DOCUMENT RESUME

ED 039 756 | 56 | EM 008 092

AUTHOR Spencer, Richard E.
TITLE The Role of Measurement and Evaluation in Instructional Technology.
INSTITUTION Academy for Educational Development, Inc., Washington, D.C.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-8-0571
PUB DATE [70]
NOTE 24p.; This is one of the support papers for "To Improve Learning; a Report to the President and the Congress of the United States by the Commission on Instructional Technology", ED 034 905

EDRS PRICE EDRS Price MF-$0.25 HC-$1.30
DESCRIPTORS *Educational Technology, *Evaluation Criteria, *Measurement Techniques

ABSTRACT

In this paper, an editorial comment on the field of instructional technology as a whole is made, followed by a brief discussion of research findings about instructional technology and assumptions underlying the use of instructional technology. The role of measurement and evaluation in the field are examined, and implications and recommendations for the future of educational technology conclude the paper. (SP)

The Role of Measurement and Evaluation in Instructional Technology

Richard E. Spencer*

There are, in existence, innumerable reviews of the research dealing with instructional technology (Lumsdaine, Torkelson, Allen and Finn). Included under the rubric of instructional technology are educational television, computer assisted instruction, programmed materials, teaching machines, simulation games, instructional films, filmstrips, slides and other materials, textbooks, etc. All such variables, as well as many more, since the introduction of the hornbook in the 16th century, have been researched, compared, evaluated, accepted, rejected, or ignored. Much of the research has been temporal or cyclical in nature; emphasis occurs on one topic or another as a specific piece of equipment is made available in the market place, or as federal funds reward certain types of innovation.

It is not the purpose of this paper to review specific research in the area. It is assumed that the reader can find such material elsewhere. Here, the purpose is to attempt to make an editorial comment on the field as a whole, to indicate some sources of assumptions, causes of the findings, the role of measurement and evaluation in the area, and a look to the future of educational technology.

If one looks at the research, one can easily estimate the popularity of various topics of concern. Discussions of the Big Three--Television, Computer Assisted Instruction (CAI), and Programmed Instruction (PI)--compose a majority of the writings in the area. In general, one finds that the early research efforts in each area tend to be focused on the hardware; i.e. Can TV Teach? Do teaching machines work? Later, concentration turns to theoretical developments, and more system or operational considerations. To a considerable extent, however, the hardware systems of instructional technology have been used in some operational settings with little regard for research findings.

* Richard E. Spencer is head of the Measurement and Research Division, University

All of the studies have some basic assumptions, but it should be apparent that there is a basic problem with instructional technology research and use which has been generally ignored. In the first place, the definition of instructional technology tends to emphasize hardware, generally to the exclusion of instructional content, or teaching strategies or methodologies. These items are variables which exist on ETV, or films, or are included in PI, textbook writing, or any other presentation of instruction. Publications seem to have been concentrated on the assumption that the hardware can do something unique, or operate independently of what someone puts into it. Such research normally results in an all too common finding in this area--no significant difference. The hardware (the media end of instructional technology) as used, results in learning to about the same level as a real, live instructor does.

As one progresses from concentration on the hardware to what is performed with the hardware (sequence of learning, audio-visual stimuli, reinforcement, overlearning, etc.) one finds small but statistically significant differences between experimental and control groups. Unfortunately such gains are usually temporary and disappear upon delayed recall evaluation. Thus, the general state of the art can be defined in either an optimistic or pessimistic way:

Instructional technology can teach <u>just as well</u> as regular, conventional methods; or

instructional technology can teach <u>no better</u> than regular, conventional methods.

Thus far, however, we have not defined "what" is taught, how something is taught, or what learning is considered to be. Secondly, we have not concerned ourselves with student "attitudes" toward the hardware style of teaching. Most of the effort in instructional technology has been put into one-shot programs, often developed for research use only, which rarely see the light of day in a continuous on-going instructional operation. Those that do develop into

operational instructional programs show no better results (in achievement) than does regular instruction, and the students generally dislike it more. Reports of non-involvement, separation from the instructor, uninteresting materials, etc. are common for ETV; for CAI and Programmed Instruction the most common student response is that it is boring. What is the basis for these rather discouraging results? Is instructional technology doomed to play only a supportive role in American education? Why hasn't instructional technology fulfilled its potential?

