible bases for classification. Some of those it would be of interest to give are:

(a) sequence characteristics, including main flow structure,

(b) general didule characteristics, and

(c) special characteristics of the components, i.e. of stimules, respules and corrules.

(D) Characteristics of the subject-matter

It should be stated here to what degree the program concurs with the curricula concerned, and also how far and in what way it differs from already existing textbooks (or programs) in content.

(E) Program effects

This is, in a way, the most important category of information. It is here that the intending purchaser of the program is informed of the demonstrated teaching effectivity of the program and of the conditions under which this empirical demonstration was carried out. In other words, the manual should contain the following information:

(a) Scope and conditions of the field try-out. This should be stated in some detail and comprise not only information about the research sample (its size, age of students, intelligence, previous knowledge etc.), but also external circumstances both as regards the teaching situation (e.g., how time was allotted, possible teacher co-operation) and as regards the test situations.

(b) The preliminary and final tests used in the field try-out should be described in the manual and in addition the complete tests should be available from the constructors or the publishers if desired. (Tests are sometimes sold together with self-instructional material. According to "Programs '63", there were then final tests for 46% of commercially available programs, but preliminary tests for only 14%. Were the publishers afraid that the tests might show that the student did not need the program in question?)

(c) The final results of the field try-out should obviously be clearly stated. It would be an advantage to give this both verbally and in the form of adequate statistics. It should be possible to compare the students' final performance with their previous knowledge, and there should be supplementary information about the amount of study time used. If other aims than simply the learning of subject-matter have been mentioned in the description of the goal of the course, the manual should also give an account of results of the try-out which confirm that these aims have been attained. Otherwise, the manual should clearly state that these points in the general goal description have not yet been confirmed by empirical results, and that therefore the effect of the program as far as they are concerned is, for the time being, hypothetical.

The task of producing a program manual is the responsibility of the program producer and publisher. This can hardly be considered an excessive burden, provided that the program constructor has really developed his program empirically and carried out a proper field-test.

Unfortunately, this kind of account is at present supplied with very few of the commercially available programs. This is not just due to an unwillingness to do the extra work involved in preparing the manual. It is mainly because many of these programs have obviously not had a thorough try-out in the field, a fact which may well be obscured by a flood of sales-talk about continuing research and delighted students, but which would look bad in print. When, in one enquiry, publishers of programs in the USA were asked about the size of the groups on which the programs had been tried out, over 80% avoided answering precisely that question! (Cf. "Programs '63", p. xiii.) In an examination of about 700 programs in 1967 at Northeastern University, it was found that only about 30% were accompanied by any information about field-tests. (Cf. Markle, 1968, p. 17.)

So it is important that program consumers are aware of the demands they should make, and that they do not hesitate to make them. Reviews such as "Programs '62" and "Programs '63" are undoubtedly useful in this respect, in that they give - without further evaluation, it is true - the nature and scope of the try-out. Publishers who are forced to leave a blank as far as this is conerned can scarcely avoid wondering what it will mean for their sales if their competitors can give fuller information. In addition to such survey lists, a standard forum for regular reviewing would be desirable. To take an example from a somewhat different field, one can point to the influence which Buros' "Mental Measurements Yearbooks", with their persistent pointing out of lack of reliability and validity checks, have had.

But apart from this, every buyer of a program naturally has the right to demand detailed information from the seller. Advertising generalisations are not much use to the consumer. The Bulletin of the National Association of Secondary School Principals (Belton, 1962) gives examples of a number of questions which a head teacher should put to a program salesman who has no program manual to show. Their tone and general tendency can be seen from the following (considerably shortened and modified) version:

"We thank you for your offer to demonstrate at our school the programmed study courses which your company is at present marketing. Our study material committee will be glad to see you next Monday at 3:30 p.m. Our time, however, is limited, and we should appreciate it if, when you visit us, you would be prepared to answer briefly the following questions about each of the programmed courses you intend to demonstrate:

1. Is there a teacher's manual or guidebook with instructions for using the course?

2. Are the goals stated in detail in a description of the course?
3. Has the course material been produced as a result of a series of try-outs and revisions? If this is the case, where can we find a description of this development work?
4. Who were those responsible for the development and trying out of the course? What are their qualifications?
5. To what extent has the completed course been tried out in the field, and where can we find an account of how the field trial was arranged and the results of it?
6. Which schools or organizations can we ask about their experiences in using the course?
7. Is the course accompanied by a pre-test and a post-test, or can such pre- and post-tests be ordered separately from your company? Where can we find out about the construction and evaluation of these tests?
8. What is the approximate average time taken for a student to complete the whole course?
9. Is the study material avialable both in book form and for machine presentation? Is the study material consumable, or can it be used again?
10. What is the cost of the material per student hour? Does your company allow a discount for a first purchase to try out the material?

When we have this information, we should be in a better position to decide on purchasing."

If, after receiving such a letter, programs salesmen with deficient programs fail to turn up for the propsed demonstration, no harm will have been done.

17. ASSESSMENT AND COMPARISON OF EXISTING PROGRAMS

We have previously studied self-instructional material mainly from the point of view of the constructor and how the teaching program is successively re-examined, tried out and revised. Finally, we shall treat the program in its finished state, above all with the following question in mind: How can we decide whether a teaching program fulfils reasonable demands and fits into a certain teaching situation?

This is mainly the program consumer's concern, but the question is also one for the program constructor. It usually turns up when the programming group has reached the point when analysis of pre-requisites has been made, i.e. when one is clear what the goal is and what type of students one will be dealing with. The questions one should ask oneself at this stage (but scarcely before are: Is there already a program on the market which corresponds to the goal which has been specified and which is aimed at the same group of students? Is there nevertheless a sufficient reason for a new program?

(Figure 17.1 shows schematically some factors in the situation about which the programming group has to make a decision.)

In the evaluation of a program, two kinds of judgment have to be made: (1) judgments based on the program's "pheno-characteristics" which can be observed by a visual examination and (2) judgments based on the program's "functional characteristics", i.e. on experience of how the program works in practice.

Questions about "pheno-criteria" in their turn are of two types: (a) Has the content of the program just those qualities which make it fit for the teaching purpose in question? (b) Does the structure of the program conform with the "theory" of programmed teaching (or with our knowledge of the principles of human learning)?

Obviously, in judging a program, one can easily come to certain conclusions by looking through it oneself. A new program should always be examined as throughly as a new textbook and primarily with regard to its contents and scope.

On the other hand (as Rothkopf, 1962, among others, has emphasized), one should beware of the idea that a program can be adequately judged merely by a visual examination of its structural characteristics. There are several reasons for this. (a) Firstly, there is scarcely a "theory" of what a maximally effective teaching program should look like, reliable enough to enable us to come to any definite conlusions. We have a certain

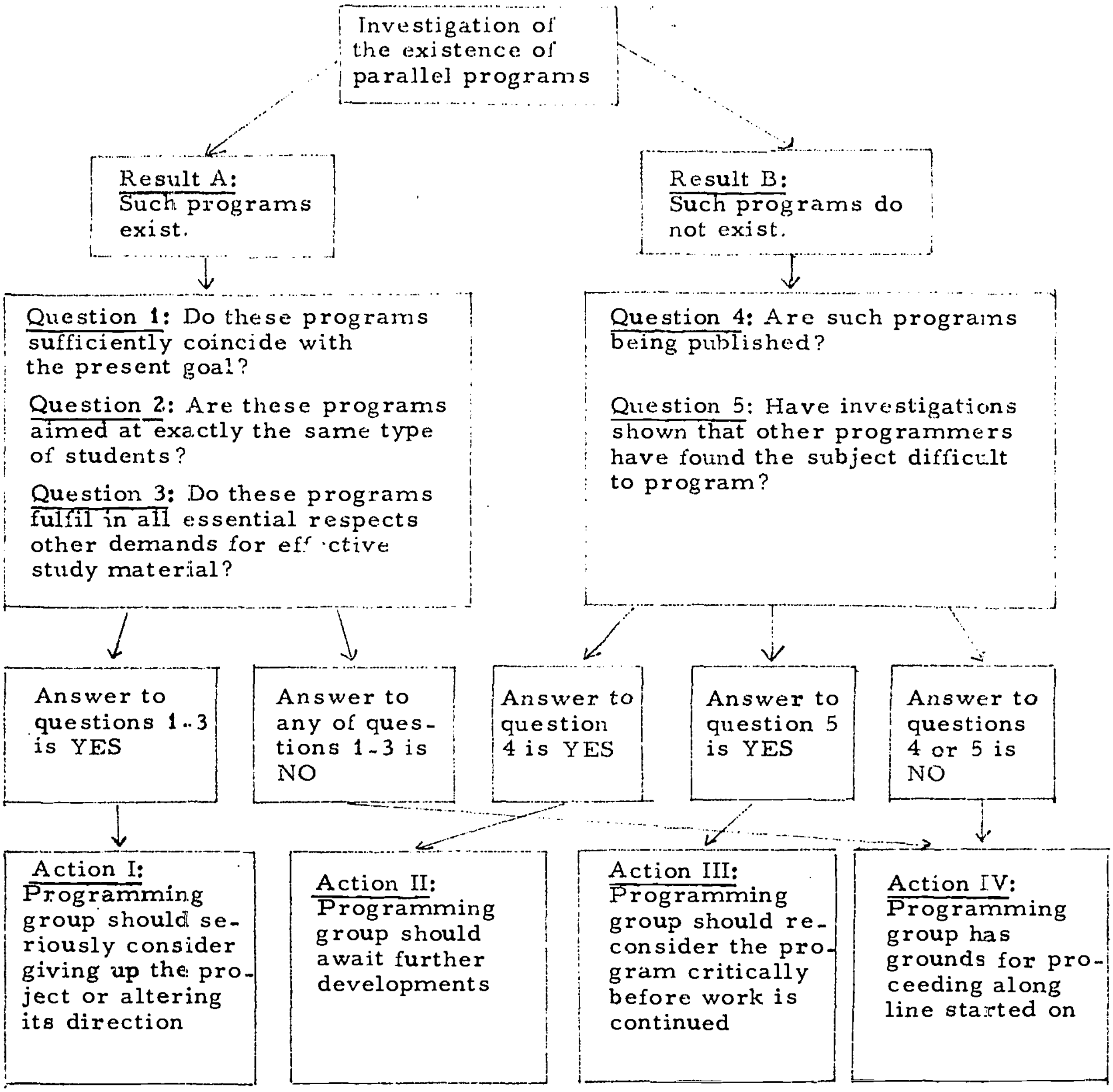


Figure 17. 1 Existence of parallel programs: A situation in which a decision must be taken during programming

body of general principles, but empirical investigations have given contradictory results on many points of detail. (b) Secondly, even if such principles were relatively clearly established, they cannot be a sufficient basis for an evaluation, since there are a whole series of other factors (e.g. organization of material) which affect effectivity. (c) Thirdly, even if the principles of programming were both clearly established and could in some way be shown to outweigh all other reasons for variation, a visual examination would still give insufficient information about the value of a program, for the simple reason that many of these principles cannot be directly observed. A low percentage of error is an example of a programming variable which can only be determined by empirical methods. (The expert in a subject often guesses wrongly about this. Precisely because of his expert knowledge, he considers the steps to be extremely "small" and consequently estimates the percentage of error as "lower" than empirical methods show it to be.)

Instead of discussing whether the teaching material conforms to certain programming principles, the experiment has been made of getting experienced teachers to judge its probable effectivity. One must, however, allow for the fact that, because of lack of training in judging these new types of material, they can go seriously astray. In one study (Rothkopf, 1963) the results even showed a negative correlation (-0.75) between teacher-estimated and real effectivity!

To sum up, we can say, then, that while a visual examination is always of interest, it must not be the only basis for the consumer's evaluation of a teaching program. He must rely primarily upon how the program has been shown to work in practice, not upon what it looks like. Possibly, in the future, we shall know so much more about the factors which produce an effective program that we shall not always need to proceed via empirical demonstration. But in the present state of programming technique this is the only completely reliable way.

