Little is known about the teaching effectiveness of instructional media, particularly film and television. Accumulated research evidence applies to a generalized "average student," and thus to no one. There has been little concern for individual differences in interaction with instructional-media variables. The problem lies with the design of experiments. In the animal lab, treatment averages are meaningful since deviations from the average are small and background variables are constant. In the case of a heterogeneous group, however, random division maintains heterogeneity, and treatment averages are therefore meaningless. Some improvement is brought to the situation if individuals are first separated into aptitude subgroups. Two major questions should be considered: 1) What aptitude variables are particularly relevant for filmed and/or televised instruction? and 2) What media attributes under what task requirements are particularly likely to interact with aptitudes? Past research has pitted one instructional medium against another without concern for differing individual responses to those media. An alternative approach would consider aptitude interactions with media variables, thus pointing up appropriate treatments for different kinds of students. (LS)
APTITUDES AND INSTRUCTIONAL MEDIA

Richard E. Snow
and
Gavriel Salomon

Technical Report No. 3
Project on Individual Differences in Learning Ability
as a Function of Instructional Variables

Lee J. Cronbach, Director
Richard E. Snow, Assoc. Director

School of Education
Stanford University
Stanford, California

May 1968
APTITUDES AND INSTRUCTIONAL MEDIA

Richard E. Snow

and

Gavriel Salomon

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.
The concern of this paper is with the nature of human aptitude and its relevance for instructional research and instructional practice, with particular reference to film and television. It begins with the bald statement of a truly null hypothesis: virtually nothing is known, as of today, about the teaching effectiveness of instructional media.

There are probably many complicated reasons which might be offered in support of such a view, and as many justifications for rejecting it. The simplest, most extreme, but hopefully an instructive supporting argument has been chosen for consideration here: almost all of the research evidence accumulated to date applies to some generalized "average student," and thus to no one.

The argument is not new and, as usually overstated, it is not a particularly constructive one. Most behavioral research problems seem intractable without some kind of averaging and it is possible, after all, to specify conditions under which the procedure is more or less justified. Learning theorists and laboratory experimenters have faced the issue intermittently for decades. Since their ultimate objective is general theory, they have deemed it appropriate in this pursuit to average out intra-species, or individual, differences. Many experimentalists believe that such differences, though ignored for the present, can eventually be treated as parameters in established theoretic equations. Currently, a

---

A portion of this paper was presented at the National Association of Educational Broadcasters-Instructional Division, Santa Barbara, California, April 4, 1966.
reluctant but relatively broad attack on the problem seems to be mounting in some of these circles (Gagné, 1967).

Among educational experimenters, also, there is growing awareness that instructional technology must be conceptualized as some combination of learning theory and individual differences, among other things. Individualized teaching, in the form of programmed or computer-based instruction, appears not to reduce the effects of individual differences, as originally hoped. Gagné (1964), in fact, has suggested that individual aptitudes must be ranked among the most important independent variables in the study of complex learning. In studies of instructional film and television, however, the problem seems still to be largely ignored. Faced with the need to understand an immensely complicated stimulus aggregate, usually used as a fixed global treatment or as an adjunct with unspecified instructional objectives, media specialists have settled for an undifferentiated and largely inadequate view of learner response. In the most recent annual convention of the Division of AV Instruction (Houston, Texas, March, 1968), for example, nineteen research papers were presented. Of these, only one included concern for individual differences in interaction with instructional-media variables.

A Problem of Experimental Design

The point of the argument is not, of course, directed at the operation of averaging, but rather against the use of undifferentiated averages as the final arbiters of instructional practice. The problem is rooted in the design of experiments.

Research on instructional media, as educational research in general, has relied somewhat blindly on classical principles of experimental
design borrowed from the animal laboratory and the agricultural experiment station. It is not that these design principles are wrong -- they are absolutely essential -- but their use hardly obviates the need for an investigator to think carefully about what he is doing. Their indiscriminate application in educational research frequently includes an implicit, bad metaphor. A man is not a rat, nor is he a plot of farmland. He was not bred from a single strain, nor is he in any important sense adjacent in space to other men. Breeding and plot-splitting make individuals homogeneous and random division maintains homogeneity. Treatment averages are meaningful here because deviations from the averages are small in such groups and because most background variables can be assumed constant. But random division of a heterogeneous group maintains heterogeneity.

