DOCUMENT RESUME

ED 098 971	95	IR 001 355

AUTHOR	Salomon, Gavriel; Clark, Richard E.
TITLE	Re-Examining the Methodology of Research on Media and Technology in Education.
INSTITUTION	Stanford Univ., Calif. ERIC Clearinghouse on Information Resources.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Dec 74
CONTRACT	NIE-C-74-0027
NOTE	37p.; Draft copy
EDRS PRICE	MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS	Instructional Media; Instructional Technology; *Media Research; *Research Design; *Research Methodology; *Research Problems; *State of the Art Reviews

ABSTRACT

Research on media and technology in education appears in many forms and styles. The results of this research so far has been quite disappointing, in that researchers have failed to point out the difference between research on media and research with media. Research also has tended to become highly specific, thus losing representativeness. A conflict between better control over specific variables and representativeness could be solved by some new research methods. In field studies a pseudo-experimental design or a staged innovative design allow research to be conducted in naturalistic surroundings. In strictly experimental studies the rotation design and the ecological design can be used. The major theme of these techniques is to move media research out of its traditional either internal-or-external validity conflict toward better and more fruitful research. (JY)

002

RE-EXAMINING THE METHODOLOGY OF RESEARCH ON MEDIA AND TECHNOLOGY IN EDUCATION

Gavriel Salomon
The Hebrew University of Jerusalem

Richard E. Clark
Stanford University

RE-EXAMINING THE METHODOLOGY OF RESEARCH ON MEDIA AND TECHNOLOGY IN EDUCATION

Gavriel Salomon
The Hebrew University of Jerusalem

and

Richard E. Clark
Stanford University

Introduction

Research on media and technology in education covers a wide and diverse range of topics, settings, and domains. Its methodologies also vary from basic experimental work in the laboratory through field experiments in educational institutions, to large scale evaluations of programs and products in current use by schools. No single research methodology can accurately cover this range, nor can the field be made uniform to fit some arbitrary conception of THE best approach to research.

However, underlying this diversity there are common research objectives which guide the field and allow us to classify diverse studies, such as those resulting from a system's approach, together with studies of one vs. two channel inputs, under the heading of media research. Three major objectives of media research can be identified. The first is to obtain knowledge about the educational or instructional effectiveness of a chosen medium, or technology. Researchers who emphasize this objective attempt to answer the question of "how can various media best be used for instruction?" (Gagné, 1974). Typical of such attempts is the comparison between two media, or between alternative versions of the same medium (e.g. Allen and Weintraub, 1968). Other studies, in which the specific merit of a particular attribute or technique associated with a medium are investigated, can be included here.

The second objective of media research is to increase our understanding of how media and technology function, and what psychological effects they have on learners. This objective is, of course, related to the previously mentioned one, but differs from it in one respect: It is more concerned with how media function psychologically than with how effective they are. A study by Miller and Hess (1972) that investigated the provoking and engaging characteristics of CAI is typical of this goal.

The third objective of media research is to improve the practice of education through the provision and evaluation of better materials, media, procedures and technologies. The evaluation studies of Sesame Street (Ball & Bogatz, 1970; Bogatz & Ball, 1971) are examples of such studies.

It becomes evident that most researchers aim at more than one objective. For example, the above mentioned study by Miller and Hess was designed not only to gain better understanding of the motivational functions of CAI, but also to compare experimentally the methods of increasing the engagement of learners in the program. The latter is an attempt along the lines of the first type of objective mentioned above. Similarly, Van de Bogart (1972) made a case study of a series of television simulation presentations and evaluated it. He also attempted to reach some more general conclusions which could be applicable to other cases. Doing so, he aimed at both our first and third objectives.

It would appear that such diverse studies, often aiming at more than one objective, carried out over more than half a century, would yield valuable practical and theoretical results. This, however, seems not to be the case. Recent summaries of and comments on media research unanimously agree that media research has, in fact, yielded very little (Saettler, 1968; Snow & Salomon, 1968; Gordon, 1969, 1970; Allen, 1971; Jamison, Suppes & Wells, 1974; and others). Olson (1974) commented on the accumulated research on media by stating that --

"We know neither how to describe the psychological effects of ... technologies nor how to adapt them to the purposes of education. The impact of technologies both ancient and modern on children's learning is either negligible or unknown".