When the consumer wishes to examine a program thoroughly, both visually and by studying the information given in the manual, it is advisable to use some form of check-list. This makes the task easier and also avoids a one-sided examination. The contents of the list can vary according to the needs of the particular teaching situation. We shall now give examples - partly based on an American "Checklist for selecting programs" (NSPI Journal, 1963) - of questions one often finds it necessary to put (cf. also e.g. Cox et al., 1962; Tracey, 1963;

Wark, 1963; and Vanderschmidt, 1964): The questions are so formulated that the answer "yes" usually gives the better alternative. In other words, a large number of positive answers is a better indication than few positive answers. However, some questions are obviously more important than others. These must naturally be given the greatest weight in drawing the final conclusion.

1. Goal and curriculum:

1.1. Does the publisher or constructor give a clearly formulated description (in the form of observable behavior changes) of what the course is intended to teach?

1.2 Does a visual examination show that the description of the goal is fully covered by the course?

1.3 Does the goal of the programmed course (as it appears in points 1.1 and 1.2) coincide in all essential respects with the curriculum or general goal for your special teaching situation?

1.4. If there is any divergence, is it of such a character that it is pedagogically advantageous and would therefore be accepted as a meaningful change in the course?

1.5. Is there reliable information about the time taken for the course by students of the type you meet in your teaching situation?

1.6. Will your students be able to get through the course, or as much of it as is desirable, in the time at your disposal?

To sum up: Does the general goal of the course sufficiently coincide with the demands of your teaching situation?

2. Student characteristics:

2.1. To get the best out of a course, certain skills and previous knowledge are often needed. Are these "starting requirements" clearly described?

2.2. Is there a pre-test, which can show whether a student has the required skills and knowledge to start the course?

2.3. Do these starting requirements coincide with what you could reasonably expect of a student in your teaching situation?

2.4. Is there a pre-test which will show how much a student already knows of the material the course is intended to teach?

2.5. Is the course material so organized that the student with such previous knowledge can start the course at some other point than the beginning or otherwise go through the material in a way suited to his special qualifications?

To sum up: Does the general design of the course reasonably suit the group of students you are teaching?

3. Specific instructional effectivity of the study material:

3.1. Is there a detailed and reliable report of the field try-out of the material in its finished state on representative groups of students (including pre- and post-tests)?

3.2. Are the results of this field try-out satisfactory?

3.3. Were the conditions in the try-out sufficiently like those of your own teaching situation for you to be able to assume a similar result in your own case?

3.4. Has a test been published with which you can check, by a satisfactory measuring procedure, whether the goal of the course has been achieved?

3.5. Are there grounds for assuming that the knowledge and skills acquired during the course are retained for a reasonable time afterwards (i.e. has the try-out also studied retention after a certain interval, with satisfactory results)?

To sum up: Are there both (a) good grounds for supposing that the study material gives the desired specific results with typical groups of students and (b) the possibility of checking in your own teaching situation (suitable final tests)?

4. General educational character of study material:

4.1. Does the material give the student training in desirable study behavior (e.g. does it involve more logical reasoning and induction from experience than copying and formal mechanical working)?

4.2. Does the material give the student sufficient opportunity for varied repetition (as against both insufficient and routine repetition)?

4.3. Does the material give the student sufficient opportunity to integrate his knowledge (e.g. by problems which promote integration, work with supplementary and source material etc.)?

4.4. Is there a teacher's handbook which gives examples of how to integrate the self-instructional material with other teaching methods?

4.5. Are there interim tests built into or accompanying the course which will be a good basis for possible additional help or teaching during the course?

4.6. Are there grounds (as a result of the field try-out or research) for supposing that the material created an interest for the subject?

To sum up: Has the study material such educationally desirable characteristics that, in addition to its specific instructional effect, it also promotes such "fringe benefits" as good study technique, positive attitude to further study etc. ?

5. Economic questions:

5. 1. Is the cost of the complete material, including the necessary accessories such as teaching machines and/or printed forms for answers, covered by available resources?

5. 2. Will it be easy for you to attend to the storage and upkeep of the material?

5. 3. If an expensive machine is necessary for the presentation of the material, will it be possible for you to use it for other courses?

To sum up: Is the purchase of the course material economically defensible?

In certain cases the consumer may consider it desirable to try out the material himself. This is perfectly legitimate, if the program is thought to fulfil all the consumer's requirements, but has been tried out on a group of students which differs essentially from those for whom the teacher intends to use it. In such a case, it is good strategy to try it out on a smaller group before the program is introduced on a large scale.

REFERENCES

AERA-APA-DAVI Joint Committee on Programed Instruction and Teaching Machines. Criteria for assessing programed instructional materials. Audiovisual Instruction, 1963, 8, 84-89.

Anschütz, H. Die Verteilung der Begriffe in Lehrprogrammtexten. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht, 3. Stuttgart: Klett, 1965. Pp. 104-113.

Barlow, J.A. Conversational chaining in teaching machine program. Psychol. Rep., 1960, 7, 187-193.

Basescu, B. More on multiple-choice frames. Programed Instruction, 1962, 2 (1), 3.

Becker, J.L. A programmed guide to writing auto-instructional programs. Camden, N.J.: RCA, 1963.

Belton, J.R. Programmed learning - 20 questions. The Bulletin of the National Association of Secondary School Principals, 1962, 46 (No. 278), 77-78.

Bjerstedt, Å. Självinstruerande studiematerial. Lund: Uniskol, 1963.

Björkman, M. Inlärningspsykologiska principer i programmerad undervisning - några frågetecken. In: Skolöverstyrelsen, Utredningar i skolfrågor 11, 1963. Pp. 96-105.

Bloom, B.S. (Ed.) Taxonomy of educational objectives. Handbook I: Cognitive domain. New York: Longmans, 1956.

Bloom, B.S. Individual differences in school achievement: A vanishing point? (Address, AERA, New York, Febr. 1971.)

Bushnell, D.D. & Allen, D.W. (Eds.) The computer in American education. New York: Wiley, 1967.

Cavanagh, P. & Jones, C. Programmes in print 1966. London: Association for Programmed Learning, 1966.

Cavanagh, P. & Jones, C. (Eds.) Yearbook of educational and instructional technology 1969/70. London: Cornmarket, 1969.

Center for Programed Instruction. Programs, '62. A guide to programed instructional materials available to educators by September 1962. Washington, D.C.: U.S. Government Printing Office, 1962.

Center for Programed Instruction. Programs, '63. A guide to programed instructional materials available to educators by September 1963. Washington, D.C.: U.S. Government Printing Office, 1963.

Center for Programed Instruction. Programed instruction materials 1964-'65. A guide to programed instruction materials available for use in elementary and secondary schools as of April, 1965. New York: CPI, Columbia Univer., 1965.

Cook, D.A. Behavior theory and the automation of instruction. In: J.E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962. Pp. 120-128. (a)

Cook, D.A. Studying the performance of a program. Programed Instruction, 1962, 2 (2), 4-8. (b)

Cox, A. et al. Suggested criteria for selection of programed materials. In: Summary Report of the Eighth Lake Okoboji Audiovisual Leadership Conference. Iowa City: State Univer. Iowa, 1962. Pp. 80-88.

Cram, D. Explaining "teaching machines" and programming. San Francisco: Fearon, 1961.

Dahllöf, U. The contents of education with regard to demands for different jobs and for further studies. (2nd ed., mimeo.) Stockholm: School of Education, 1965.

Deutsch, W. Programed learning: An overview of personnel and financial requirements. In: S. Margulies & L.D. Eigen (Eds.), Applied programed instruction. New York: Wiley, 1962. Pp. 208-211.

Eigen, L.D. & Margulies, S. Response characteristics as a function of information level. Journal of programed Instruction, 1963, 2 (1), 45-54.

Evans, J.L., Glaser, R. & Homme, L.E. An investigation of "teaching machine" variables using learning programs in symbolic logic. J. educ. Res., 1962, 55, 433-452.

Evans, J.L., Homme, L.E. & Glaser, R. The Ruleg system for the construction of programmed verbal learning sequences. J. educ. Res., 1962, 55, 513-518.

Fry, E.B. Teaching machines and programmed instruction. New York: McGraw-Hill, 1963.

Gage, W.W. Contrastive studies in linguistics. Washington: Center for Applied Linguistics, MLA, 1961.

Geis, G.L. Some considerations in the evaluation of programs. AV Communication Review, 1962, 10, 64-69.

Gestrelius, K. Utbildningsplanering med befattningsanalytiska metoder. Lund: Uniskol, 1970.

Gilbert, Th.F. Mathetics: An explicit theory for the design of teaching programmes. (Rev. educ. Cybernetics & appl. Linguistics, Supplement 1.) London: Longmac, 1969.

Glaser, R. Some research problems in automated instruction. In: J.E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962. Pp. 67-85.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. Amer. Psychologist, 1963, 18, 519-521.

Gordon, J.M. The evaluative response: A rate problem. AID, 1963, 3, 240.

Grace, H.A. Programming and readability. Calif. J. educ. Res., 1963, 14, 164-166.

Greenberg, M. Conversational chaining in programmed instruction. Progr. Learn. educ. Technol., 1968, 5, 115-121.

Holland, J.G. Evaluating teaching machines and programs. Teachers College Record, 1961, 63, 56-65.

Holland, J.G. & Porter, D. The influence of repetition of incorrectly answered items in a teaching-machine program. Journal of the experimental Analysis of Behavior, 1961, 4, 305-307.

Hughes, J.L. Programed instruction for schools and industry. Chicago: Science Research Associates, 1962.

Husén, T. & Dahllöf, U. Mathematics and communication skills in school and society. Stockholm: The Industrial Council for Social and Economic Studies, 1960.

Härnqvist, K. & Grahm, Å. Vägen genom gymnasiet. Stockholm: SOU, 1963.

Jacobs, P.I. Item difficulty and programed learning. Journal of programed Instruction, 1963, 2 (2), 21-38.

Klaus, D.J. & Lumsdaine, A. A. Some economic realities of teaching machine instruction. In: S. Margulies & L.D. Eigen (Eds.), Applied programed instruction. New York: Wiley, 1962. Pp. 198-207.

Kopstein, F.F. & Cave, R.T. Preliminary cost comparison of technical training by conventional and programmed learning methods. Wright-Patterson Air Force Base, Ohio: MRL, 1962.

Krathwohl, D.R., Bloom, B.S. & Masia, B.B. Taxonomy of educational objevtives, Handbook II: Affective domain. New York: McKay, 1964.

Larsson, I. Tyska språket i delar av svenskt näringsliv: En studie i enkätmetodik och en undersökning av språkbehoven. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 101, 1969.

Lehnert, U. (Ed.) Elektronische Datenverarbeitung in Schule und Ausbildung. München: Oldenbourg, 1970.

Lewis, B.N. & Pask, G. The theory and practice of adaptive teaching systems. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 213-266.

Lindell, E. A system of study materials for the teaching of foreign languages: Principles and examples. Didakometry and Sociometry, 1969, 1, 91-124.

Lindsten, C. Elevpreferenser vid stoffval i hembygdskunskap, årskurs 3. Pedagogisk-psykologiska problem (Malmö. School of Education), No. 95, 1969.

Little, J.K. Results of use of machines for testing and for drill upon learning in educational psychology. J. exp. Educ., 19-4, 3, 45-49.

Löfqvist, G. Analys av officiella texter för bestämning av kraven på ämneslärarbefattningen. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 103, 1969.

Mager, R.F. Preparing objectives for programmed instruction. San Francisco: Fearon, 1961.

Mager, R.F. & Clark, C. Explorations in student-controlled instruction. In: G.D. Ofiesh & W. C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 235-238.

Markle, S.M. Faulty branching frames: A model for maximizing feedback to the programer. Programed Instruction, 1963, 3 (1), 4-8.

Markle, S.M. Five answers in search of the right questions: The bibliographies of programed instruction. NSPI Journal, 1968, 7 (3), 15-21.

Markle, S.M. Good frames and bad. (Rev. ed.) New York: Wiley, 1969.

Markle, S.M. A programed primer on programing. Vol. I. Principles. New York: Center for Programed Instruction, 1961.

Marton, F. Evaluering av evalueringar. Rapporter från Pedag. inst., Göteborgs univer., No. 42, 1969.

Pennington, D.F. & Slack, C.W. The mathetical design of effective lessons. In: S. Margulies & L.D. Eigen (Eds.), Applied programed instruction. New York: Wiley, 1962. Pp. 298-310.