College sophomores may look alike psychologically relative to differences observed in the general population. Relative to rats or plots however -- and here a bad metaphor can be used in the other direction -- a collection of sophomores looks more like a zoo or a farmers' market. Treatment averages here are meaningless. It is like comparing "tigers" or "applanges." Some improvement is brought to the situation if individuals are first separated into "tigers" and alligators" or "apples" and "oranges" so that average comparisons can be meaningfully interpreted within these subgroups. This clearly does not solve the whole problem, but if the variables used to stratify the group are well chosen, then at least the stage has been set for a new kind of instructional improvement, one based on the hypothesis that there is no "one best way" to teach anything.

Many investigations of learning under different modes of instruction simply assign students randomly to two or more treatments, compare average
performance on some criterion, and find no significant differences. No one familiar with the research literature on instructional film and television needs reminding that the great bulk of TV vs. live comparisons and many film vs. live comparisons, have ended this way. To underline the general problem and to introduce the solution proposed here, the comments of two prominent psychologists seem particularly apt. Hilgard (1965, p. 3) stated the problem as follows:

"It is surprising that, after all these years of doing it, we know so very little about effective teaching. The payoff of careful studies...is very slight indeed. It is surprising that studies of class size, discussion vs. lecture, and teaching aids such as motion pictures and TV point to so few differences in the effectiveness of teaching. These studies, therefore, give us little guidance. It is not that these studies are poorly done, and even studies which show little differences in effectiveness leave us with freedom of choice. My guess is that they fail, however, to understand the subtle differences made by kind of student, kind of teaching setting, and kind of long-range goals that are operative."

And Cronbach (1957, p. 631) proposed the solution as follows:

"Applied psychologists should deal with treatments and persons simultaneously. Treatments are characterized by many dimensions; so are persons. The two sets of dimensions together determine a payoff surface...We should design treatments, not to fit the average person, but to fit groups of students with particular aptitude patterns. Conversely, we should seek out the aptitudes which correspond to (interact with) modifiable aspects of the treatment."

An Alternative

Some illustrations will explicate the proposed alternative more fully and provide implications which can then be treated in more detail. The traditional academic prediction paradigm, in which some aptitude variable is correlated with achievement in a single instructional treatment, serves as a starting point. In Figure 1, the aptitude variable is positively related to learning in treatment T1. Individuals with higher aptitude scores learn more than do individuals with lower aptitude scores. Most
prior work which has considered aptitudes at all has been limited to

Insert Figure 1 about here

this outcome. But this finding is not particularly helpful because we are not simply interested in selecting students who will learn more, we are interested in increasing the learning of everybody. Suppose, however, that another instructional treatment (T2) can be found or designed in which the same aptitude is differently, even negatively, related to learning -- that is, where low aptitude students do especially well.

Figure 2 shows what has been called a disordinal interaction. The two regression lines intersect, in this case, near the average of each. An overall comparison would yield no significant difference. But it is

Insert Figure 2 about here

quite clear that if we use the aptitude variable to divide the group and assign the two subgroups to different instructional treatments, we can greatly improve the learning of each kind of person. Note that, to use such a finding, it is only necessary that the regression lines for the two treatments intersect (hence, are disordinal) somewhere within the obtained ranges of the aptitude and criterion variables. Note also that, to find the intersection, the regression slopes rather than the correlations alone must be studied.

Had we gone on trying to improve treatment T1, ignoring the aptitude interaction, some average increase might have been obtained, but it is
likely that some students would still have been better off in a different treatment. If, instead, we seek to improve both treatments with specific reference to the functioning of the aptitude variable, then it is likely that the regression slopes rather than the averages will be increased. Selected placement of students in the appropriately tailored treatment condition then maximizes the payoff for both groups.

A few examples from previous work (Snow, Tiffin & Seibert, 1963) should dispel doubts that such findings can actually be obtained. First, in Figure 3, the aptitude is a personality variable called "ascendancy," the dotted line represents live presentation of introductory college physics demonstrations while the solid line represents filmed demonstrations. The ordinate is learning as measured by immediate posttests. Note that there

Insert Figure 3 about here

... to no difference for the middle group but that for highly active, self-assured, assertive individuals the live condition is best and for students on the low end, characterized as passive, observers rather than participants, lacking self-confidence, and dependent on others, the film condition is best. Here, in Figure 4, the aptitude variable is "responsibility." Here, low responsibility students, described as unable to stick to tasks