Some of the assumptions underlying the use of instructional technology could stand critical inspection. There is a considerable "religiosity" associated with instructional technology--those that are in the field seem to believe that the potential is just lying there waiting to be tapped. With only a slight degree of effort, the technology will prove its worth. Faith is built into many research designs, but satisfaction occurs with no significant regularity. An evident assumption appears to be that the media is in fact the message (McLuhan and Fiore). It is certainly true that one can study content, as a student, only according to what content is available. What university courses exist define for the student what he may study--he is, in part at least, led through an a priori system of content. "...the students rate of learning is restricted to the teacher's rate of presenting; the amount to be learned by the student is limited by the amount the teacher presents; and the pattern of instruction by which the student is supposed to learn is determined by the pattern the teacher decides to use for the whole class." (Dial-Access, p. 5) Similarly he may only read from the reservoir of what has been written; he may see films only related to that which film makers consider worth filming. Practical applications of ETV, on college campuses at least, have resulted in the video taping of hours and hours of the head and shoulders of some professor performing precisely the way he does in class. If the media is the message, the message on this TV media is that education is a bore. This attitude has become so fixed in the minds of the students,

perhaps because of the rather obvious comparison to commercial TV, that in order to obtain significant increments in learning with ETV, one has to exert considerable effort, and consider many different variables and factors in the learning situation.

A second apparent assumption is that one must concentrate on the INPUT to the instructional system rather than OUTPUT; i.e. the assumption is made that we know what the objectives of a course really are. The method by which an instructional segment becomes involved with instructional technology is to start with how it is done now--in the regular (i.e. conventional) classroom instructional setting. This methodology of approach may set the stage for the all too popular "no significant differences" that so often result. Instructional technology has been constrained by the already existing structural pattern. Knowledge, or content, has long since been carved up into segments, placed at levels, and defined by a rendition of the Sapir-Whorf hypothesis. As Sapir states (Sapir, p. 578):

> "Language constitutes a sort of logic, a general frame of reference, and so molds the thought of its habitual users."

Such a theory may be applied to an educational system--one which has defined reality and the real world (see Jerome), divided it up into language arts, arithmetic, social studies, physical education, etc.; housed it in a certain manner, supplied the media (i.e. teacher and text), financed it, dated it (from September to June), timed it (8:30 to 3:15) to such an extent that the system becomes a reality as Benjamin so aptly points out in the Saber-toothed Curriculum. It is difficult to develop a independent developmental procedure, and a capability to stand aside and view the system, define its objectives and evaluate its outcomes without the constraint of the existing structure within which instructional technology has been operating. The organization and the institutionalization of education itself has formed the greatest barrier of the use of instructional technology. The research on instructional technology has consistently occurred <u>within</u> the system--affected and controlled by the systemization already in

existence. This system controls what we may experiment upon, how we may experiment, and how we are to treat the results. The potential for significant differences may not exist; or those things upon which we determine significant differences aren't the important ones. The system itself may be the most important variable related to the future of instructional technology.

Those variables on which instructional technology research is conducted are largely based on those factors held dear by individual professors or teachers. An analysis of these variables can be conducted in two ways---(1) ask what educational objectives are, or (2) inspect the examinations used. The former approach yields descriptions of everything from "improving thinking" to "reducing the myoptic view of our culture." The latter indicates that most examinations contain a majority of factual, memory-type test items (essay or objective) toward which learning must be oriented.

The type of examination system used in an educational program represents to a considerable degree the ultimate objectives in that learning system. The examination, therefore, serves to direct, influence, and determine what is measured and evaluated---and thus the concentration by design, intent or accident, that the student will place on certain elements to the exclusion of others---regardless of stated educational objectives. The examinations state the objectives very concretely, and with more meaningfulness to the student, than does the instructor, course outline, or textbook. If the system is composed largely of factual material, that which will be learned will be factual, at the expense of attitudes, appreciations, or problem solving concepts and understandings. To the extent that the evaluative instruments represent a known goal, that which occurs in the classroom (regardless of media or technology) is that which the student will learn (Spencer). The student himself controls what he will pay attention to or learn; partly on the basis of his own interest and motivation, partly (but only slightly) on the degree to which the instructional program can develop

interest, partly on the basis of his a priori attitudes toward the content of the instruction (he likes math but hates English), and on the basis of what he sees as the way he can maneuver through the system. Whether technology is used or not is irrelevant at this point--the overriding influence is the system. Any alteration in the system, such as moving from conventional instruction to some sort of instructional technology involves a relationship with an already existing institutionalization and reward system. If these later variables are related to instruction in such a manner that they reward convention more than newer procedures, the results of the newer procedures will tend to be depressed irrespective of content. This effect can operate on content, faculty participation, interest, and student response. An assumption underlying instructional technology experimentation exists which is relative to this area--that the system is adaptable to instructional technology, and that operations in this area will be welcome. Such an assumption has not been wholeheartedly validated.