Indeed, it could be both. Yet, that the effects may be noticeable is illustrated by research done on television and aggression (Fuchs & Lyle, 1972), advertisement and children's consumer behavior (Liebert, Neale & Davidson, 1973) and the like. It seems more likely that educational media research has failed to detect such effects. More important, it has failed to explore ways of realizing yet unexplored potentialities of media, and converting them to educational purposes. We prefer therefore to re-examine the developmental trends of media research rather than blame the media themselves. In the following sections, major trends in past and present studies on media are examined, and some methodological suggestions are made.

Inter-Media Comparisons

Only a few years ago, research was mainly engaged in attaining the goal of finding the best medium or technology for instruction. This gave birth to a rather large number of inter-media comparisons, most salient of which were the medium vs. face-to-face studies.

The typical question -- is one medium (or technology, procedure, etc.) superior to another -- appeared to be a rather straightforward valid question to raise. However, as Knowlton (1964) and Mielke (1968, 1971) have convincingly shown, it was in fact an invalid question, leading to uninterpretable results, if results were obtained at all.

Methodologically, a comparison between two media calls for a well controlled experiment in which all variables, except a media variable, are held constant. The content, mode of presentation, structure, didactics, situation, and the like need to be equalized between the experimental conditions. In the typical ITV vs. face-to-face comparison the design would be as follows: A teacher presents the material in the face-to-face condition (no interaction with students permitted to avoid a new variable from entering) while another group of learners watches the same presentation on a TV monitor. Conditions, indeed, are equal and only the medium of presentation is allowed to vary.

However, as Mielke (1968) has shown, if all other variables have been controlled for, what was left to vary? What, then, was the independent variable whose effects were studied? All that remained to vary in such a study was the technology of transmission, since, indeed, other things were equal.

> The only reasonable conclusion [of such a study] would be that the mediation, and the mediation alone, caused the significant differences in [say] achievement (Mielke, 1968).

But such differences were rarely found. And if found, how could they be interpreted? The answer is offered by Gordon (1969):

> Most research in this area has been designed merely to measure the influence of technology (not mediums) upon academic grades, rather than determine the real difference between the mediums themselves. That these experiments have shown that the same kind of teaching operates more or less the same way with and without technological aids ... might have been anticipated before experimentation began (p. 118).

In short, when only the least significant aspects of instruction are allowed to vary, nothing of interest could, and did, result.

Research With and Research On Media

A similar criticism applied to a corollary of the media-comparison strategy, namely the study of classroom management procedures which in fact were unrelated to media. Yet, studies concerned with active vs. passive learning, immediate or delayed feedback to students (see e.g. Gay, 1972), and the like, appeared as if dealing with media. Knowlton (1964) has commented on such studies claiming that if, say, overt active participation of learners has been found to enhance their learning, why should it be any different when film is used? Indeed, no new knowledge was gained with respect to either human learning or the merits of film when research with media posed as if it was research on media.

There is a major difference of focus between the two. Whereas research with media only employs media as convenient devices for stimulus presentation or data collection, research on media involves inherent qualities or attributes of media as the major focus of investigation.

Research with media differs from research on media not only in focus but also in approach. Research concerned with the management of learning, for instance -- by presenting programmed materials on a TV screen (Carpenter 1968), investigates a "natural" given. Thus, Carpenter studied the tolerance of students to externally paced material, and Gay (1972) studied one's progress with immediate or delayed feedback when involved in difficult tasks. Research on media, on the other hand, need not be concerned with any available given. It can create the most contrived experiences, extract the most creative potentialities of media, and study how they affect learning, even if they are not typical of any existing instructional package.