Payne, D.A. The specification and measurement of learning outcomes. Waltham, Mass.: Blaisdell, 1968.

Popham, W.J. et al. Instructional objectives. Chicago: Rand McNally, 1969.

Porter, D. Some effects of year-long teaching machine instruction. In: E. Galanter (Ed.), Automatic teaching: The state of the art. New York: Wiley, 1959. Pp. 85-90.

Randell, G.A. Adaptive teaching machines. (Report, Berlin 1963.)

Roe, A. Automated teaching methods using linear programs. J. appl. Psychol., 1962, 46, 198-201.

Rothkopf, E.Z. Criteria for the acceptance of self-instructional programs. In: A.E. Traxler (Ed.), Improving the efficiency and quality of learning. Washington: Amer. Council on Educ., 1962. Pp. 30-38.

Rothkopf, E.Z. Some observations on predicting instructional effectiveness by simple inspection. Journal of programed Instruction. 1963, 2 (2), 19-20.

Schaefer, H.H. A vocabulary program using "language redundancy". Journal of programed instruction, 1963, 2 (3), 9-16.

Shay, C.B. Relationship of intelligence to step size on a teaching machine program. J. educ. Psychol., 1961, 52, 98-103.

Shettel, H.H. & Lindley, R.H. An experimental comparison of two types of self-instructional programs. Pittsburgh: AIR, 1961.

Silberman, H.F. et al. Fixed sequence versus branching auto-instructional methods. J. educ. Psychol., 1961, 52, 166-172.

Silverman, R.E. The evaluation of programmed instruction: A problem in decision making. Psychology in the Schools, 1964, 1, 74-78.

Sorenson, A.G. The use of teaching machines in developing an alternative to the concept of intelligence. Educ. psychol. Measmt, 1963, 23, 323-329.

Standards Committee of the NSPI. Checklist for selecting programs. NSPI Journal, 1963, 2 (6), 4.

Stolurow, L.M. & Davis, D. Teaching machines and computer-based systems. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 162-212.

Swets, J.A. & Feurzeig, W. Computer-aided instruction. Science, 1965, 150, 572-576.

Thorndike, E.L. Education. New York: Macmillan, 1912.

Tingsten, H. Gud och fosterlandet. Studier i hundra års skolpropaganda. Stockholm: Norstedt, 1969.

Tracey, W.K. Content and construction review. NSPI Journal, 1963, 2 (10), 13-14.

Trittipoe, W.J., Trittipoe, T.G. & Hahn, C.P. The effectiveness of three programed-learning frame styles. J. progr. Instr., 1963, 2 (1), 35-43.

Vanderschmidt, H. Validation data for programmed texts: A checklist for evaluation of testing. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 210-212.

Wark, D. et al. Checklist comments and criticism. NSPI Newsletter, 1963, 2 (1), 7-8.

Weltner, K. Der Shannonsche Ratetest in der Praxis der programmierten Instruktion. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht, 4. Stuttgart: Klett, 1966. Pp. 40-53.

Weltner, K. Lernziele unter dem Aspekt der Informationstheorie. In: B. Rollett & K. Weltner (Eds.), Fortschritte und Ergebnisse der Unterrichtstechnologie. München: Ehrenwirth, 1971. Pp. 26-35.

Zaccaria, M.A. & Olsen, J. Reappraisal of achievement measures. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 213-214

Öreberg, C. et al. IMU-projektet: Målbeskrivning för ett självinstruerande studiematerial i matematik för grundskolans högstadium. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 43, 1966.

APPENDIX I

THE TERMINOLOGY OF PROGRAMMED INSTRUCTION AND EDUCATIONAL TECHNOLOGY: A SELECTIVE LIST

Some terms within the area of educational technology and programmed instruction are listed with brief definitions. Naturally, complete coverage is not aimed at. Instead, some attempt has been made to include the major terms used in the present survey and/or that together form a relatively consistent frame of reference.

Adaptive instructional system An instructional system which can adjust itself to varying student behavior by, for example, giving simpler tasks if the student has not been successful, and more difficult tasks when the student is successful.

Algorithm General and non-ambiguous prescription for solving problems within a specific problem area. Part of the subject-matter analysis (as one step in the pre-writing phase when constructing self-instructional study material) may be focused upon the questions: to what extent is it possible to formulate algorithms for the subject-area involved, and to what extent can and should these influence the instructional sequence and method?

Amplified T-tabulation A "T-tabulation" (see that term) to which, besides the usual data from the goal analysis, has been added a list of items of information about the target population with point of departure from the student analysis. Reports on the special preliminary knowledge and existing skills of the students have then been added, as well as possibly a judgment as to how the ordinary occupation of the students after their specific educational training can affect the results of the instructional program in the long run. An influence of this kind can naturally be both a support and a hindrance; hence, an appraisement in this respect can modify the programmer's treatment of the question of repetition and overlearning. (For illustration, see chapter 6 above.)

Anschütz diagram A diagrammatic method (developed by H. Anschütz) for the evaluation and comparison of instructional programs, utilizing an examination of the distribution of concepts over the instructional micro-units. (For details, see Anschütz, 1965.)

Branching model A course layout which makes it possible for different categories of students (for example, students with different preparatory knowledge) to take separate paths through the study material.

C-catalog A collection of fundamental concepts (C) in a certain subject area. Systematizing these logical "basic units" within a hierarchical system results in a "structured C-catalog". (A card file with the separate units on separate cards is often a practical set-up for such a catalog.)

C. C-catalog (or "CxC-catalog") A collection of instructionally interesting relationships between the logical "basic units" (C) within a certain subject area. A catalog of this type (which in concrete form is ordinarily set up as a card file) can be made up with the help of a "C-matrix" (see that word).

Chain model ("conversational chaining") A course layout in which the individual instructional units are linked together in such a way that the feedback component of the preceding instructional unit is incorporated as a part of the stimule component of the following unit. In other words, the key answer is not given a separate presentation, but is embedded in the next unit, ordinarily as one or more words printed in italics.

C-matrix An aid in mapping the instructionally interesting relationships between various fundamental concepts within a certain subject area. In principle it consists of a square-ruled diagram, the row and column headings of which are the key words for the various "basic units" in the C-catalog, and every cell in the matrix is thought to symbolize all possible relationships between two such units. The constructor can, for example, go through the matrix cell by cell and ask himself questions of this type: What connection is there between these two concepts, and is this connection such that it can and should be used pedagogically? (For illustration, see chapter 7 above.)

Computer-assisted instruction Usually a special case of programmed instruction, viz, the case when a computer handles information storage, registrering, response comparison and feedback, during a highly individualized instructional program.

Correlative stimuli Information, instructions and evaluations, which in an instructional situation are presented in close connection with the principal points of the content but are not themselves meant to be learned. When these correlative stimuli have the chief function of facilitating the learning of certain other items, they can be called "prompts". (Compare "Focal informative stimuli".)

Corrule See "Didule".

Crowder model A course layout which combines a branched flow-model (in which the student's working path is successively determined by his accomplishments) with use of multiple-choice responses. The model has been realized partly in book form (see the term "Scrambled book"), partly in the form of mechanical presentation (e. g. Auto-Tutor, Mark II). Crowder's most frequent use of the branching technizue seems to be for explanation of an error just made.

Density of information A measure of the effective communication of information ("reduction of uncertainty" for the student; compare this word) in relation to some quantitative characteristic of the study material, e. g., the number of instructional units. For example, the fewer "didules" (see that word!) which are used to present a certain amount of information to the student, the greater the density of information can be said to be.

Didactogram A diagram (used during the empirical testing of instructional programs) showing one or several students' progress through a programmed study material. As background pattern for the didactogram, the previously worked out "didactoplan" may sometimes be used (see word "Didactoplan"), after the necessary revisions and specifications have been incorporated. On such a pattern (showing the basic structure of the study material) are presented the behavior data: in part, information as to the study routes used; in part, information as to right and wrong performances. (For illustrations, see chapter 14 above; especially Figures 14. 4 and 14. 5.)

Didactoplan A graphic visualization (mainly made use of during the prewriting phase) of the internal order in which one plans to bring up different subject-matter units (e. g., concepts) in the auto-instructional study material. By the use of special symbols, it can be shown on the chart which points in the material will be gone through by all, and which will be presented only to certain students.

Didule The fundamental instructional unit in programmed study material, ordinarily consisting of three components: the "stimule" (the information component), the "respule" (the response request), and the "corrule" (the feedback component). The didule can be said to represent the contribution of the program to the fundamental unit of interaction between the pupil and the program, which ordinarily has the temporal sequence: s-r-R-c. (The letters s, r and c refer here to the three components mentioned above; capital R symbolizes the student's answer reaction or "response"). - The didule is an instructional unit in programmed study material, that is, it is not identical with a presentation unit. (The most nearly equivalent term "frame" has the disadvantage of being used both for the concept 'instructional unit' and for the concept 'presentation unit'.)

Didule protocol An aid in the examination of an instructional program. The protocol consists of a chart in which each separate didule has a matching review column and the different review points are entered as row headings. (For illustrations, see chapter 12 above, especially Figures 12.2-12.5.)

Educand The individual who is subjected to educational influence (instruction, attitudinal influence etc.). The "receiving" component in an educational process. In school, "educand" = "student", but "educand" is a more general term.

Educator Someone or something which causes educational influence (instruction, attitudinal influence etc.). The "sending" component in an educational process. In school, "educator" often = "teacher", but "educator" is sometimes used in a more general sense, referring also to non-personal sources of influence (self-instructional materials-and methods systems etc.).

Egrul See "Ruleg".

Error rate The percentage of incorrect responses on a single instructional unit or a set of such units when tested on a group of students. Low error rate during work on program has often been considered desirable, but can also be obtained by non-desirable design (such as overprompting) and cannot therefore be used as an evaluation criterion by itself.

Fading Successive diminishing of the number of elements in a stimulus complex. In programmed teaching the term is used in two contexts. (1) = Successive cutting down of prompts (so that the student gradually can give a correct answer without help, while he was able to give the correct answer in the beginning only with the help of certain clues;

"fading of correlative stimuli"). - (2) = Successive disconnecting of separate parts of a larger stimulus complex, which is meant to be learned as a whole ("fading of focal stimuli"). Example: The student learns a poem by repeated recitations. As a support for his memory he has a text for use in recital, which successively is cut down as words and sentences are deleted. - For meaning (2) the term "vanishing" is sometimes used.

Feedback technique The interplay between the student and the study material, which tells the student how far he has succeded with each separate task. Ordinarily the feedback takes the form of a key answer presentation, but simple information as to whether the student response was right or wrong, or more detailed error explanations occur in some arrangements. It is typical of programmed instruction that feedback is almost immediate and very frequent.

Filter didule A unit in the study material which informs us as to the best possible working track (branch) for the individual student. Depending on what the more specific purpose is, we can speak of "precontrolling" or "post-controlling" didules (see these words).

Focal informative stimuli The main points in the content which in an instructional situation are presented to a student to be "learned" (compare "Correlative stimuli").

Formal prompts Stimuli, which provide knowledge about the form of the expected response and thereby make it easier for the student to give the correct answer. The student can, for example, be allowed to see the first letter of the correct answer word or to know how many letters are in the word. (Since overuse of this type of clue technique can promote thoughtless answering with an inadequate direction of attention, it should as a rule be used sparingly.)

Goal adjustment Successive changes of earlier goal descriptions guided by empirical controls of the realism of the goal level (in concrete cases where the initial behavioral repertoire of the target group and certain resources are defined). Is sometimes considered to be the final phase of a complete goal-analytic process (cf. "Goal seeking" and "Goal focusing").

Goal analysis Analysis of instructional objectives. A careful and detailed review of which knowledges and skills the students should be able to display after going through a certain study material. (Compare

"Terminal behavior repertoire".) Sub-components: "goal seeking", "goal focusing", and "goal adjustment".

Goal focusing Decisions about and formulation of the specific components of a goal description. "Goal focusing" can often be regarded as an intermediate link between "goal seeking" and "goal adjustment" in a total goal-analytic procedure.

Goal-related assessment See "Norm-related assessment".