As we have seen, examinations themselves may be classified as an instructional technology. CAI and Programmed Instruction have emphasized this variable in their operations, by attempting to supply immediate feedback of results to students as they work through an instructional sequence, in providing positive reinforcement (getting the test item correct), and in testing only that which has been presented. One major problem comparing instructional technology with conventional systems is the introduction of different examination variables into the picture. If one proceeds to "program" a course, and supply positive reinforcement and immediate feedback, one is varying a considerable number of factors other than the method of presentation (i.e. by computer, self-instructional programmed text, or T.V. film or whatever). One is also introducing a methodological difference (in the psychological sense); the manner in which the material is presented is different. One can find in the literature that changes in these methodological factors, within a conventional classroom system, can produce a

significant difference. The media may have little to do with learning--what's on the media is the thing.

There has been considerable emphasis on statistical treatment of instructional technology research, primarily because many of the assumptions underlying the statistics used are not met by the research design (Lumsdaine, p. 593). A second problem is a determination of significance for what? What type of significance is one really after?

If we imagine ourselves setting up an educational enterprise from scratch, with no previous institution ever having been in existence, we see that we must get the naive (student) connected to the all-knowing (professor). When we accomplish this purpose we can note by tests of significance that the student/professor system results in more learning (especially if the professor makes the examinations) than does the lack of such a combination. Thus we conclude that IF what the student/professor system has developed is of value to the society, we recommend its continuance, elaboration, and extension. Some systematic education is better than random ... no education. Now if we want to experiment with different types of educational systemization, we tend to prejudice our thinking in the direction of very slight changes (relatively), each within the student/professor system. Let's put the professor on T.V., or film, or radio; add a textbook, or program the instruction, or use a computer; self-pace or time-pace the student, put him into homogeneous groups of slow, average, or fast learners, etc. The amount of potential effect these variables can possibly have upon student learning is indeed small. The results validate this argument--but perhaps, our criteria have changed. The IF statements (if we can slip into computer language for a moment) have been changed. Now, we may want significance tests like IF we can do it cheaper, faster, to more people at once, with less effort, etc., the significance of the results is tested by a valid "no significant difference," i.e. we are successful if there are no differences. If however we are interested

in the students learning more, we have other things to consider. Now we have stated a specific direction; we must define "more" of what, and our effort (time, money, personnel) is directly related to the extent of the "more" that we want. If we decide to do in half the time (one semester) what was done before (in two semesters), again no significant difference in criterion test scores is a significant difference--our real criterion is time, which (2 X 1) is significantly different. But this design assumes we merely want to do the same kind of job. This is another primary problem with our assumptions. We assume that we know what we need to get accomplished, and that the educators know how to (pedologically) get it done.

As McCleary states (in another context, but applicable here):

1. "The immediate applicability of research findings becomes a primary criterion in the identification of the problem, research design, and support."

2. "The field of investigation is limited prematurely and may exclude essential criteria and/or experimental variables."

3. "The factor of self-correction--control and verification of activities and conclusions--is suspect."

4. "The haste to state conclusions crystalizes thinking and precludes examination of assumptions."

5. "Projects are too ambitious for the personnel and resources, thus they contribute to loose speculation about observations made in different situations." (McCleary, p. 7)

Educators are constantly arguing about what is a good teacher, and what is a "good" method of instruction. We are concerned about how to evaluate when a student has reached success or mastery. Our examination, evaluation, or grading system again gets in our way. Is an "A" given by a professor sufficient indication of mastery? Is one "A" equal to another "A" in another course or subject? In instructional technology research, one notices that criterion tests tend to be rather easy (as most classroom tests are). Most of the variance among students is determined by a spread of scores among the top 25% of the scale (between 75-100% items right). If we want to make comparisons, we obviously don't have much

room to operate. A football team (on offense) doesn't often get a long run for a touchdown because most of the time they are very near the goal line and do not have very far to go.