Dealing with what media or technology might do, rather than with what they are doing implies that particular media attributes are singled out and built into well constructed experimental conditions to show what effects they have. This was done, for instance, by Salomon (1972) who studied the mental-supplanting functions of specific filmic techniques in interaction with individual differences. Such studies are gradually replacing inter-media comparisons as well as other types of research with media, as can be witnessed in Levie & Dickie's review of the field (1972). Thus, instead of gross media comparisons, research was focusing upon specific attributes whose instructional potentialities were being investigated.

Conceptualization and Interaction Studies

The realization that gross media comparisons and research with media approaches were of little practical or theoretical value to the field, was accompanied by appeals for more conceptualization to replace the strict empirical a-theoretical approach. Conceptualization was expected to focus researchers' attention to more valid questions on the *inherent* merits of media, and to imbed the research in theoretical contexts.

Indeed, a number of researchers have suggested ways to deal with media in terms of their inherent attributes and derive new hypotheses from the analyses of their merits (e.g., Pryluck & Snow, 1967; Salomon & Snow, 1968; Snow, 1970; Frederiksen, 1970; Clark, 1974).

It was generally suggested that attributes of media and technology be conceptualized in terms of their psychological or instructional effects and functions, rather than in terms of their physical appearance. Thus, it could be found that a number of physicalistically different attributes of media or technology have a high probability of eliciting similar responses in learners. By virtue of sharing a common effect on learners they then could be grouped together and be conceptually treated as equivalents. Studies, which would seem to deal with entirely different entities, such as line drawings and randomly ordered frames in a programed instruction text, could be found to bear upon the same cognitions, such as (say) response uncertainty.

Conceptualization of media attributes in terms of their psychological effects and functions points to a seldom made distinction between research concerned with psychological effects, and research concerned with instructional effectiveness. The former addresses itself to the information processing activities, mental or emotional states, brought about by different media attributes as they interact with learners' psychological characteristics. Such research is most typically found in the fields of psychology and communication. Berlyne's work (e.g. 1965) on the psychological effects of complex, ambiguous or asymetrical stimuli, and the psychological analyses of pictures by Gombrich (1974), are examples of research on effects.

Research on instructional effectiveness, on the other hand, is concerned with the contribution of some media attribute to learning in light of the demands of a specific task. It is, of course, the traditional dominion of educational research.

It was the fusion of research on effects with research on effectiveness which was so badly needed for conceptualization to emerge. Indeed, studies which point to psychological effects of media attributes are highly suggestive of hypotheses which point to the learning tasks for which the studied attributes could be most effective. The studies of Samuels on the role of pictures in the acquisition process of reading (Samuel, 1967; 1970;

Samuels, Biesbrock & Terry, 1974) illustrate this point.

More generally, knowing what effect a mode of stimulation has, and how it interacts with different types of learners, aids us in comprehending its more fundamental nature. The mode of stimulation can then be conceptualized in terms of its psychological function as was done for instance by Conway (1968) with regard to multiple-channel inputs, or by O'Neil (1970) with regard to negative feedback in CAI courses. A detailed analysis of the more fundamental requirements of learning tasks might consequently tell us the extent to which the documented effects become also instructionally effective.

That different media attributes may have different psychological effects on individual learners is quite self-evident. Less self-evident is the fact that instructional effectiveness is the result of the interaction of psychological effects with the requirements of the desired learning outcome. It becomes clearer, however, if we accept the idea that media or technology attributes facilitate learning to the extent that they activate, elicit or arouse in specific learners those mental states and processes which are relevant to the requirements of the task (Salomon, 1974a). The hypotheses concerning differential learning outcomes which can be derived from recent work on concrete-stimuli and their imagery arousing effects (e.g., Paivio, 1971) exemplify this point.

Generalizations

This shift from gross media comparisons to studies in which interactions of media with learners and tasks are sought implies that no overall generalizations of results, such as "CAI is superior to programmed texts", are formulated. It also leads research away from a strict trial-and-error empiricism and into a more conceptual realm.

It is interesting to note that limited, rather than total generalizations, were already sought after by the first generation of media researchers. Freeman (1924, quoted by Seattler, 1968) was apparently the first to conduct a systematic series of experiments with film. One of his conclusions follows:

> The relative effectiveness of verbal instruction as contrasted with the various forms of concrete or realistic material in visual media depends on the nature of the instruction to be given and the character of the learner's previous experience with objective materials (Seattler, 1968, p. 116).