Goal seeking Collection of data as a basis for goal decisions and goal formulation. Normally, the goal seeking should involve empirical data collections, be as many-sided as possible, and also give starting-points for priority ranking among various sub-goals.

Horizontal panel-book A programmed textbook in which the pages are divided into panels which are not read in the usual reading sequence (from top to bottom) but from page to page (or from sheet to sheet). The student reads, in other words, the top part of the pages through the whole book (or a large part of the book) before he goes over to another section of the same pages. The key answer for a given task does not usually appear in the same section or on the same page, but separately on the back of the section or in connection with the next information unit.

Instructional effectivity Some measure of the instructional result in relation partly to a specific goal, partly to instructional time expended. An effective instructional situation is, therefore, a situation in which an optimal balancing of behavioral products (as defined by the goal analysis) is achieved with the minimum expenditure of time.

Integrator A teaching aid which offers the student a constantly alternating sequence of (a) presentation of material and (b) the possibility of response reaction. Programmed textbooks and teaching machines are typical examples. An integrator, through this double function, is different from the much more usual aids in conventional teaching, which either have just the function of "presenting" or only give the only give the student an opportunity for a response reaction. The term "presentator" describes an aid in teaching, the chief task of which is to present to the student a stimulus-sequence prepared in advance (an ordinary textbook or educational film would be typical examples). The term "reactor, on the other hand, refers to a teaching aid, the chief function of which is to offer the student opportunities to give response reactions. (Example: Question lists.)

Intermittent model A layout of the auto-instructional material which demands a fairly independent treatment of the material by the student. He is presented with a "guide-book" containing a detailed study plan and self-correcting part tests, but the main material is collected by the student from different sources outside the guide-book. If the result of a test is good, the student is directed to new material. If it is not good, he is sent back to the same source (or one parallel to it). This shifting between instructions, independent study, and diagnostic review tests places the intermittent model somewhere between a completely programmed course and Pressey's "tail model" (see that word).

Iteration model A course layout in which the student always goes through the various instructional micro-units in a pre-established order, the same for all students, but where he is directed to repeat single units (or groups of units) under certain achievement conditions.

Linear model A course layout with the same task sequence for all students.

Mathetics A systematic approach to the analysis of instructional components and the construction of instructional units and sequences, devised by T. Gilbert.

Micro-branching model A "branching model" in which a certain amount of student differentiation may take place within the separate instructional unit (e.g., different error explanation), while these units as a whole are taken in the same order by all students. Most scrambled books, for instance, are of this type. (See "Branching model" and "Scrambled books".)

Multiple success strategy An attempt to create a dual motivation (possibly for different categories of students simultaneously). This can mean, for example, that on the one hand the student works through the study material with a high probability of getting the correct answer to separate tasks ("sure winner"), while on the other hand he is offered "extra success possibilities" (for example, possibilities for fast-track procedure and/or the satisfaction of answering correctly without clues, "individualized clue feeding").

Norm-related assessment An assessment of the instructional result in which the single student's result is related to information about other students' results. (For example: "On this test, student A has obtained a result which places him in the best 25 % of group G.") In contrast, "goal-related assessments" relate the instructional result to a specific series of goal descriptions. (For example: "Out of the items of this test, designed to cover all relevant goal descriptions, student B has mastered 80 %".)

Objective A statement describing the behavior the student is expected to acquire from a course.

Parallel track arrangement A branching model (see that word) which allows two or more separate categories of students to go through differently designed study plans with the same final goal. The same principle or group of problems can, for example, be treated from various points of view, depending on the varying preliminary attitudes of the different students.

Passage criterion The lowest possible point which the student can be allowed to reach with reference to a certain terminal behavior. In goal analysis it is important to fix specific values of this kind, as soon as we are dealing with a continuum of skill within a certain area of behavior. Without a clearly defined "minimum for acceptability" it is impossible to judge how well the goal set by the course has been reached. The passage criteria can suitably be included in the "T-tabulations". (Compare "Goal analysis", "Terminal behavior", and "T-tabulation".)

Post-controlling didule A unit in the study material which attempts to ascertain whether the student has learned carefully enough what has just been studied. If his responses show that this is not the case, the student is directed back to a section which has already been gone through or to an alternative repeat section (with similar, although not identical content).

Pre-controlling didule A unit in the study material with which an attempt is made to determine whether the student has sufficient preparatory knowledge to be able to skip a certain section of the course.

Presentator See "Integrator".

Programming Construction of auto-instructional study-material according to a three-phase procedure consisting of (a) preparatory work ("pre-writing" phase) with, among other things, emphasis on goal analysis and subject-matter analysis, (b) composition of a preliminary version utilizing certain psychological principles of learning and (c) a "post-writing phase", including successive empirical try-outs and revisions in the preliminary versions until the goal determined in advance has been reached.

Prompt Some stimulus added to the terminal stimulus to make a correct response more likely during a learning process. Approximate synonyms (at least in some texts): Cue and hint. Sometimes two major types are distinguished: "Formal prompts" and "Thematic prompts" (see these words).

Reactor See "Integrator".

Reduction of uncertainty A measure of the average effectivity of an instructional material, expressed as a reduction of the number of errors (within a specific student group). If we examine the ability of the students to solve a related terminal test item before going through the material ("error percentage A"), as well as their ability to solve the problem after going through the material ("error percentage B"), then the measure of reduction of uncertainty would be "error percentage A - error percentage B". (Compare terms "Speed of information" and "Density of information".)

Response request That part of an instructional micro-unit which presents the student with the task of giving an active response reaction (to solve a problem in mathematics, to fill in a word which has been left blank, choose the right alternative from several which have been presented, etc.). It is typical of the programmed instruction that these response requests are very frequent and continually alternate with the presentation of new material.

Respule See "Didule".

Reverse accumulation A system for practising a temporal behavior sequence as follows: the final phase of the series is practised first, after this the next to the last plus the last, etc. Besides the fact that the student practises sequences (instead of separate phases only), the advantage of this arrangement is that the student experiences a motivation in being able, at an early stage, to perform the final phase in a behavior series. This is especially true if this final phase has the character of a desirable goal.

Ruleg (1) A sequence in programmed study material in which new information is introduced according to the formula: first general rule (law, principle, definition or the like), then a complete example, then an incomplete example for the student to work with. That is, the "ru" (rule) precedes the "eg's" (examples). (2) The term is also used (and originally) for a special system of constructing self-instructional study material, developed by Evans, Homme & Glaser (1962), in which the use of the sequence type metntioned above is an important part, but that also has a series of other typical characteristics. - The term "Egrul" is sometimes used for the more inductive sequence (leading the student through some examples before letting him try to formulate the rule himself); hence, the opposite of Ruleg (1).

Scrambled book A programmed textbook characterized by branches incorporated for the purpose of explaining errors. After going through the first information unit, the student must answer a multiple-choice question as control. The answer he chooses determines the page to which he is directed next. If he has chosen the right answer, the new page presents new information and a new task. If his choice has been wrong, on the other hand, he is given an explanation of his error on the new page and as a rule is directed back to the original task page with the instruction to choose a better answer.

Skinner model A course layout which combines a non-branching flow model (linear model where all students go through the same study material in the same order) with a demand for self-constructed responses. The name derives from B. F. Skinner but is somewhat misleading, since Skinner worked with other models also (among them models with extreme individualization of the repetition procedure).

Speed of information A measure of the effective communication of information ("reduction of uncertainty" for the student; compare this word) in relation to the period of time needed for learning. A measure which indicates, in other words, how much effective information the students assimilate per time unit in going through a certain instructional program, and which thereby gives us one meaningful point of departure in evaluating the effectivity of the program.

Spiral progression A sequence of material set up in such a way that the student goes through different subject-matter areas at a lower grade of difficulty before going on to these different areas of material at a higher level of difficulty.

Stimule See "Didule"

Student-paced flow of information The fact that in programmed instruction the student himself most often decides when a new information unit shall be presented to him (as opposed to lesson and lecture teaching where the flow of information is directed by the teacher).

Surplus learning An extra dividend in the learning result, over and above the specific knowledges and skills. Acquired study techniques, attitudes to a certain field of knowledge and the like may be considered "surplus learning".

System analysis A general term for various types of studies on the different components of a system. A system analysis is usually a very important preparatory phase of work in constructing an instructional methods-and-materials system and then includes such things as goal analysis, student analysis, situation analysis, subject-matter analysis and media analysis.

System synthesis A general term for putting together the components of a system. In constructing a methods-and-materials system for instruction, we usually start with an analytic phase (studying the prerequisites one by one) and follow up with a system synthesis in constructing a preliminary version.

"Tail model" A layout of auto-instructional material which presupposes that the student has already in some other way acquired the main points of the study material (for instance, in an ordinary textbook). The auto-instructional material functions then as post-study control and remedial teaching where needed. The foremost advocate of this partial type of auto-instruction has been S. L. Pressey.

Target population That group of students to which a certain study material is directed.

Terminal behavior or terminal behavior repertoire The knowledge and skill which the student should be able to display after going through a certain study material. A careful mapping of the desired terminal behavior is an important part of the preliminary work in drawing up a programmed course. (Compare "Goal analysis".)

Terminal stimulus The unprompted question or problem to which the student is taught to respond. (Compare the word "prompt".)

Terminal test An examination which makes it possible to ascertain whether a student who has gone through a given study material can demonstrate the knowledge and skill which are specified in the goal description of the course. (Also: "criterion test".)

Thematic prompts Stimuli which with the help of meaningful (content-related) associations make it easier for the student to give the correct answer. Examples and logical deductions can be counted as belonging to this group of hints.

T-tabulation A tabulation (used as an aid in goal analysis), with the help of which the programmer gets a survey of (a) the specific "terminal behaviors" (see that word) which the student should display at the end of the course as well as (b) under which detailed conditions this is to take place. The T-tabulation consists of a cross-classification chart of specific terminal behaviors on the one side and a list of "specifications" (which describe approved aids, acceptable minimum speed, maximum tolerance of error etc.) on the other. (For illustration, see chapter 5 above.)

Tutofor A general term for an aid in presenting a teaching program to a student. This aid, or tool, can take the form of a printed book, which we then refer to as a "programmed textbook" (or "tutoprint"). But it can also be of mechanical character, and we then speak of a "teaching machine" (or "tutomat").

Vertical panel-book A programmed textbook, in which the pages are divided into panels which (in contrast to the case of the "horizontal panel-book", compare that word) are read in the usual reading sequence (from top to bottom of the page). Each panel usually contains a short information unit to which is linked a response request. A key answer is also usually part of the instructional micro-unit, and this can be placed in the same panel ("simple marginal-key"), in connection with a later panel ("delayed marginal-key"), or also on the back of the information page.

Weltner technique A technique for assessment of the instructional effect of a particular program, utilizing the difference in subjective information of a basic text before and after the instruction. The difference is established by means of a specific technique of text predictions, similar to one introduced by C. E. Shannon. (For details, see Weltner, 1966.)

Zig-zag book A programmed textbook in which the pages are divided into cut sections which can be turned forward and backward independently of each other. This provides a better stimulus situation than that offered by the ordinary forms of horizontal and vertical panel-books. That is, the key answer is not in the visual field when the student answers the problem, but both the problem and the key answer are in the visual field after the student has answered. (For illustration, see chapter 8 above, especially Figures 8. 7 and 8. 8.)

APPENDIX II

PROGRAMMED INSTRUCTION AND EDUCATIONAL TECHNOLOGY: A SELECTIVE BIBLIOGRAPHIC GUIDE

Introductory note

The present bibliographic guide is selective, that is, it does not at all aim at being comprehensive. Instead, we have tried to cover a number of basic topics and, within each of these, to list a sample of pertinent references. For some of the categories we have also attempted to give some further guidance by means of brief comments.

Outline of contents

1. SURVEY PUBLICATIONS

1.1 Integrated Surveys in Book Form

Becker, J. L. A programed guide to writing auto-instructional programs. Camden, N. J.: RCA, 1963.

Bernmalm, S. Programmerad undervisning: En granskning och sammanställning av amerikanska forskningsrapporter. Göteborg: IPU, 1965.

Bjerstedt, Å. Programmerad undervisning, språklaboratorier och gruppdynamisk kartläggning. (4th ed.) Lund: Gleerups, 1968.