An ideal test for comparing differences between treatment groups would be one which distributes the students around an optimum mean (midway between chance score and the maximum possible score). Thus, on an objective test with five alternatives per item, and 100 items, an optimum mean would be the maximum score (100) minus chance score (20) = 80. A range of 80 between 20 and 100 would put the optimum mean at 60. Most classroom tests have means higher than this; and thus, the room to show the differences is depressed artifically by the type of examination used. There is not enough range for the students to show the gains they may have made.

In programmed instruction, and CAI, the tendency is to devise tests which have a mean score no lower than 95% in order to retain the assumption of positive reinforcement. This theory is seriously open to question, both on theoretical and statistical grounds (if one wants to compare groups). Tests are approximations of criterion behaviors sought by an instructional program. One instructs or presents a specific type of content, to which students respond. In an open situation (i.e. instructive) as traditional classroom instruction is, learnings are permitted to vary two-dimensionally--in amount of material that is learned (horizontally on a sequence ladder); and in kind (vertically across sets of ladders). Learning includes a variety of types, including factual knowledge or vocabulary (a neat, sequential, additive type of learning), to appreciation (very imprecise). In a closed, structured situation (P.I., C.A.I.) that which is taught is more prescribed and defined. There is less opportunity for horizontal movement. The two situations present very different instructional settings within which wide variation may occur.

There is certainly a value to be gained by the use of instructional technology. It presents an opportunity to investigate and possibly change an instructional unit or program and it exposes the system to view. This in itself may yield changes which benefit the student, but the one who is changing is most probably the individual faculty member; thus, not only does he change his procedures for the TV or C.A.I. or P.I. course, he also goes back and changes the conventional course. This system has not been investigated, because of a hypothesis that arranging a course for TV, for example, is sufficient reason to yield significant differences.

The potential for instructional technology is far greater than its use to date. Although wildly expensive, for educational minds and pocketbooks, instructional technology could be used to produce extensive "significant differences." In the first place let us acknowledge that the forgetting curve for school learning is monumental--most factual learning is forgotten within a very short period of time and most school learning is factual. Instructional technology can be used quite effectively for providing a tachistoscopic-memory device that could almost guarantee set levels of memorization--even overlearning. One could flood the student's perception with mechanical systems which he could scarcely escape from (1984 is only 16 years away). Attitudes, understanding, problem solving, evaluation, concept formations tasks are more difficult--but again readily possible given the appropriate resources to accomplish the task. If one looks at instructional technology installations across the country, one is reminded of where data processing was first housed--in basements, garages, and converted steam rooms. Present instructional technology setups indicate that instructional technology has evidently not been accepted by society. As Coombs states:

> "Education's technology, by and large, has made surprisingly little progress beyond the handicraft stage, whereas remarkable strides have been made in the technology and productivity of many other sectors of human activity, such as medicine, transportation, mining, communications, and manufacturing." (p. 7)

Jackson agrees:

> "...there is reason to suspect that many of the bolder forecasts concerning technological change in education will not be fulfilled." (p. 3)

Coombs goes on to suggest four reasons which might contribute to this problem:

1. "First is the sharp increase in popular aspirations for education, which has laid seige to existing schools and universities."

2. "Second is the acute scarcity of resources, which has constrained educational systems from responding more fully to new demands."

3. "Third is the inherent inertia of educational systems, which has caused them to respond too sluggishly in adopting their internal affairs to new external necessities, even when resources have not been the main obstacle to adaptation."

4. "Fourth is the inertia of societies themselves--the heavy weight of traditional attitudes...."

Instructional technology offers the potential of communicating the necessary facts and information into the public domain, rather than keeping educational materials locked inside classroom walls. Confining ETV to closed-circuit is an example of the degree to which financial restrictions and poor planning eliminate wide sources of education from public view. Perhaps the public is not interested; they certainly don't rush to watch NET, but the reason may be that the programs are

1. highly "arty"--drama, dance, music (and cooking)

2. highly academic, and by this I mean dismally boring and uninteresting for the potential audience.

That the Beverly Hillbillies is more popular than NET should not cast blame on the populace, but on the programming of NET!

Given the necessary and adequate funding, with creative and imaginative producers and writers, ETV can be as interesting and communicative as commerical TV. ETV has gone underground--perhaps because the public would be horribly shocked at

the state of teaching in American education. Why is there the attitude that we must not "entertain" our student audiences? Why are textbooks dull? There is simply no support, either financial or in the faculty reward system, for making an instructional program exciting and stimulating. There is no system to reward the innovator, and thus no particular reason to evaluate the results. Why evaluate if nothing different is going to occur? The academic institutions may be competent, but they may also be impotent.