It is unfortunate that the idea of interactions of media with learning tasks and learners was lost, giving way to studies which sought "the best medium."

Limited, rather than total, generalization of findings yield more meaningful and more logically acceptable results. Knowing how diverse learners and tasks are, and having noticed the wide range of messages and situations that media and technology can provide, it would be difficult to accept all-encompassing generalizations. (Snow (1973) maintains in this respect that –

> Interactions limit generalization of treatment effects. If an important personal characteristic of learners is sampled too narrowly . . . or if that characteristic is ignored . . . then generalization is rendered uncertain at best and patently wrong at worst (p. 8).

The design of interaction studies thus calls for a wide representation of learners and of learning tasks. In a study concerning motion in films, conducted by Allen and Weintraub (1968) this was the case. Three types of learning tasks were studied, and 582 learners differing in age, sex, ability and specific knowledge were tested. The researchers reached the conclusion that motion in films facilitates learning more than still pictures. This generalization appears to be warranted since this was the case regardless of learner or task differences. More often, interactions are found with either learners, tasks, or both, demonstrating that generalizations on the basis of restricted samples of learners or tasks are unwarranted.

In general, it might be said that as media attributes become better conceptualized, better rationales are developed, somewhat better understanding of specific media functions can be gained, and more justified limited generalizations are reached. Generalizations of this kind have recently been formulated by Allen (1974, in press) as a result of his extensive survey of the literature.

A Dilemma

Experimental work, as just described, has recently gained increasing prominence in the field of media and technology. However, the more it moved into the deeper layers of understanding media, the farther away it went from the world of education.[1] And in spite of its improved quality it nevertheless fell short of accomplishing the objective of improving educational practice.

There is a major reason for this failure. The research described above is by necessity highly analytic and detached, thus it is -- by its very

---

[1] The reader is invited to compare a 1963 summary of media research (Lumsdaine, 1963) with a similar survey published nine years later (Levie & Dickie, 1972).

nature -- unrepresentative of the real world of education.

One of the major purposes of media research is to deepen understanding of what functions media attributes can accomplish for different learners and different tasks. It must emphasize, first and foremost, internal validity. If the researcher wishes to ascribe a particular effect or function to a particular attribute, neatness of experimental comparison is necessarily called for. This calls for carefully arranged experiments in which only the desired variables are allowed to vary according to the researcher's rationale. However, when such is carefully done according to the canons of methodology, something of utmost importance is lost, namely: representativeness, or external validity.

A number of recent papers have been addressed to this issue (Bracht & Glass, 1968; Shulman, 1970; Snow, 1973). Briefly stated, there is a constant competition between the control of variables to allow accurate inferences, (emphasis on internal validity), and generalizability of findings to real life settings (emphasis on external validity). Indeed, what could be generalized to the real world of instruction from an experiment in which cooperative students were shown experimentally manufactured stimuli and asked to learn experimentally designed materials under ideal conditions, introduced as a relaxing break in a tedious schoolyear?

Acknowledging this, it is often argued that the research needs to be carried into the real education world and conducted with real materials and stimuli. Foxall (1972) compared the relative instructional effectiveness of radiovision (tape recorded sound track accompanied by a filmstrip) with television in a course of new mathematics. Each medium of instruction was allowed to manifest all its qualities. Thus, there were no time limitations on the radiovision, while repetitions, discussions and the like were allowed. The television presentation, on the other hand, contained much animation but did not permit interruptions. No significant differences between the groups were found, leading to the conclusion that radiovision, which is far less costly than television, facilitates learning as much as television does.

Internal validity is, of course, sacrificed in such a study as in most evaluation studies. A host of known and unknown variables are involved. Even if results would favor one medium, no clear explanation could be provided, nor would it be possible to suggest which of all the participating variables is responsible for the outcomes. This, however, could still be sufficient for a summative evaluation study which, as Guttenberg (1971) has pointed out, involves a "judgment of worthwhileness of some activity."