Bjerstedt, Å. Självinstruerande studiematerial. (Rev. ed.) Lund: Uniskol, 1969.

Brethower, D. M. Programed instruction: a manual of programing techniques. Chicago: Educational Methods, 1963.

Brethower, D. M. Programed instruction: a manual of programing techniques. Arbor, Mich.: Ann Arbor Publishers, 1965.

Dodd, B. Programmed instruction for industrial training. London: Heinemann, 1967.

Espich, J. E. & Williams, B. Developing programmed instructional materials. Palo Alto, Calif.: Fearon, 1967.

Filby, Y. Teaching machines. A review of theory and research. Copenhagen: Munksgaard, 1961.

Fry, E. B. Teaching machines and programmed instruction. An introduction. New York: McGraw-Hill, 1963.

Garner, W. L. Programed instruction. New York: Center for Applied Research in Education, 1966.

Green, E. J. The learning process and programmed instruction. New York: Holt, 1962.

Hughes, J. L. Programed instruction for schools and industry. Chicago: Science Research Associates, 1962.

Kay, H., Dodd, B. & Sime, M. Teaching machines and programmed instruction. Harmondsworth: Penguin Books, 1968.

Leedham, J. & Unwin, D. Programmed learning in schools. London: Longmans, 1965.

Leith, G. O. M. et al. A handbook of programmed learning. (Rev. ed.) Birmingham, England: University of Birmingham, 1966.

Leith, G. O. M. Seconds thoughts on programmed learning. London: NCET, 1969.

Lysaught, J. P. & Williams, C. M. A guide to programmed instruction. New York: Wiley, 1963.

Markle, S. M., Eigen, L. D. & Komoski, P. K. A programed primer on programing. Vol. I. Principles. New York: Center for Programed Instruction, 1961.

Markle, S. M. A programed primer on programing. Vol. II. Practical problems. New York: Center for programed Instruction, 1961.

Markle, S.M. Good frames and bad. (Rev. ed.) New York: Wiley, 1969.

Oficsh, G.D. Programed instruction: A guide for management. New York: Amer. Management Ass., 1965.

Rowntree, D. Basically branching. A handbook for programmers. London: MacDonald, 1966.

Schiefele, H. & Huber, G.L. Programmierte Unterweisung - programmiert. München: Ehrenwirth, 1969.

Schramm, W. Programed instruction today and tomorrow. New York: The Fund for the Advancement of Education, 1962.

Shirley-Smith, K. Programmed learning in integrated industrial training. London: Gower Pr., 1963.

Silvern, L.C. Fundamentals of teaching machine and programmed learning systems. Los Angeles: Education and Training Consultants, 1964.

Skinner, B.F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

Stiernborg, M. Programmerad undervisning - analys och översikt. Stockholm: Dept. Educ., 1966.

Stolurow, L.M. Teaching by machine. Washington: U.S. Government Printing Office, 1961.

Taber, J.I., Glaser, R. & Schaefer, H.H. Learning and programmed instruction. Reading, Mass.: Addison-Wesley, 1965.

Thomas, C.A. et al. Programmed learning in perspective. A guide to programme writing. Barking, Essex: Adelphi, 1963.

Undervisningsmaskiner och programmerat studiematerial. Kungl. Skolöverstyrelsen: Utredningar i skolfrågor 11, 1963.

Zielinski, J. & Schöler, W. Methodik des programmierten Unterrichts. Ratingen: Henn, 1965.

Zielke, W. Programmierte Instruktion in der Wirtschaft. München: Verlag Moderne Industrie, 1970.

Note: Some of these books are very elementary and popular in style; whereas others - for instance, the publications by Filby and Stolurow - are more adapted to readers with some familiarity of the terminology and style of scientific psychology. In some cases the presentation form is that of a programmed textbook and, hence, functions both as introduction and illustration (for instance, Becker, 1963; Brethower, 1963; Markle et al., 1961; Schiefele & Huber, 1969). Those who are most interested in the program construction process will find Fry (1963), Lysaught & Williams (1963), and Markle (1969) to be some of the important references. The aspects of learning theory are given special emphasis by Green (1962).

1.2 Survey Articles

Holland, J.G. Research on programming variables. In: R. Glaser (Ed.), Teaching machines and programmed learning, II. Washington, D.C.: NEA, 1965. Pp. 66-117.

Klaus, D.J. An analysis of programming techniques. In: R. Glaser (Ed.), Teaching machines and programmed learning, II. Washington, D.C.: NEA, 1965. Pp. 118-161.

Leib, J.W. et al. Teaching machines and programmed instruction. Psychol. Bull., 1967, 67, 12-26.

Lumsdaine, A.A. Instruments and media of instruction. In: N.L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 583-682.

Morrill, C.S. Teaching machines: A review. Psychol. Bull., 1961, 58, 363-375.

Skinner, B.F. Why we need teaching machines. Harvard educ. Rev., 1961, 31, 377-398.

Skinner, B.F. Reflections on a decade of teaching machines. In: R. Glaser (Ed.), Teaching machines and programmed learning, II. Washington, D.C.: NEA, 1965. Pp 5-20.

Spaulding, S. Advanced educational technologies. Prospects in Education (Unesco), 1970, 1 (3), 9-19.

1.3 Introductory Films and Filmstrips

"Example of a teaching machine program". Filmstrip prepared by D.H. Luxton & R.E. Corrigan. Color, 69 frames. Pasadena, Calif.: Basic Skill Films.

"Learning and behavior". Interviews with B.F. Skinner and R.J. Herrnstein at Harvard's Psychological Laboratory (CBS). B & w, 16 mm sound, 27 min. New York: Carousel Films.

"One step at a time". Produced by D. Klaus et al., American Institute for Research. Color, 16 mm sound, 28 min. Pittsburgh, Penna.: AIR.

"Programming is a process". Filmstrip (color, 80 frames) and sound tape. Produced by S.M. Markle and Ph. W. Tiemann. Chicago: Tiemann Associates.

"Selection and use of programmed materials". Filmstrip (color, 64 frames) and sound record, 16 min. Washington: NEA.

"Teaching machines". Filmstrip prepared by W.H. Allen. Techn. advisor: A.A. Lumsdaine. Color, 62 frames. Pasadena, Calif.: Basic Skill Films.

"Teaching machines and programmed learning". Demonstrations and comments by B.F. Skinner, A.A. Lumsdaine, and R. Glaser (NEA). B & w, 16 mm sound, 29 min. Washington: Norwood Films.

1.4 Inventory-Type Surveys

Note: Many bibliographical surveys have appeared, but they too often have the disadvantage of becoming obsolescent almost before leaving the printing press, due to the rapid increase in literature. Only some examples are therefore listed here. For information about the most recent publications, the reader should consult the latest issues of the journals in the field, where surveys and reviews often appear. - As is apparent from the titles, the inventories cover different things. Inventories of machines (not quite up to date any longer, of course) are given by Finn & Perrin and by Ross, whereas program inventories are presented in "Programs, '62", "Programs, '63" etc. and in Hendershot. Finished research is presented in Schramm. A more general bibliographic aim is represented by Gee.

Bjerstedt, Å. The terminology of programmed instruction. Didakometry, No. 13, 1966.

Cavanagh, P. & Jones, C. (Eds.) Programmes in print 1966. London: Association for Programmed Learning, 1966.

Cavanagh, P. & Jones, C. (Eds.) Yearbook of educational and instructional technology 1969/70. London: Cornmarket, 1969.

Center for Programed Instruction. Programs, '62. A guide to programed instructional materials available to educators by September 1962. Washington: U.S. Government Printing Office, 1962.

Center for Programed Instruction. Programs, '63. A guide to programed instructional materials available to educators by September 1963. Washington: U.S. Government Printing Office, 1963.

Center for Programed Instruction. Programed instruction materials 1964-'65. A guide to programed instruction materials available for use in elementary and secondary schools as of April, 1965. New York: CPI, Columbia Univer., 1965.

Centre de Documentation sur l'Enseignement Programmé. Bibliographie sur l'enseignement programmé et les machines à enseigner. I-II. Paris: Institut Pédagogique National, 1966.

Centre de Documentation sur l'Enseignement Programmé. Catalogue analytique des cours programmés. (Langue française.) Paris: Institut Pédagogique National, 1967.

Committee for Out-of-school Education, Council for Cultural Co-operation. Programmed instruction: Institutions and their activities. Strasbourg: Council of Europe, 1970.

Deutschsprachige Lehrprogramme. Stand: Mai 1969. Berlin: Pädagogisches Zentrum, 1969.

Finn, J.D. & Perrin, D.G. Teaching machines and programed learning. A survey of the industry. Washington: U.S. Government Printing Office, 1962.

Gee, R.D. Teaching machines and programmed learning. A guide to the literature. Hatfield: HERTIS, 1965.

Gesellschaft für programmierte Instruktion. Kontakt-Register der pädagogischen Technik. (3. Aufl.) Quickborn: Schnelle, 1970.

Hendershot, C.H. Programmed learning: A bibliography of programs and presentation devices. (4th ed. & suppl. 1-6.) Bay City, Mich.: Author, 1970.

Hintermaier, R. Ed.) Lernprogramme ''68. München: Ehrenwirth, 1968.

Hintermaier, R. (Ed.) Lernprogramme ''70. München: Ehrenwirth, 1970.

Müller, D.D. Jahreskatalog: Kybernetik, Automation, Programmierter Unterricht, Grenzgebiete. Berlin: Elwert und Meurer, 1968.

National Society for Programmed Instruction. Directory of members. San Antonio, Texas: NSPI, 1971.

Programmed instruction guide. (2nd ed.) Newburyport, Mass.: Entelek, 1968.

Programmed learning and teaching machines: Bibliographical references. Antwerpen: International Audio-Visual Technical Centre. (Undated issues.)

Ross, W.L. m.fl. Teaching machines: Industry survey and buyers' guide. New York: Center for Programed Instruction, 1962.

Schramm, W. The research on programed instruction. Washington: U.S. Government Printing Office, 1964.

Spaulding, S. Programmed instruction: An international directory. Paris: Unesco, 1967.

2. COLLECTIONS

2.1 Collections of Articles in Book Form

Aagaard, K., Dohn, H. & Marckmann, W. (Eds.) Konference om programmeret undervisning. Copenhagen: Lærerforeningernes Materialudvalg, 1968.

Bung, K. (Ed.) Programmed learning and the language laboratory (1). London: Longmac, 1968.

Coulson, J.E. (Ed.) Programmed learning and computer-based instruction. New York: Wiley, 1962.

Dahllöf, U. & Wallin, E. (Eds.) Läromedelsforskning och undervisningsplanering. Stockholm: Nord. utredningsserie, 1969.

DeCecco, J. P. (Ed.) Educational technology. Readings in programmed instruction. New York: Holt, 1964.

Dolmatch, Th. B. m. fl. (Eds.) Revolution in training. Programed instruction in industry. New York: American Management Ass., 1962.

Dunn, W.R. & Holroyd, L. (Eds.) Aspects of educational technology. Vol. 2. London: Methuen, 1969.

Filep, R. T. (Ed.) Prospectives in programing. New York: Macmillan, 1963.

Frank, H. (Ed.) Lehrmaschinen in kybernetischer und pädagogischer Sicht. 1-4. München: Oldenbourg, 1963-66.

Galanter, E. (Ed.) Automatic teaching: The state of the art. New York: Wiley, 1959.

Gesellschaft für Programmierte Instruktion. 8. Internationales Symposion über Programmierte Instruktion und Lehrmaschinen: Kurzfassungen. Berlin: GPI, 1970.

Glaser, R. (Ed.) Teaching machines and programmed learning, II: Data and directions. Washington: NEA, 1965.

Goldsmith, M. (Ed.) Mechanisation in the classroom. An introduction to teaching machines and programmed learning. London: Souvenir Pr., 1963.

Goodman, E.H. (Ed.) Automated education handbook. Detroit, Mich.: Automated Education Center, 1965.

Hughes, J.L. (Ed.) Programed learning: A critical evaluation. Chicago: Educational Methods, 1963.

Lambert, Ph. (Ed.) The teacher and the machine. Madison, Wisc.: Dembar, 1962.