There are certain other variables which demand consideration in regard to instructional technology, particularly in the area of specific evaluation procedures and research designs used. Objective testing has been considered the exemplification of empiricical research. Atkin (1963) for example indicated that innovation or creativity may develop beyond the capability of objective criteria to evaluate them. How then are the innovators with instructional technology to calculate their effects? The question resides, unfortunately, in the case of objective criteria, or subjective judgment, with WHO does the evaluation, and HOW the criteria are arrived at. Obviously, if one picks objective instruments which depress the capability for the measurement of creativity (whatever that is), it will not be discovered in the treatment or experimental design. Similarly, with subjective judgment as the method, one is dependent upon the capability of individuals to observe creativity. The criteria, however, remains constant in both cases--to develop a measuring instrument (be it observational, experiencial, objective, or performance) which will consistently (reliably) be sensitive to the behavior we are interested in developing. Secondly, it must also do the job the next time, or with different subjects, or different operators of the project.

The development of an instrument always is a less than perfect approximation of the behavior we might desire to measure--tests and people are sensitive and insensitive. Care must certainly be applied in order to prevent the instrument from concentrating, or over-structuring the intent of a project, or determining

in advance the results one may potentially render. If only a single, objective instrument is used as the criterion measure, we are placing all our marbles in one bag. A multi-media evaluation model (using objective tests, observational records, performance criterion, etc.) would be more likely to discover changes in students than a single-media approach to evaluation.

The type of evaluation system employed also depends on what it is we are attempting to do. If the object of the design is to show that one system of instruction is better than another, evaluation design, and media may differ considerably from a project which is determining what is the best way to accomplish a specific objective, or an exploratory instructional system which needs feedback into the system as the instruction develops.

In the same vein, much of the research with instructional technology has been concerned with a "defensive" research design; i.e. ETV is as good, or is better than conventional instruction (considering learning as the only criterion). A more rational approach to instructional technology would seem to be in the direction of "How can instructional technology be used to improve instruction?. When, at what age, in what areas, and with what types of learning? (See Carpenter, 1968)

It should be apparent from how instructional technology is used in education today that whole programs tend to be presented rather than single-variable stimuli; i.e. French I in 6th grade is presented as an entity, rather than, say, the concept of "carrying in two-digit addition." There has been little effort to discover where a particular media or technology would be of help. In general, this is primarily due to the fact that: (1) we don't know where certain procedures would help, and (2) we don't know where, in our conventional educational system, the most help is needed.

We do not normally include in our educational system a constant evaluation program to inform administration diagnostically where the good and poor points of

the program are located. For example, can we answer questions like: (1) Is high school chemistry taught better than history? (2) Is long division learned to the same extent as multiplication? (3) Is the English they learn in high school enough?

The university level is particularly vulnerable to questions of this sort since there are few criteria for judgment. The graduate schools can make judgments about undergraduate education--but few undergraduates go to graduate school. Since, however, this relationship is one of the few which can lead to evaluations of undergraduate education, the lower level system tends, perhaps, to orient the program in the direction of graduate school--even when only a minority of students ever arrive there.

The college can evaluate high schools--but again not all high school students go to college. The high school can evaluate junior high and they, in turn, the elementary. In general this results in a system of constantly casting blame downward. Each in turn sets its own criteria for selection and promotion on the ever present evaluation lying one step higher. The evaluation system is one which merely exists--no one thought it out, planned it, or adopted it--it grew. Is it the proper one? If one were to determine the practical objectives of the use of instructional technology, one would have to conclude that the criterion are already built into the system--get more people into college or graduate school! This has lead, one could presume, to an over-intellectualization of instructional technology. From observation (if I may use Atkin;s method of evaluation) one observes the primary resources going into college preparatory courses, material pre-determined by already existing course and curricular structure, required rather than elective courses, etc. Very little has been done with motivation, self-evaluation, vocational education, visual communication, interpersonal cooperation, selection and interest, reading cross-field educational designs, etc. We still operate in the structure developed a hundred years ago--the assumption must be that this is the way it ought to be done.