Foxall wishes, however, to carry out an investigation into the relative merits of two teaching media, not just two highly specific instructional packages. He aimed at external validity, or generalizability, but because of poor internal validity the study failed.

It becomes evident that sacrificing internal validity, allowing media to exploit their best qualities, and moving out of the highly controlled laboratory-like setting, are insufficient conditions to secure represenativeness.

Thus, one faces a dilemma. On the one hand,

> If complex behaviour is assumed to be both probalistic and multidimensional, "stripping" the environment down to a minimum in order to control, to determine the role of a few variables, may be a potentially self-defeating process (Perebum, 1971, p. 445).

A similar argument was raised by Ebel (1967), leading him to conclude that the search for generalizations based on analytic research is futile.

Yet, on the other hand,

> The empiricist in education stops with comparison of gross effects. He asks . . . whether an educational film produces more learning than a verbal presentation. Such studies play a significant part in the engineering phase of curriculum preparation. But research that is no more than an empirical check on the effectiveness of some educational package does not get to the heart of the matter. One cannot generalize, for example, about the advantages of educational films. The strict empirical approach would force us to assess each film in turn, and, indeed, to test its effects when applied in each of several different ways. We will never have enough investigators to carry out exhaustive studies of this kind, *so that the only practical approach is to search for explanatory principles* (Cronbach, 1966, p. 543, italics added).

Differently stated, better control over variables provides better internal validity, allows better conceptualization and understanding, but little representativeness, and hence -- has remote relevance to educational practice. On the other hand, studies in the real world of education, dealing with complex variables, are most often highly specific and do not warrant generalization. Given also the complexity of the phenomena they deal with -- their internal validity is very poor.

This dilemma has become a focus of interest to researchers, and a number of extremely potent methodological solutions have been suggested (e.g., Campbell, 1969, Bracht & Glass, 1968; Shulman, 1970; Snow, 1973; Buss, 1974). In the following pages we will discuss ways to apply some of these solutions to media research. Essentially, two major kinds of research strategies are suggested: Strategies to make the field study, concerned with real events in the natural setting more generalizable and internally valid, and strategies to make experiments more externally valid without losing internal validity.

Improving Media Research in Natural Settings

Natural settings provide the researcher with a host of situations, outstanding events, and highly innovative projects which deserve to be carefully studied. Such events are particularly prominent in the world of media and technology. Their investigation is important inasmuch as their quality, imaginativeness and complexity far exceed the events typically studied by researchers. Most research concerned with innovative, yet complex, real-life instructional media, if conducted at all, is usually limited to simple gross evaluations that lack internal validity. The problem is therefore how to conduct research on real-world events, including large-scale program evaluations as well as research into the effects of outstanding innovations, while attaining satisfactory internal validity.

The Statistical Pseudo-Experimental Design

Salomon (1971) suggested that media research might sometimes be more successful if it started out with events in the real world and worked backwards into the laboratory by gradually analyzing them into ever smaller components. Earlier, Shulman (1970) proposed a similar approach labeled the

"Epidemiological Strategy" to accomplish the same end for research in teaching and learning. Essentially, outcomes are compared after learners have been differentially exposed to an external, natural, factor such as a TV program. Gathering data on numerous individual difference variables including background, abilities, achievement, and the like, it should be possible to distinguish between those who are more and those who are less affected by the program.

This strategy needs to be supplemented by a careful analysis of the various components comprising the program, or the newly introduced technology. Identifying such significant components, the researcher should be able to generate hypotheses as to their possible effects and effectiveness. He should thus be able to investigate not only who was more and who was less affected, but also what caused the effect. In this way, a real life event could be studied as if carefully controlled experimental conditions were present, while in fact they were not.