La recherche en enseignement programmé. Paris: Dunod, 1969.

Lehnert, U. (Ed.) Elektronische Datenverarbeitung in Schule und Ausbildung. München: Oldenbourg, 1970.

Lumsdaine, A.A. (Ed.) Student response in programmed instruction. Washington: National Academy of Sciences, 1961.

Lumsdaine, A.A. & Glaser, R. (Eds.) Teaching machines and programmed learning: A source book. Washington: NEA, 1960.

Lysaught, J.P. (Ed.) Programmed learning. Evolving principles and industrial applications. Ann Arbor, Mich.: The Foundation for Research on Human Behavior, 1961.

Mann, A.P. & Brunstrom, C.K. (Eds.) Aspects of educational technology. Vol. 3. London: Pitman, 1969.

Margulies, S. & Eigen, L.D. (Eds.) Applied programed instruction. New York: Wiley, 1962.

Ofiesh, G.D. & Meierhenry, W.C. (Eds.) Trends in programmed instruction. Washington: NEA, 1964.

Praxis und Perspektiven des programmierten Unterrichts. (Referate des 3. Nürtinger Symposions über Lehrmaschinen.) Quickborn: Schnelle, 1965.

Praxis und Perspektiven des programmierten Unterrichts, Band II. Referate des V. Symposions über Lehrmaschinen.) Quickborn: Schnelle, 1967.

Programmierter Unterricht und Lehrmaschinen. Bericht. Internationale Konferenz. Berlin: Pädag. Zentrum, 1964.

Rollett, B. (Ed.) Praxis und Theorie des programmierten Unterrichtes. Stuttgart: Klett, 1970.

Rollett, B. & Weltner, K. (Eds.) Perspektiven des programmiertes Unterrichts. Wien: Österreichischer Bundesverlag, 1970.

Rollett, B. & Weltner, K. Fortschritte und Ergebnisse der Unterrichtstechnologie. München: Ehrenwirth, 1971.

Roucek, J.S. (Ed.) Programmed teaching. New York: Philosophical Library, 1965.

Schestakow, A.W. (Ed.) Programmiertes Lernen und Lehrmaschinen. (Transl. fr. Russian.) Berlin: VEB Verlag Technik, 1965.

Smith, W.I. & Moore, J.W. (Eds.) Programmed learning: Theory and research. Princeton, N.J.: Nostrand, 1962

Teal, G.E. (Ed.) Programmed instruction in industry and education. Stamford, Conn.: Public Service Research, 1963.

Tobin, M.J. (Ed.) Problems and methods in programmed learning, 1-5. Birmingham, England: The National Centre for Programmed Learning, 1967.

Unwin, D. & Leedham, J. (Eds.) Aspects of educational technology. Vol. 1. London: Methuen, 1967.

Váňa, J. & Tollingerová, D. (Eds.) Programované učeni. (Programmed learning and teaching machines, I-II. From the Liblice Conference, 1965.) Praha: Pedagogický ústav JAK CSAV, 1966.

Wallin, E. (Ed.) Undervisning - konst eller teknik? Stockholm: Almqvist & Wiksell, 1968.

Zifreund, W. (Ed.) Schulmodelle, programmierte Instruktionen und technische Medien. München: Ehrenwirth, 1968.

Note: Among the many collections of articles available, the early one by Lumdaine & Glaser (1960) is probably still one of the best sources for anyone who wants to have a survey of the brief history of programmed instruction and to read some of the most important early articles by the originators in the field: Pressey, Skinner, Crowder, and others. This book is also a good bibliographic source for "old" publications (up to 1960).

A continuation volume, containing among other things a retrospective essay by B.F. Skinner, is Glaser (1965). The development in Great Britain can be followed in, for example, Unwin & Leedham (1967), Dunn & Holroyd (1969), and Mann & Brunstrom (1969). The development in German-speaking countries is mirrored in the many reports from the symposia of Gesellschaft für programmierte Instruktion, for example: Frank, 1963-66; Praxis und Perspektiven des programmierten Unterrichts, 1965 and 1967; Rollett, 1970; Rollett & Weltner, 1970 and 1971.

2.2 Journals and Bulletins

AID: Auto-Instructional Devices. Has been published by Institute of International Research and Development, Lubbock, Texas. (Cf. NSPI Journal below)

Audiovisual Instruction. Ten issues per year. Department of Audiovisual Instruction, NEA, Washington.

Automated Education Letter. Monthly (since 1965). Automated Education Center, Detroit. (Supplements E.H. Goodman (Ed.), Automated Education Handbook. Detroit: AE Center, 1965.)

AV Communication Review. Quarterly. Department of Audiovisual Instruction, NEA, Washington.

Bibliographie: Programmierter Unterricht. Berlin: Pädagogisches Zentrum.

Deutsche Lehrprogramme für Schule und Praxis. Has been published quarterly by Manz Verlag, München, Germany.

Didakometry. Mimeographed bulletin. Department of Educational and Psychological Research, School of Education, Malmö, Sweden.

Educational Technology. (Formerly: Teaching Aids News.) Monthly. Educational News Service, Saddle Brook, New Jersey.

Enseignement Programmé. Quarterly. Published by Centre de Documentation sur l'Enseignement Programmé, Paris.

Journal of Educational Technology. Published three times a year (since 1970) by the National Council for Educational Technology, England.

Journal of Programmed Instruction. Quarterly. Has been published by Center for Programmed Instruction, Teachers College, Columbia University, New York.

La Cybernétique et la Pédagogie Cybernétique. L'Association de Pédagogie Cybernétique, Paris.

L'Enseignement Programmé. Quarterly. Institut Pédagogique National, Paris.

NSPI Journal. Ten issues per year. The National Society for Programmed Instruction, USA. A combination of AID (cf. above) and the earlier NSPI Newsletter.

Programmed Instruction. Nine issues per year. Has been published by Center for Programmed Instruction, Teachers College, Columbia University, New York.

Programmed Learning and Educational Technology. Association for Programmed Learning, London.

Programmiertes Lernen, Unterrichtstechnologie und Unterrichtsforschung. Quarterly. Cornelsen, Berlin.

Visual Education (Incorporating "Programmed Learning News") Monthly. London: National Committee for Audio-Visual Aids in Education.

Zentralblatt der Gesellschaft für programmierte Instruktion. Abstract journal. Published by GPI, Berlin.

Note: As is obvious from the titles, "AV Communication Review" and "Audiovisual Instruction" are designed to cover the total audio-visual field, but these journals have published much material of interest in the area of didactic programming. Several differences in style of presentation are seen in the group of journals mentioned above. For instance, "AV Communication Review", and "Programmed Learning" favor strict scientific reports, whereas "NSPI Journal", for example, has the style of an information bulletin, in which news and brief notes of various kinds take up more space. The journal activity has been very lively, but it has also shown signs of instability. Several of the "early" journals have already been discontinued (e.g., AID, Automated Teaching Bulletin, Deutsche Lehrprogramme, Journal of Programmed Instruction, and Programmed Instruction). - For more detailed information on U.S. periodicals within the area of educational technology, see: Assmann, I. Bibliographie amerikanischer Zeitschriften aus dem Bereich der Bildungstechnologie. Programmiertes Lernen, 1970, 7, 174-179.

3. SPECIAL-TOPIC PUBLICATIONS

3.1 The Pre-Writing Process: Introductory Analyses and Decisions ("System Analysis")

3.1.1 Analyses of terminal objectives, target populations, and subject-matter structure

Bjerstedt, Å. Goal seeking, goal focusing, and goal adjustment. Progr. Learn. educ. Technol., 1970, 7, 268-279.

Bloom, B.J. (Ed.) Taxonomy of educational objectives. New York: Longmans, 1956.

Boeckmann, K. Basaltext und operationale Lernzieldefinition - Eine vergleichende Betrachtung ihrer Möglichkeiten. In: B. Rollett & K. Weltner (Eds.) Fortschritte und Ergebnisse der Unterrichtstechnologie. München: Ehrenwirth, 1971. Pp. 17-25.

Bullock, D.H. Strukturanalys av en undervisningsenhet. In: E. Wallin (Ed.), Undervisning - konst eller teknik? Stockholm: Almqvist & Wiksell, 1968. Pp. 74-87.

Davies, I.K. The mathetics style of programming. In: K. Bung (Ed.), Programmed learning and the language laboratory (I). London: Longmac, 1968. Pp. 29-50.

Evans, J.I., Homme, L.E. & Glaser, R. The Ruleg system for the construction of programmed verbal learning sequences. J. educ. Res., 1962, 55, 513-518.

Flechsig, K.-H. Probleme der Entscheidung über Lernziele. Programmiertes Lernen, 1970, 7, 1-32.

Frank, H. & Graf, K.D. ALZUDI - Beispiel einer formalen Didaktik. Z. erziehungswiss. Forsch., 1967, 1 (1), 27-34.

Gagné, R.M. The analysis of instructional objectives for the design of instruction. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 21-65.

Glaser, R. Some research problems in automated instruction: instructional programming and subject-matter structure. In: J.E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962. Pp. 67-85.

Hartley, J. Factors affecting the efficiency of learning from programmed instruction. (I). Visual Education, 1971, May, 33-35.

Kibler, R.J., Barker, L.L. & Miles, D.T. Behavioral objectives and instruction. Boston: Allyn & Bacon, 1970.

Leith, G.O.M. Learning and personality. In: W.R. Dunn & C. Holroyd (Eds.), Aspects of educational technology, Vol. 2. Methuen, 1969. Pp. 101-110.

Mager, R.F. Preparing objectives for programmed instruction. San Francisco: Fearon, 1961.

von Mentzer, C.H. Studier i empirisk målanalys. Stockholm: School of Education, 1968.

Miller, R.B. Task description and analysis. In: R.M. Gagné (Ed.) Psychological principles in system development. New York: Holt, 1962. Pp. 187-228.

Popham, W.J. et al. Instructional objectives. Chicago: Rand McNally, 1969.

Riedel, H. Psychostruktur: Psychostruktur und Lehrprogrammierung. Quickborn: Schnelle, 1967.

Roe, A. & Moon, H. Analysis of course content for individual learning. Automated Teaching Bulletin, 1960, 1 (3), 3-11.

Weltner, K. Zur Bestimmung der subjektiven Information durch Ratetests. In: Praxis und Perspektiven des programmierten Unterrichts, Band II. Quickborn: Schnelle, 1967. Pp. 69-74.

Weltner, K. Information theory and programmed instruction. Rev. educ. Cybernetics & appl. Linguistics, 1969, 1 (1), 25-41.

Weltner, K. Informationstheorie und Erziehungswissenschaft. Quickborn: Schnelle, 1970.

Weltner, K. Lernziele unter dem Aspekt der Informationstheorie. In: B. Rollett & K. Weltner (Eds.), Fortschritte und Ergebnisse der Unterrichtstechnologie. München: Ehrenwirth, 1971. Pp. 26-35.

Zielke, W. Stichwortschaubild und Faktenanalyse als Hilfsmittel rationellerer Lehrprogramerstellung. In: Gesellschaft für Programmierte Instruktion. 8. Internationales Symposion . . . : Kurzfassungen. Berlin: GPI, 1970. P. 20.

3.1.2 Decisions on Presentation Media

Ahlström, K.-G. & Amcoff, S. Feedback functions in teaching machines. Scand. J. Psychol., 1967, 8, 243-249.

Atkinson, R.C. Learning to read under computer control. Progr. Learn. educ. Technol., 1968, 5, 25-37.

Bjerstedt, Å. Schwierigkeiten und beobachtungstechnologische Möglichkeiten in der Lehrerausbildung. Z. erziehungswiss. Forsch., 1968, 2, 59-82.

Brown, J.W. Student response systems. Audiovisual Instruction, 1963, 8, 214-218.

Bung, K. Programmed learning and the language laboratory (2): Collected papers. London: Longmac, 1967.

Bushnell, D.D. The role of the computer in future instructional systems. Washington: DAVI, 1963.

Goldstein, L.S. & Gotkin, L.G. Teaching machines vs. programed textbooks as presentation modes. Journal of programed Instruction, 1962, 1, 29-36.

Hayes, A.S. Language laboratory facilities. Washington: U.S. Government Printing Office, 1963.