Insert Figure 4 about here

... that do not interest them, and as flighty or irresponsible, seem to profit more from live than from filmed demonstrations. Again, there is a tendency toward reversal at the other end of the aptitude continuum.
Such findings are merely suggestive at this point; they represent only scattered unreplicated hunches. With these considerations, however, there seems to be sufficient reason to justify a search for other relationships of this kind. There are two major questions to be answered: 1) What aptitude variables are particularly relevant for filmed and/or televised instruction? 2) What media-attributes under what taks requirements are particularly likely to interact with aptitudes? There is no theoretical framework currently available to direct the work and there has been no general review of the literature on the subject. While a number of studies reviewed or abstracted by Hoben and Van Ormer (1950), Allen (1960), and MacLennan and Reid (1964) have included aptitude variables, most suffer from the inadequacies previously mentioned. Interactions which have been found previously were, for the most part, unanticipated outcomes of studies designed for other purposes. Some evidence is available, but an adequate accounting of this literature would require reanalysis of reported and unreported data, in most cases, to dredge out all of the relevant contrasts. While a general review of the literature is clearly beyond the scope of the present paper, a rough summary and some guidelines may tentatively be offered.

Aptitude as an Input

Upon entering instruction, an individual may be characterized in terms of an aptitude pattern which reflects the prior history and development of that individual. The term "aptitude" refers to any individual difference variable which functions selectively with respect to learning, that is, which appears to facilitate learning in some students and some instructional treatments while limiting or interfering with learning in
other students and other instructional treatments. The term does not mean "general mental ability." Of the previous studies recognizing the possibility of aptitude-treatment interactions, too many have simply collected IQ information and left it at that. Interactions between general ability and film-live or TV-live comparisons, may be expected when the live condition represents a differently paced or self-paced treatment like script-reading (see Kress and Cropper, 1966), but this aptitude variable could be insensitive to many other variations in treatment characteristics. Some 64 studies abstracted by MacLennan and Reid (1964) included some form of general ability or IQ measure. Of these, about one third reported positive interactions between aptitude measures and film or TV-mediated vs. conventional instruction. Of these, however, only a handful showed clearly disordinal interactions. In one study (Gordon, 1959), televised instruction was found superior to conventional instruction for high GPA children while the reverse was obtained for low GPA children. Even this one study involved statistical problems which limit the generalizability of its results. Similar observations are reported by Kanner and Wesley (1963), using Air Force Qualification Test scores as a measure of general ability, and by Jacobs and Bollenbacher (1959) who found high ability students learning more from TV and low ability students learning more from face-to-face presentation (Figure 5).

Other reported interactions all show differences between media at one ability level with no differences at another. Methodological problems,
sampling differences, and other shortcomings make construction of consistent patterns among these findings extremely difficult, but a rough overview suggests that one hypothesis advanced earlier may deserve further consideration. Both Westley and Barrow (1959) and Schramm (1962) have noted the possibility that, in an unrestricted population representing the full intelligence range, televised instruction might well be found superior to conventional presentations at both high and low IQ levels, with no differences between methods appearing in the large middle range of the general ability continuum. The rationale for this view rests on the presumed value of the more rigorously organized and attention-focussing character of most televised instruction, for individuals of lower intelligence, and its tendency to contain more information per time unit for reception by individuals of higher intelligence. Some support for the attentional hypothesis may be derivable from research on discrimination learning in retardates (Zeaman and House, 1967). Also, in a study by Kanner and Rosenstein (1960), it was found that low-ability groups of trainees tended to learn more from color than from black and white TV presentations while high-ability groups obtained more from the black and white mode. For the low-ability Ss, apparently, clear presentations were needed in which the displayed electronic particles could be easily discriminated from one another. High-ability Ss, on the other hand, seemed not to need the attention-directing addition of colors in that particular learning task. Perhaps also, low-ability Ss needed colors instead of, or in support of, their own internal labeling responses. The redundancy thus introduced could have functioned to reduce the amount of material learned by high ability individuals. In another study, Gropper and
Lumsdaine, 1961), high general ability students exposed to a programmed televised lesson obtained significantly higher criterion scores than high ability students exposed to conventionally televised material, using both immediate and delayed posttests. Low ability students, on the other hand, achieved higher scores when exposed to the conventionally televised lesson, but this difference was judged significant only for the delayed posttest. Several other studies (Kanner and Rosenstein, 1960; Kanner, Runyon and Desiderato, 1954; Kraft, 1961) have reported that students of high ability learn more than students of low ability regardless of method of instruction or mode of presentation. This conclusion would suggest that general ability measures provide a global assessment sometimes correlating positively with achievement regardless of the instructional treatments used.