What instructional technology is used, how it is used, and on what is it applied "...is guided largely by force of personal conviction, and evaluation research is mainly of use in helping consumers decide where personal convictions to buy." (Bareiter, 1967, p. 192)

With the advent of instructional technology, some institutions have been altered in ways they had not foreseen. Film usage requires projectors, screens, and projectionists, as well as some delivery method for getting the supplies to where the students are. C.A.I. requires cables, a student station or more, a computer; TV requires a channel, a tower, a studio, etc., etc. To the extent that such mechanics are bought, or subscribed to, the students receive a different sort of instruction. The instructor may be a machine and not a person; it may be a self-controlled rather than teacher structured, written not spoken, or presented by a stranger on a screen rather than a teacher within the womb of the classroom. It is obvious that the individual teacher, therefore, looses direct control of what is presented, how it is presented, and when it is presented. The teacher becomes more of a participant than a director of the learning. This may result in many problems and disturbances---from some teachers turning the TV set off, not ordering films, not fitting the material into the classroom sequence of things, etc. The author observed an elementary class receiving TV instruction in the New Math and was interested in how the teacher was going to handle the content after the presentation was completed. The teacher responded with considerable aplomb--"That's enough of that New Math stuff," she said, "Now let's get back to OUR Mathematics--open your arithmetic Books to page 73; do exercises 1-10!"

There are several ways in which instructional technology can be used to improve education. Generally, the "replacement" system has been used, i.e. replace a course taught by conventional means with one taught on or by TV. This procedure merely accomplishes the same basic purpose with a little greater

hostility from the students. Some procedures, notably the audio-visual people, use a "supportive" system--assist an ongoing program to do the job a little better (which is difficult to find validity for). Other methods do exist, but seem difficult for administrative, legal, or institutional means. These methods include concepts of repetition, time, difference, applied, review, availability.

For example, if something is repeated often enough, it is more probable that it will be learned. Why not repeat the presentation again, over a different media perhaps, with greater clarity for those who didn't get it the first time, or in abstract for those who want merely a summary and review? It can be presented by another person or system--programmed instruction and ETV, radio or film. It can be presented during supper hours, over open circuit TV, offered at night, dial-access, or on the telephone. One can flood the market place and hit the students with many ways of presentation, many times, at their convenience.

A large issue in this whole argument is one of criteria--what do we want the student to do after the instructional program that he couldn't do beforehand. Generally we keep this a secret--we don't very often meaningfully communicate to him exactly what is expected of him. Many students fail. Should we blame the student or ourselves? If there are methods which will improve the number who reach the criteria, shouldn't we try it? Many times failure occurs because the student doesn't know what he doesn't know. Procedures can be developed which could answer this problem, through the use of instructional technology. One example will be presented to indicate an application of instructional technology in areas in which very little now exists in order to acknowledge the fact that instructional technology possibly would obtain better and more usable results if the area of attack was a novel one rather than an already existing one.

If we may assume that most learning will be evaluated by an examination, or series of examinations, and we further assume that the items on these examinations represent the content of the course, and similarly the capability of the

student to indicate the degree to which he has learned, then our objective becomes clear--i.e. to enable him to: (1) know the material well enough so that (2) he will be able to answer correctly the examination questions we will use to permit him to show us that he has learned the content (such questions may be objective, essay, performance, what have you).

The capability of the student to know what these test items are can certainly have an effect on his learning behavior. If the items are good ones and he knows what the items are (barring memorizing of the examination for the moment), his direction toward learning the correct answers to the items is the same as learning the content. That this is not always true should not alter the fact that if our assumptions are followed thus far, it should be true.

Given that he can obtain experience with these test items (takes a test, is graded, and learns his relative position in his class and the instructors opinion of his efforts), the feedback will inform him of what he knows or doesn't know, or how much he does or doesn't know. Such information, if presented diagnostically, can offer him specific knowledge of his particular strengths and weaknesses, dependent on the feedback of which items he has missed and which ones he got right. Similarly, information could be communicated dependent upon his item results as to where he may find a discussion about the area that that item represented.

Usually there are few rather than many examinations given during an instructional program--the procedure takes away from instructional time, someone had to score them, etc. Instructional technology provides means whereby such problems need not exist. Optical Scanning equipment can read answer sheets, or student input stations can be used to record answers to objective questions, scored by computer, and fedback (by printout or video display) to indicate his evaluation. The student may desire many questions, or few, testing often or seldom; if the facilities are readily available, he may guide himself, or be directed to pursue a line of evaluation most necessary for him or the teacher.