Hocking, E. Language laboratory and language learning. Washington: NEA, 1964.

Holling, K. The feedback classroom in use. In: D. Unwin & J. Leedham (Eds.), Aspects of educational technology. London: Methuen, 1967. Pp. 275-306.

Lewis, B.N. & Pask, G. The theory and practice of adaptive teaching systems. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 213-266.

Malmquist, E. & Grundin, H.U. Olika former för bjudning av programmerat undervisningsmaterial. Forskningsrapporter från Statens Försöksskola i Linköping, Nr 8, 1966.

Pask, G. Teaching machines. In: 2nd International Congress on Cybernetics, Sept. 1958. Namur, Belg.: Assoc. Intern. de Cybernétique. Pp. 962-978.

Pask, G. The teaching machine as a control mechanism. Transactions of the Society of Instrument Technology, June 1960. Pp. 72-89.

Richter, H. Lehrautomaten - Beispiele und Entwicklungstendenzen. In: B. Rollett & K. Weltner (Eds.), Fortschritte und Ergebnisse der Unterrichtstechnologie. München: Ehrenwirth, 1971. Pp. 173-192.

Ryans, D.G., Bushnell, D.D. & Cogswell, J.F. A computer-based laboratory for automation in school systems. Santa Monica, Calif.: System Development Corp., 1962.

Stolurow, L.M. & Davis, D. Teaching machines and computer-based systems. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 162-212.

3.2 Problems Related to the Writing Process ("System Synthesis")

3.2.1 The problem of "flow structure"

Chapman, B. Phases of mathetical procedure. Programed Instruction, 1963-64, 3 (3-4), 6-9.

Coulson, J.E. et al. Effects of branching in a computer controlled autoinstructional device. J. appl. Psychol., 1962, 46, 389-392.

Crowder, N.A. On the differences between linear and intrinsic programming. Phi Delta Kappan, 1963, 44, 250-254.

Davies, I.K. Mathetics - a functional approach. In: D. Unwin & J. Leedham (Eds.). Aspects of educational technology. London: Methuen, 1967. Pp. 205-216.

Evans, J.L., Glaser, R. & Homme, L.E. An investigation of "teaching machine" variables using learning programs in symbolic logic. J. educ. Res., 1962, 55, 433-452.

Gilbert, Th.F. Mathetics: An explicit theory for the design of teaching programmes. (Rev. educ. Cybernetics & appl. Linguistics. Supplement 1.) London: Longmac, 1969.

Graf, K.-D. Algoritmische Zuordningsdidaktik und Dialog-Didaktik. In: U. Lehnert (Ed.), Elektronische Datenverarbeitung in Schule und Ausbildung. München: Oldenbourg, 1970. Pp. 64-73.

Holland, J.G. & Porter, D. The influence of repetition of incorrectly answered items in a teaching-machine program. Journal of the experimental Analysis of Behavior, 1961, 4, 305-307.

Jivén, L.M. & Öreberg, C. The IMU Project: Preliminary plan for investigating the effects of a system for individualized mathematics teaching. Didakometry, No. 22, 1968.

Larkin, T.C. & Leith, G.O.M. The effects of linear and branching methods of programmed instruction on learning and retention of a topic in elementary science. Programmed Learning, 1964, 1, 12-16.

Leith, G.O.M. Conflict and interference. Programmed Learning, 1971, 8, 41-50.

Mager, R.F. & Clark, C. Explorations in student-controlled instruction. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 235-238.

Pennington, D.F. & Slack, C.W. The mathetical design of effective lessons. In: S. Margulies & L.D. Eigen (Eds.), Applied programed instruction. New York: Wiley, 1962. Pp. 298-310.

Pressey, S.L. Basic unresolved teaching-machine problems. Theory into Practice, 1962, 1, 30-37.

Pressey, S.L. Autopresentation vs. autoelucidation. Programed Instruction, 1963, 2 (4), 6-7.

Roe, A. A comparison of branching methods for programmed learning. J. educ. Res., 1962, 55, 407-416.

Schröter, G. Das zweistufige Programm als integrierender Übergang vom linearen zum verzweigten System. In: B. Rollett (Ed.), Praxis und Theorie des programmierten Unterrichtes. Stuttgart: Klett, 1970. Pp. 32-36.

Shettel, H.H. Individual differences in subject matter knowledge and programmed instructional format. Pittsburgh: AIR, 1963.

Silberman, H.F. et al. Fixed sequence versus branching autoinstructional methods. J. educ. Psychol., 1961, 52, 166-172.

Smallwood, R.D. A decision structure for teaching machines. Cambridge. Mass.: MIT Pr., 1962.

3.2.2 The characteristics of the single instructional units

Barlow, J.A. Conversational chaining in teaching machine programs. Psychol. Rep., 1960, 7, 187-193.

Basescu, B. More on multiple-choice frames. Programed Instruction, 1962, 2 (1), 3.

Basescu, B. & Horn, R.E. A response analysis system. Programed Instruction, 1964, 4 (3), 6-9.

Brethower, D.M. et al. Programmed learning: A practicum. Ann Arbor, Mich.: Ann Arbor Publ., 1965.

Eigen, L.D. & Margulies, S. Response characteristics as a function of information level. Journal of programed Instruction, 1963, 2 (1), 45-54.

Ferraro, D.P. et al. Some considerations on the main features of programmed instruction. In. G.E. Teal (Ed.), Programmed instruction in industry and education. Stamford, Conn.: Public Service Research, 1963. Pp. 43-67.

Goldbeck, R.A. & Campbell, V.N. The effects of response mode and response difficulty on programed learning. J. educ. Psychol., 1962, 53, 110-118.

Grace, H.A. Programming and readability. Calif. J. educ. Res., 1963, 14, 164-166.

Klaus, D.J. The art of auto-instructional programming. AV Communication Review, 1961, 9, 130-142.

Krumboltz, J.D. & Kiesler, Ch.A. The partial reinforcement paradigm and programed instruction. J. progr. Instr., 1965, 3 (2), 9-14.

Krumboltz, J.D. & Weisman, R.G. The effect of overt versus covert responding to programed instruction on immediate and delayed retention. J. educ. Psychol., 1962, 53, 89-92.

Ludwig, E.H. Die Technik zur Herstellung von Lehrprogrammen. Ratingen: Henn, 1965.

Margulies, S. Some general rules of frame construction. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 185-188.

Roe, A. Automated teaching methods using linear programs. J. appl. Psychol., 1962, 46, 198-201.

Schaefer, H.H. A vocabulary program using "language redundancy". Journal of programed Instruction, 1963, 2 (3), 9-16.

Trittipoe, W.J., Trittipoe, T.G. & Hahn, C.P. The effectiveness of three programed-learning frame styles. Journal of programed Instruction, 1963, 2 (1), 35-43.

3.3 The Post-Writing Process ("System Modification and Evaluation")

AERA-APA-DAVI Joint Committee on Programed Instruction and Teaching Machines. Criteria for assessing programed instructional materials. Audiovisual Instruction, 1963, 8, 84-89.

AERA-APA-DAVI Joint Committee on Programmed Instruction and Teaching Machines. Recommendations for reporting the effectiveness of programmed-instruction materials. NSPI J., 1966, 5 (3), 3-9.

Bjerstedt, Å. Mapping the effect-structure of self-instructional materials. Programmed Learning, 1965, 2, 99-109.

Bjerstedt, Å. Mapping the pheno-structure of self-instructional materials. Progr. Learn. educ. Technol., 1967, 4, 87-102.

Bjerstedt, Å. Phäno-Struktur und Effekt-Struktur didaktischer Sequenzen: Beiträge zur Analyse und Beurteilung von Lehrprogrammen. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht, 4. Stuttgart: Klett, 1966. Pp. 17-39.

Bloom, B.S., Hastings, J.Th. & Madaus, G.F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

Cook, D.A. Studying the performance of a program. Programed Instruction, 1962, 2 (2), 4-8.

Cox, A. et al. Suggested criteria for selection of programed materials. In: Summary Report of the Eighth Lake Okoboji Audiovisual Leadership Conference. Iowa City: State Univer. Iowa, 1962. Pp. 80-88.

Geis, G.L. Some considerations in the evaluation of programs. AV Communication Review, 1962, 10, 64-69.

Glaser, R. Instructional technology and the measurement of learning outcomes: Some questions. Amer. Psychologist, 1963, 18, 519-521.

Gordon, J.M. The evaluative response: A rate problem. AID, 1963, 3, 240.

Holland, J.G. Evaluating teaching machines and programs. Teachers College Record, 1961, 63, 56-65.

Jacobs, P.I. Item difficulty and programed learning. Journal of programed Instruction, 1963, 2 (2), 21-38.

Jacobs, P.I., Maier, M.H. & Stolurow, L.M. A guide to evaluating self-instructional programs, New York: Holt, 1966.

Lumsdaine, A.A. Some problems in assessing instructional programs. In: R.T. Filep (Ed.), Prospectives in programing. New York: Macmillan, 1963. Pp. 228-262.

Lumsdaine, A.A. Assessing the effectiveness of instructional programs. In: R. Glaser (Ed.), Teaching machines and programed learning, II. Washington, D.C.: NEA, 1965. Pp. 267-320.

Markle, S.M. Faulty branching frames: A model for maximizing feedback to the programer. Programed Instruction, 1963, 3 (1), 4-8.

Markle, S.M. The lowest common denominator: A persistent problem in programing. Programed Instruction, 1963, 2 (3), 4-5.

Rollett, B. Untersuchungen zur semantischen Transinformation. In: B. Rollett & K. Weltner (Eds.), Perspektiven des programmierten Unterrichts. Wien: Österreich. Bundesverlag, 1970. Pp. 70-76.

Rothkopf, E.Z. Criteria for the acceptance of self-instructional programs. In: A.E. Traxler (Ed.), Improving the efficiency and quality of learning. Washington: Amer. Council on Educ., 1962. Pp. 30-38.

Rothkopf, E.Z. Some observations on predicting instructional effectiveness by simple inspection. Journal of programed Instruction, 1963, 2 (2), 19-20.

Silverman, R.E. The evaluation of programmed instruction: A problem in decision making. Psychology in the Schools, 1964, 1, 74-78.

Standards Committee of the NSPI. Checklist for selecting programs. NSPI Journal, 1963, 2 (6), 4.

Vanderschmidt, H. Validation data for programmed texts: A checklist for evaluation of testing. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 210-212.

Wheatly, R.A. Measurement in programmed learning experiments - problems encountered and some possible solutions. In: A.P. Mann & C.K. Brunstrom (Eds.), Aspects of educational technology, III. London: Pitman, 1969. Pp. 115-120.

Zielinski, J. Empirisch-pädagogische Aspekte der Programmtestung und -validierung. Z. erziehungswiss. Forsch., 1967, 1 (3), 38-62.

3.4 Theoretical Discussions

Björkman, M. Inlärningspsykologiska principer i programmerad undervisning - några frågetecken. In: Skolöverstyrelsen, Utredningar i skolfrågor II, 1963. Pp. 96-105.

Cook, D.A. Behavior theory and the automation of instruction. In: J.E. Coulson (Ed.), Programmed learning and computer-based instruction. New York: Wiley, 1962. Pp. 120-128.

Corrigan, R.E. Programmed instruction as a systems approach to education. In: G.D. Ofiesh & W.C. Meierhenry (Eds.), Trends in programmed instruction. Washington: NEA, 1964. Pp. 36-45.

Ely, D.P. (Ed.) The changing role of the audiovisual process in education: A definition and a glossary of related terms. AV Communication Review, 1963, 11, No. 1.

Frank, H. Kybernetik und Lehrmaschinen. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht. München: Oldenbourg, 1963. Pp. 13-26.

Frank, H. Kybernetische Grundlagen der Pädagogik, I-II. (2. Aufl.) Baden-Baden: Agis-Verlag, 1969.

Hilgard, E.R. Issues within learning theory and programmed learning. Psychology in the Schools, 1964, 1, 129-139.

Landa, L.N. Algoritmierung im Unterricht. (Transl. fr. Russian.) Berlin: Volk u. Wissen, 1969.

Landa, L.N. Diagnostik und programmierter Unterricht. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht, 4. Stuttgart: Klett, 1966. Pp. 54-65.