Other individual difference measures, almost as readily available as general ability indices, are sex, age, and attitude toward the medium. These variables have not typically provided interactions in the past (see, for example, Almstead and Graf, 1960), and Amirian (1962). And even if obtained, such interactions would likely be considered indirect. That is, it is not these variables per se that would interact with learning but rather some other "hidden" variables correlated with sex, age, and attitude. There are still other potential aptitude variables correlated with the content of instruction rather than with the method or medium of instruction. These also are not good candidates since they correlate with an aspect of instruction that is common to all possible presentations. One must beware, however, of such a general distinction between form and content since, for example, prior knowledge of the subject-matter, a
content aptitude, may be a moderator of other interactions. One study (Snow, Tiffin and Seibert, 1955) found an interaction between numerical aptitude and film vs. live presentation in a low prior knowledge group. The effect was also apparent in a high prior knowledge group but its directions were reversed. A similar finding occurred using attitude toward instructional films as the aptitude variable. Although such higher-order relationships pose special problems for interpretation, they are to be expected especially with intellectual aptitudes.

Potential interactions are likely to reside in three classes of aptitude variables: 1) specific intellectual abilities like those defined in the work of Guilford (1967), 2) specific personality traits like those defined in the work of Cattell (1959), and perhaps most importantly 3) aptitudes in a poorly defined group of cognitive styles and preferences, learning sets, information processing and coding strategies, and other subtle experiential variables.

Taking the personality class first, consider the two variables mentioned earlier; "ascendancy" and "responsibility." If the interpretation is correct, then similar interactions might be expected in TV vs. live comparisons, particularly if the television condition included no talk-back link and no discussion leader or monitor. Within the TV condition, talk-back vs. no talk-back and discussion leader vs. no discussion leader may also interact with these aptitudes. Taken together, these aptitudes suggest a distinction between an interpersonal instructional atmosphere and an intrapersonal atmosphere. Any TV or film utilization variable which produces this contrast is likely to interact with such temperamental characteristics in students. Also noteworthy is the old
Hovland, Janis, and Kelley (1953) work on persuasibility and self-esteem. Instructional presentations which differ in authority reference or credibility should interact with high vs. low self-esteem. Another personality variable appears here also. Recent research directed by McGuire has shown relations between kinds of anxiety and persuasibility. In turn, differences in anxiety might relate to instructional treatments differing in pace, complexity or presentation, or reliance upon short-term memory (Sieber, 1967).

It is now appropriate to make a rather important distinction regarding the use of film or television simply as a distribution system as opposed to their use as unique visual media. The personality variables considered previously seem fairly persuasive in their effects. One might expect them to operate even when live lectures are compared with televised lectures or film recordings of lectures. Cognitive and experiential variables, on the other hand, should interact increasingly with treatment variables as these treatment variables depart from simple recording of live teacher behavior. When film or television is used to present contrived visual experience, we can expect many complex interactions with perceptual and cognitive aptitudes. This domain is virtually untouched but one hypothetical example has been suggested elsewhere (Pryluck and Snow, 1967). In a comparison between motion pictures and film strips in teaching map reading, Hovland, Lumsdaine, and Sheffield (1949) found a difference in learning on an item concerned with interpreting contour lines. The motion picture had used a visual transformation, from horizontal to vertical, to show the lines as projections of differences in elevation. In effect, a spatial visualization ability had been supplanted
by a cinematographic technique. Although this aptitude was not included in the study, it would be reasonable to expect that students with highly developed visualization ability would not need or want the filmic presentation while students without this capacity would find film essential. It should be added, however, that the visual transformation from horizontal to vertical, used in the film, could be used equally well by the filmstrip. The difference between the two would then be only with regard to continuous motion presented by the former and absent in the latter. It can be hypothesized that in such a comparison another aptitude would interact with the presentations, namely, an ability to visualize an integrated continuity from a series of still shots. The effects of other treatment variables, such as auditory vs. visual or single vs. multiple channel information transmission should also find their clearest expression in interaction with relevant aptitudes. In general, aptitudes involved in receiving, organizing, coding, manipulating, storing, and retrieving visual images should interact in many complex ways with variables reflecting different cinematic and editing practices.

Given even the limited available knowledge about individual differences in quantity of information intake, sensitivity to uncertainty, tendency to seek information, preferred level of stimulus complexity, etc. (Schroder, Driver, and Strufert, 1967), we are able to formulate specific hypotheses as to possible interactions. It is reasonable to hypothesize, for example, that individuals who are easily overwhelmed by uncertainty will learn more readily from straightforward "linearly" edited film, while others who prefer a higher level of uncertainty will be bored and thus inhibited by such a film. On the other hand, when a film is edited to
leave wide gaps in the course of action or to carry several story lines simultaneously, those who experience more response uncertainty and reach higher arousal levels should display improved performance, while those who are less aroused should be overwhelmed. Hypotheses of this sort require, however, a clear specification of the amount of uncertainty contained in each film version (Salomon and Snow, 1968). On the basis of results obtained by Suedfeld and Struelfert (1966), it may be suggested that such individual differences will interact also with the amount of new information given per unit of time in the film. Individuals who experience a high degree of response uncertainty tend to search for new information. Hence, highly loaded films should serve them best. Those who experience less response uncertainty, and thus need more correcting feedback (see MacLennan and Reid, 1964), will profit more from presentations that provide such feedback or repetition.