Many instructors would find such a situation intolerable because students may see test questions before instruction related to them has been presented, may be able to memorize answers to test questions, or may be able to find right answers by answering incorrectly and be corrected. This should be considered improper only to the degree that test questions are limited in their number, for if thousands of questions are available representing all parts of the instructional program, these intolerablenesses become motivational and acknowledging variables to relate the student to the material. Knowledge of the test, type of test items, and self-analysis (anonomyous) capabilities ought to improve learning. The only problem is in supplying items in sufficient quantity to form an ITEM BANK, which the student can sample for his purposes, and from which the teacher can sample for examination time.

The computer can store ALL multiplication, division, addition, and subtraction problems; can vary the numbers in word problems ad infinitum, can create questions on the basis of a test item model, can collect, file, and retrieve ALL questions ever used on tests. What is more, data on what students do with these questions can also be stored (how many, at what stage pass this item), and analysis information given to the student. The instructor can also obtain such data indicating where students are having problems. In the same manner a student can address the ITEM BANK with direction for what kind of test item he wishes to try:

Give me some items on photosyntheses used in Biology 101.

Give me two digit division problems.

Give me French vocabulary review items Chapter I, II, and III.

Etc., Etc.

Depending upon the sophistication of the instructional technology, the items could be visually displayed, could include graphics, charts, pictures, and configurations, could be solved by writing, light pen, or response to a multiple-choice question. The test can become a learning tool, rather than a punishment

or negative reinforcing device. The computer can be programmed to evaluate what a student asks for, how well he does, and on what he ought to be tested. Thus, a student who asks for a test item beyond the appropriate stage of the course can be automatically advised of this fact. If he gets the item incorrect, communication can occur to yield appropriate statements indicating he should, perhaps, try item 1489. If he gets an item correct he can be praised (within the limits of the programmer's skill in thinking of appropriate positive reinforcement phrases, like "good," "great," "wonderful," "very few students at your stage get this item right," or "Sorry--that's not the right answer; you didn't divide the last number by 2--would you like to try another problem of this type?", etc.).

Another aspect of instructional technology which is interesting to observe is the degree to which dullness has been supported so resolutely. No manifestation of Madison Avenue ever crops across the ETV screen or is included in programmed instruction or C.A.I. It is obvious from the life and death struggle on the commercial market and the glassy stares of many of the youth of America that commercial TV has something to offer; that commercial advertising--even for aspirin--does in fact work. Why haven't any of these procedures and systems been used in educational technology? Does commerce ask educational technologists for assistance in selling Volkswagens, or creating the Pepsi-generation, or creating an image of a political candidate? The image of education suffers by comparison. We are in the business of selling knowledge, perhaps to a much greater extent than we realize--learning is not becoming an "in" thing. The establishment is getting criticized from right and left, and perhaps rightly so. Can we use as a viable criterion that we ought to be able to teach Geometry to the same extent that Crest can sell toothpaste? Or are we satisfied and content that what we are doing is good and right, and thus should rest on its own laurels?

We have discussed both evaluations of instructional technology and how to evaluate instructional technology. The picture is indeed less than adequate. What then can be recommended?

One of our major problems with the use of instructional technology is the educational system itself in terms of credit allocations, management, and administrative systems, methods of showing proficiency, and product control. A system does not exist which gives support (i.e. promotion, salary increase) to those interested in improving instruction. Administrative details and incapability to change structures the manner and mode which yields us indices of learning. The learning appears, sometimes, less important than how the individual attained it. In one instance we require 16 hours of a foreign language for a B.A. degree, and on the other hand, prevent a foreign student who is in our university studying in English from presenting either English or his native language as fulfillments for this requirement. We pass a student into the next level course if he obtains a "D" or better grade in the prerequisite but we refuse to grant proficiency credit to someone who already had a "D" knowledge in that course. The system leads to a considerable amount of "playing the game" rather than learning. Competitive systems are not permitted. The system in fact works antagonistically to innovation which includes instructional technology.

Recommendations:

1. Develop funding which will lead educational institutions to build appropriate reward systems for the improvement of instruction--careers in teaching, particularly at the college level, are prerequisites to improved instruction.
2. Supply means whereby politically separate educational institutions become interested and rewarded for producing instructional programs useful in more than one institution--duplication of effort is wasteful and costly; efficiency needs to become an educational criteria.
3. Provide year long grants-in-aid for faculty to pursue research toward the improvement of instruction; in order to improve,

senior personal need to be released for concentration on the teaching process.