Landa, L.N. The construction of algorithmic and heuristic models of thinking activity and some problems in programmed learning. In: W.R. Dunn & C. Holroyd (Eds.). Aspects of educational technology, Vol. 2. London: Methuen, 1969. Pp. 133-141.

Newman, S.E. A mediation model for paired-associate learning. In: J.P. De Cecco (Ed.), Educational technology. New York: Holt, 1964. Pp. 172-186.

Pressey, L. Teaching machine (and learning theory) crisis. J. appl. Psychol., 1963, 47, 1-6.

Pressey, S.L. Autoinstruction: Perspectives, problems, potentials. In: Theories of learning and instruction (Sixty-third Yearbook). Chicago: The National Society for the Study of Education, 1964. Pp. 354-370.

Skinner, B.F. Verbal behavior. New York: Appleton-Century-Crofts, 1957.

Skinner, B.F. Contingencies of reinforcement: A theoretical analysis. New York: Appleton-Century-Crofts, 1969.

Thompson, R.L. Programed instruction and reinforcement theory: A view from the laboratory. In: R.T. Filep (Ed.), Prospectives in programing. New York: Macmillan, 1963. Pp. 16-35.

3.5 Programming Activities in Various Countries

Bjerstedt, Å. Auto-instructional methods in Great Britain. Audio-Visual Media, 1968, 2 (2), 89-93.

Bjerstedt, Å. Notes on programmed instruction in Sweden. Didakometry, No. 5, 1966.

Bjerstedt, Å. Educational technology in Sweden. Educational Technology, 1969, 9 (11), 48-51.

Bremer, W. et al. Programmierte Instruktion in der Berufsausbildung der USA. Weinheim: Beltz, 1967.

Brown, J.W. & Thornton, J.W., Jr. (Eds.) New media in higher education. Washington: NEA, 1963.

Cavanagh, P. & Jones, C. Further trends in programmes in print in the United Kingdom. Programmed Learning, 1970, 7, 56-62.

Center for Programed Instruction. The use of programed instruction in U.S. schools. Washington: U.S. Government Printing Office, 1963.

Connor, D.V. Educational technology in Australia. Journal of educational Technology, 1970, 1, 207-216.

Décoté, G. L'enseignement programmé en France. (In: Décoté, Vers l'enseignement programmé. Paris: Gauthier-Villars, 1963. Chap. IV:)

Hartley, J. Programmed learning in emerging nations. Programmed Learning, 1964, 1, 21-25.

Hawkridge, D. Programmed learning in Africa. Journal of educational Technology, 1970, 1, 217-225.

Issing, L.J. (Ed.) Der programmierte Unterricht in den USA heute. Weinheim: Beltz, 1967.

Kishi, T. & Numano, I. The present state of programmed instruction in Japan. The Science of Learning (Tokyo), 1967, 3 (1), 1-10.

Komoski, P.K. & Green, E.J. Programmed instruction in West Africa and the Arab States. A report on two training workshops. (Educational Studies and Documents, No. 52.) Paris: Unesco, 1964.

Landa, L.N. Programmed instruction in the Soviet Union. Educational Technology, 1970, 10 (7), 40-45.

Müller, D.D. Five stages of development of programmed instruction in Germany. Progr. Learn. educ. Technol., 1967, 4, 216-222.

Romiszowski, A.J. The international conference on educational technology 1971. Visual Education, 1971, May, 29-33.

Vogt, H. Programmierte Instruktion und Pädagogische Kybernetik im sowjetischen Bildungswesen. Progr. Lern. progr. Unterr., 1967, 4, 54-65.

Vogt, H., Zänker, U. & Schellenberg, E. Pädagogische Kybernetik und programmierte Instruktion in der Sowjetunion und in der DDR. Weinheim: Beltz, 1968.

3.6 From the Discussions and Research on Programmed Instruction and Educational Technology in Sweden

Ahlström, K.-G. SMID - självinstruerande metoder i dövundervisningen. In: E. Wallin (Ed.), Undervisning - konst eller teknik? Stockholm: Almqvist & Wiksell, 1968. Pp. 115-126.

Andersson, B.-E., Parknäs, L. & Wallin, E. Erfarenheter från ett försök med programmerad undervisning. Nordisk Psykologi, 1962, 14, 261-269.

Berglund, L.G. & Jivén, L.M. IMU-projektet: Försök i hjälpklass rörande informationsgivning med två typer av svarsbekräftelse i relation till stegstorlek. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 86, 1969.

Bernmalm, S. Programmerad undervisning vid universitet och högskolor: En granskning av förutsättningarna och ett försök med ett program i njurclearance. Göteborg: Dept. Educ., 1969.

Bjerstedt, Å. Programmerad undervisning, språklaboratorier och gruppdynamisk kartläggning. Lund: Gleerups, 1963. (4th rev. ed., 1968.)

Bjerstedt, Å. Självinstruerande studiematerial. Lund: Uniskol, 1963. (Rev. ed., 1969.)

Bjerstedt, Å. Mapping the pheno-structure of didactic sequences. Didakometry, No. 1, 1964.

Bjerstedt, Å. Mapping the effect-structure of didactic sequences. Didakometry, No. 2, 1964.

Bjerstedt, Å. Undervisningsprogrammering och utvärdering. Stockholm: Sv. Bokförlaget, 1965. (Rev. ed., 1970.)

Bjerstedt, Å. Notes on programmed instruction in Sweden. Didakometry, No. 5, 1966.

Bjerstedt, Å. Phäno-Struktur und Effekt-Struktur didaktischer Sequenzen: Beiträge zur Analyse und Beurteilung von Lehrprogrammen. In: H. Frank (Ed.), Lehrmaschinen in kybernetischer und pädagogischer Sicht, IV. Stuttgart: Klett, 1966. Pp. 17-39.

Bjerstedt, Å. Att konstruera undervisningsprogram: En studieguide. Kompendieserien (Malmö: School of Education), No. 4, 1967.

Bjerstedt, Å. Konsten att vänta på fakta: Två diskussionsinlägg. Pedagogisk-pdykologiska problem (Malmö: School of Education), No. 75, 1968.

Bjerstedt, Å. Educational technology in Sweden. Didakometry, No. 23, 1969.

Bjerstedt, Å. Den pedagogiska processen. Lund: Gleerups, 1970. (Rev. ed., 1971.)

Bjerstedt, Å. Systematic analyses of instructional objectives. Didakometry, No. 28, 1970.

Bjerstedt, Å. System analysis in instructional programming: The initial phases of the program construction process. Didakometry, No. 30, 1971.

Bjerstedt, Å. System synthesis in instructional programming: The intermediate phases of the program construction process. Didakometry, No. 32, 1971.

Edfelt, Å.W. Automatic teaching - principiella synpunkter. Nordisk Psykologi, 1962, 14, 6-20.

Eklund, H. Det programmerade studiematerialets plats inom den obligatoriska skolan. Pedag. Tidskr., 1963, 99, 145-154.

Ekman, B. Teaching geometry in grade four. Uppsala: Sv. Bokförlaget, 1968.

Gestrelius, K. Utbildningsplanering med befattningsanalytiska metoder. Lund: Uniskol, 1970.

Grundin, H.U. Svarskrav och facitsvar vid programmerad undervisning. Forskningsrapport från lärarhögskolan i Linköping (Linköping: School of Education), 1970.

Holmberg, I. Effekten av programmerad undervisning på vissa prestations-, attityd- och beteendevariabler, jämfört med konventionell undervisning. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 18, 1965.

Holmberg, I. Maskinundervisning kombinerad med lärarledd smågruppsundervisning, jämfört med konventionell undervisning. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 30, 1966.

Holmberg, I., Holmberg, L., Jivén, L.M. & Svensson, J.-E. Statistiska grundbegrepp. En programmerad introduktion. Stockholm: Läromedelsförlagen/Sv. Bokförl. 1964. (Rev. ed., 3rd pr. 1970.)

Husén, T. Inlärningsmaskiner och programmerade hjälpmedel. Pedagogisk-Debatt, 1962, 10, 48-57.

Husén, T. Programmerade hjälpmedel i undervisningen. Kungl. Krigsvetenskapsakademiens Handlingar och Tidskrift, 1962, Nr 1, 3-21.

Jivén, L.M. (Ed.) IMU-projektet: Rapport från utprovning av IMU-systemet på grundskolans högstadium (I). Pedagogisk-psykologiska problem (Malmö: School of Education), No. 63, 1968.

Jivén, L.M. & Svensson, J.-E. Erfarenheter från programmerad undervisning i statistik. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 32, 1966.

Klingberg, G. A scheme for the classification of educational objectives. Dept. Educ. Res., Gothenburg School of Education, Res. Bull., No. 5, 1970.

Lane, G. Erfarenheter av ett försök med programmerad undervisning i matematik. Rapp. Pedag. inst., Göteborgs univer., No. 8, 1963.

Lewerth, A. & Stangvik, G. Läsning av socialt viktiga ord. Ett försök med programmerad undervisning för utvecklingsstörda. Rapp. Pedag. inst., Lärarhögskolan i Göteborg, No. 9, 1969.

Lindecrantz, N. Datamaskinförmedlad undervisning. Lund: Studentlitteratur, 1968.

Lindell, E. (Ed.) Ett försök med självinstruerande material i språklaboratorium och i skriftlig programmering. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 53, 1967.

Lindell, E. A system of study materials for the teaching of foreign languages: Principles and examples. Didakometry & Sociometry, 1969, 1, No. 2.

Lindell, E. (Ed.) Att arbeta med ett undervisningssystem: En presentation av UMT-projektets läromedel. Pedagogisk-psykologiska problem (Malmö: School of Education), No. 106, 1970.

Lindell, E. Främmande språk i grundskolan. Lund: Gleerups, 1971.

Lindmark, L. m.fl. Datamaskinförmedlad utbildning. Stockholm: Ecklesiastikdepartementet, 1966.

Lundin, J.E. Utprövning och revision av PI. Undervisningsteknologi, 1969, 4 (4), 16-19.

Malmquist, E. Undervisningsmaskiner och deras användning. Skola och Samhälle, 1961, 42, 175-187.

Malmquist, E. Undervisningsmaskiner och programmerat studiematerial. Svensk Skolledartidning, 1963, 22, 130-145.

Malmquist, E. & Grundin, H.U. Olika former för bjudning av programmerat undervisningsmaterial. Forskningsrapporter från Statens Försöksskola i Linköping, Nr 8, 1966.

von Mentzer, C.H. Studier i empirisk målanalys. Stockholm: School of Education, 1968.

von Mentzer, C.H. Tio delstudier från UME-projektet. En antologi med exempel på problem och metoder i ett "system"-inriktat forsknings- och utvecklingsarbete. Rapp. Pedag. inst., Lärarhögskolan i Stockholm, No. 50, 1970.

Pikas, A. Jämförelser mellan traditionell och programmerad undervisning. Pedagogisk Tidskrift, 1967, 103, 1-10, 41-49, 109-116, 115-167.

Pikas, A. Indoctrination of basic ideas of rational conflict solution. Report, The Institute of Education, Uppsala University, No. 41, 1969.

Stiernborg, M. Utbildning av värnpliktiga utbildningsprogrammerare. Stockholm: Dept. Educ., 1966.

Stukát, K.-G. Ett försök med programmering av svensk språklära. Stockholm: Biblioteksförlaget, 1963.

Stukát, K.-G. Construction and field testing of a grammar program. Programmed Learning, 1965, 2, 14-30.

Stukát, K.-G. Ett experiment rörande kännedom om resultatet och stegstorlek i ett språkläreprogram. Pedag. Forskn., 1965, 9, 93-102.

Tidland, A. Undersökningar i anslutning till undervisningsprogrammet "Engelsk skrivning". Rapp. Pedag. inst., Göteborgs univer., No. 21, 1966.

Undervisningsmaskiner och programmerat studiematerial. Kungl. Skolöverstyrelsen: Utredningar i skolfrågor 11, 1963.

Wallin, E. Om målanalys och målformulering. Rapp. Pedag. inst., Göteborgs univer., No. 49, 1970.

Öreberg, C. IMU-systemet. Undervisningsteknologi, 1968, 3 (1-2), 9-14.