The Pseudo Experimental Design (somewhat different from Shulman's Epidemiological Strategy), is based on the measurement of three kinds of independent variables: relevant individual differences of the students involved, instructionally significant components of the program, and amount of student exposure to -- or involvement in the program. Students differ as to the amount of their exposure to, or involvement in an instructional program -- thus exposure is a continuous major independent variable. The purpose of this approach is, then, to examine the extent to which amount of exposure or involvement differentially affects students. Since, however, the program is analyzed into its significant components, one can also attempt to answer the question of what elements in the program affect individual learners.

It becomes immediately clear that the examination of the program's effects, when carried out under natural conditions, is methodologically deficient. The amount of exposure to the program by each student may be the result of self-selection. Better able, more curious, more outgoing students may choose to expose themselves more to the program. The necessary condition of "other things being equal" is not met unless statistical procedures are used to partially "equalize" the students. Toward this end, background data needs to be collected and multiple-regression procedures used (c.f. Cohen, 1968). It then becomes possible to partial out initial exposure-related differences. We meet, to an extent, the condition of "other things being equal" through the Pseudo-Experiment rather than through design procedures, as will be illustrated below.

An Example

The introduction of Sesame Street to Israeli children created a unique opportunity to study the effects of a highly complex and sophisticated program on TV-naive children (Salomon, 1974b). Since, however, it was broadcasted simultaneously all over the country a traditional experimental design became impossible. No adequate control group of children who were not expected to watch the program could be formed. On the other hand, simple comparisons between heavy and light viewers of the program would be meaningless since amount of viewing could be the result of self-selection, thus threatening internal validity.

Even if this threat were removed, there was still the threat to external validity. Since the effects of only one program (40 one-hour shows) were to be studied, generalizability would be extremely limited, as in most evaluation studies. The effects of one program may not represent the possible effects of other programs.

Specific statistical methods, however, allow us to overcome such difficulties. If one wishes to talk about changes in achievement which are presumed to be the result of program viewing, then each child's amount of exposure to the program can be measured and the degree to which exposure is related to later achievement can be computed. Actually, we have a situation where the independent variable (exposure) has values distributed over a wide range: from total non-exposure, through many levels of partial exposure, to the total exposure of each and every show. In this respect we have an advantage over the traditional experimental design in which subjects are divided into groups of "viewers" and "non-viewers." The traditional method usually avoids looking into differences <u>within</u> each one of the groups, whereas here they are taken into account.

The statistical method of Multiple Regression allows us to partial out the contributions of background and initial achievement variables, thus measuring the "net" contribution of exposure to the program to the post-viewing achievements (e.g., Cohen, 1968). In other words, to some extent we are able to specify the "net" amount of post-viewing achievements which can be attributed to exposure, other things being equal.

This method of analysis allows us to compare groups inasmuch as the same background and pre-viewing measures are entered into the analyses in the same order, and exposure entered as the last predictor. It is thus possible to see in which group its "net" contribution to post-viewing achievement is larger.

Table 1 provides an abbreviated example of data on two tests resulting from such an analysis.

- - - - - - - - - -

Table 1 about here

- - - - - - - - - -

As can be seen all background and initial achievement measures accounted for 36.8 to 50.8 percent of the post-viewing variance depending on the group and the test. Exposure accounted for 4.3 to 16.3 additional percent. It can also be seen that while exposure made a significant difference for lower class children in the case of the Letter Matching test, it did not make much of a difference for middle class ones. The converse is true in the case of the Parts of the Whole test.

It becomes evident that the multiple regression analysis is not used only as a statistical procedure but as a basic element of the design. Without it, subjects can not be "equalized" and the absence of a control group, typical of the study of natural events, can not be overcome.

The question of generalizability was treated, as mentioned already, through conceptualization of specific program components followed by the generation of specific hypotheses. Thus, for example, it was hypothesized that particular presentation formats used in the program would affect specific skills in particular children. Although such components could not be experimentally manipulated, it was still possible to test such hypotheses using multiple regression procedures.