These last considerations begin to elaborate the third category of likely variables listed earlier and suggest a somewhat different approach to the whole problem. Aptitude can be viewed as the transfer of learning sets or information processing strategies from previous learning experiences. Individuals who differ in amount or kind of previous experience with film seem to differ also in the extent to which they profit from film learning treatments. Hoban and Van Ormer (1950) called this phenomenon 'film literacy' and it has been suggested that this unique aptitude represents a kind of comprehension of the 'grammar' of cinema (Pryluck and Snow, 1967). As television, both commercial and educational, develops a comparable visual language, we can expect comparable interactions with 'media literacy.' The earlier distinction between form and content apti-
tudes is essentially a distinction between aptitudinal and applicational transfer. Media literacy deserves considerable attention as a general aptitudinal transfer phenomenon, because it may represent both aptitude interaction with subsequent learning and aptitude development. Such a variable might be extremely important, also, in multi-cultural educational settings where students of different social classes or cultures differ both in degree and kind of "media literacy" (Salomon, 1968).

Aptitude as an Output

It is apparent from this view that an aptitude can affect learning and it can in turn be affected by learning. In the search for aptitudes as selective predictors, it is hoped that we will also keep an eye on aptitudes as outcomes of instruction. After all, the objectives of education are, in large part, aptitudinal in nature. We expect students to forget part of the information but to retain and expand upon the learning and thinking processes that educational experiences help develop. The instructional contributions made by film and television in this regard cannot be evaluated until aptitudinal criteria are included in the study of media effects. For instance, Salomon and McDonald (1968) have shown that teacher interns who are dissatisfied with their own teaching performance and seem to have generally low self-esteem react differently to self-viewing on videotape than interns with higher self-esteem and self-satisfaction. The attitudes of the former toward teacher education become less favorable after self-viewing, and they attend to more physical cues of their own appearance. The latter improve their self-evaluation and tend to notice more cues relevant to their teaching performance. On the basis of such findings it is reasonable to expect that the low self-esteem
teachers will change their self-esteem and their teaching behaviors by means of defense-reducing experiences, e.g., satisfying first-hand experience in classrooms. They cannot be expected to change when faced with their own image on videotape screens. The high self-esteem teachers, on the other hand, can be expected to benefit more from self-viewing than from repeated classroom teaching without such self-confrontation.

Conclusion

There was a time when it would have been ridiculous to suggest alternative filmed or televised presentations, for selected subgroups of students. But multichannel systems and videotape editing now make the suggestion feasible. Even for open-broadcast educational television, multiple channels may someday be available. Without these facilities, the aptitude-treatment interaction problem takes a different complexion. Nonetheless, more specific description of target audiences in aptitudinal terms and appropriate tailoring of program presentation could yield significantly improved results.

Hopefully, media specialists will increasingly consider the inclusion of aptitude variables in their thinking and in their work. Most previous research has pitted instructional media and methods against one another without concern for individual differences. While one cannot really argue that none of the treatment variables studied thus far produce general effects across all students, it is appropriate to ask how many such studies have masked real and important interactions by averaging in overall comparisons. Truly general effects will never be separated from important special effects until the possibility of interaction is directly investigated. The latter approach may well be the widest road currently available toward genuine instructional improvement.
References


Amirian, G. T. The retention by elementary school children of natural science material and of science attitude and interest changes following a program of science teaching by television. Unpublished Doctor's Dissertation, Boston University, Boston 1962.


Figure Captions

Figure 1. Regression of aptitude on learning in a single instructional treatment.
Figure 2. Intersecting regression slopes for alternative instructional treatments.
Figure 3. Interaction of ascendancy with film vs. live treatment (after Snow, Tiffin, and Seibert, 1965).
Figure 4. Interaction of responsibility with film vs. live treatment (after Snow, Tiffin, and Seibert, 1965).
Figure 5. Interaction of general mental ability with TV vs. live treatment (after Jacobs and Bollenbacher, 1959).
Fig. 1
Fig. 2.
Fig. 3

ASCENDANCY

LEARNING

Live

Film

Low Middle High
Fig. 4

RESPONSIBILITY

LEARNING

Low  Middle  High

Film  Live
Fig. 5.