4. Subsidize large scale educational TV networks and communal C.A.I. systems; to provide means whereby institutions can observe presentations from other institutions, or can use them cooperatively if they so desire.

5. Develop model building educational resources, whose function it is to produce an obviously better instructional program and provide such proven courses nationwide--what is lacking in the technology to date is the proof that it can offer something better.

6. Support, with emphatic funding, experimental innovations in instructional programming, which will not fail because of lack of personnel, writers, producers or directors, material or equipment.

7. Study and produce an alternative administrative system of the teacher role; to enable college departments to continue their graduate programs without relying on graduate teaching assistantships; to reduce the competitive jealousy developed between public school teachers and instructional technology.

8. Investigate the possibility of establishing a federal free-university which can develop and make available courses of study in direct competition, or support of already existing courses, possibly presented over national ETV.

9. Supply funding for the expansion of libraries to include programmed texts, C.A.I., and video taped or audio taped courses of study (remedial, advanced).

10. Develop systems to automate information retrieval for students, for single-concept film loops on long-division, to TV recordings of plays, presentations of specific points, recordings of lectures, synopses of presentations, etc.

11. Support innovative practices which deal with the management of education; i.e. credits, hours, requirements, grades, probation, required attendance, or the classroom concept.

12. Assist in the development of procedures which will serve to bring the student in closer contact with that which is required of him--such as larger scale proficiency and diagnostic testing, test item bank procedures, and self-analyses systems.

These recommendations are, of course, very general; but they are intended to point in the direction which will en le the use of technology to catch up with the engineering state of the art. The hardware development has far outstripped the software (programs, content, assimilation, structure). Most important, however, is the system itself--it is not, and unless some radical changes are made in the management/administrative structure of educational institutions, will not be able to advantageously incorporate instructional technology into the system to an extent which the potential demands.

BIBLIOGRAPHY

Allen, William H. and Finer, James D., Co-Chairmen. "Instructional Materials: Educational Media and Technology." Review of Educational Research, Vol. 32, No. 2, April 1962.

Atkin, J. M. Some evaluation problems in a course content improvement project. J. Res. Science Teaching, 1963, 1, pp. 129-132.

Benjamin, Harold R. Saber-toothed Curriculum. New York: McGraw-Hill, 1939.

Bereiter, Carl. Acceleration of Intellectual Development in Early Childhood. U. S. Department of Health, Education, and Welfare, Final Report, Project No. 2129, Contract No. OE4-10-008. University of Illinois, Champaign, Illinois, 1967.

Carpenter, C. R. Planning for Colleges and Universities with Educational Technologies. Paper presented for the Illinois Conference on Higher Education, Monticello, Illinois, Nov. 2, 1967.

Coombs, Philip H. The World Educational Crises: A Systems Analysis. New York: Oxford University Press, 1968.

Dial-Access Information Retrieval Systems for Education. Newsletter, Special Issue #4, Vol. III, No. 3, May 1966. University of Wisconsin, Madison, Wisconsin.

Jackson, Philip W. The Teacher and the Machine. Pittsburgh: University of Pittsburgh Press, 1968.

Jerome, Judson. The System Really Isn't Working. Life, Nov. 1, 1968, pp. 67-70.

Lumsdaine, A. A. Instruments and Media of Instruction. In N. L. Gage, (Ed.) Handbook of Research on Teaching. American Educational Research Association, Chicago: Rand McNally, 1963, pp. 583-682.

McCleary, Lloyd E. Status of Research on Education in Latin America. Latin American Research Review, Vol. III, No. 1, Fall 1967.

McLuhan, Marshall and Fiore, Quentin. The Medium is the Message: An Inventory of Effects. Bantam Books, 1967.

Sapir, E. "Conceptual Categories of Primitive Languages." Science, 74, 1931, p. 378.

Spencer, R. E. The Improvement of College Level Student Achievement Through Changes in Classroom Examination Systems. U. S. Department of Health, Education, and Welfare, Contract #OEC-3-7-061147-0271.

Torkelson, Gerald M., Chairman. "Instructional Materials: Educational Media and Technology." Review of Educational Research, Vol. 38, No. 2, April 1968.

Whorf, B. L. Four Articles on Metalinguistics. Washington: Foreign Service Institute, 1950.