The Staged Innovation Design

Sometimes it may happen that a new media-based program, sufficiently innovative to deserve a thorough study, is introduced into schools. Let us assume that all students are to participate in the program, making it again difficult to create adequate control groups. However, it might become possible to introduce the program in stages, thus allowing for a Staged Innovation Design (Campbell, 1969). According to this design not all schools are introduced to the program simultaneously. Some schools, if possible randomly assigned,[2] are introduced to the program earlier than others. The early beginners turn out to serve as the "experimental" group while the late beginners serve temporarily as the no-treatment "controls." Achievements, or any other dependent variables, serve to compare the "experimental" (early beginners) with the "control" (late beginners).

This design can be further developed as follows: Once the "controls" take part in the program, their achievemnts can be compared with those of the "experimentals" as measured on an earlier date. Thus a replication is built into the design. Two groups have taken part in the program, one after the other, and their results can be compared on two occasions: before the "control" schools started out with the program, and again -- after they have finished it.

One can also try to change the program before the "control" schools begin to participate in it. Comparing their post-participation results with those of the "experimental" schools, measured on an earlier date, is similar to an experimental comparison in which the newly introduced changes in the program serve as the "treatment." The outline of the Staged Innovation Design are graphically presented in Figure 1.

- - - - - - - - - - -

Figure 1 about here

- - - - - - - - - - -

An Example

Elements of the Staged Innovation Design can be found in the Age Cohert study, part of the first year's evaluation of Sasame Street (Ball & Bogatz,

---

[2] If not, schools can still be equalized using elements from the Statistical Pseudo Experimental Design.

1970). In that study, 114 children, 53-58 months old, were pretested before the program was shown, and their achievements compared with those of another group of 101 children of the same age, after the program was shown. When the posttest group was divided into viewing quartiles, it was found that those who viewed the program more achieved more than the pretest group. Thus, the conclusion was reached that viewing the program led to greater gains in scores.

It will be noted that the Age Cohort resembles the Staged Innovation Design inasmuch as it compares groups at different points in time. Those who are about to receive the "treatment" serve as the controls, and their pre-"treatment" scores are compared with the scores of another group after it has received the "treatment."

Improving Experimental Media Research

The major problem of experimental research on media and technology is, as already discussed, its remoteness from the real world of education. While it does not necessarily suffer from the absence of internal validity, it lacks representativeness. Snow (1973) in a recent paper, discusses this problem and ways to solve it in great detail.

The biggest threat to external validity, claims Snow (1973), comes -

> . . . when the experiment does not fit the nature of the behavior being studied and, furthermore, does not include the means of discovering this fact (p. 2).

To this, it may be added, external validity is badly reduced when the variables under investigation are isolated, taken as discrete elements, and examined under relatively artificial conditions. This, then, leads to the consideration of ecological validity, i.e. -

> Ecological validity concerns the extent to which the habitats or situations compared in an experiment are representative of the population of situations to which the investigator wishes to generalize (Snow, 1973, p. 10).

The consideration of ecological validity needs to be complemented by "referent generality," that is, the use of a wide range of possible-experimental outcomes (not just achievement on one multiple-choice test). These may include both expected as well as unexpected outcomes.

The Ecological Design

Snow enumerates several ways of securing ecological validity and referent generality. Among those we find the observation and recording of intra-experimental processes (what do the students do, in fact, when the treatment is applied), extra-experimental observations (how disruptive is the experiment when initiated in midschool year), description and analysis of the students taking part in the study, their preparation for the experiment, the duration of the experimental treatment, its detailed description, and the like.

Essentially, all this adds up to multivariate experiments, conducted under real life conditions, with real life treatments of more than a brief duration, with a wide range of dependent variables. Snow suggests a design by means of which one can study the interactive effects of complex treatment variables comprising an instructional sequence, complex learners' aptitudes and some learning criteria.[3] Figure 2 represents one version of Snow's design.

---

[3] The interested reader is urged to read Snow's paper (1973) for details and explanations.

Figure 2 about here

The Ecological Design when used under the conditions mentioned above, could help to reduce threats to external validity.

The Rotation Design

A somewhat simpler version of the Ecological Design may also be appropriate for media research, where the experimenter wishes to extract some outstanding qualities of media and study their effects and effectiveness in interaction with learners and learning tasks. The researcher may generate specific hypotheses concerning the effects of these qualities on particular learners, and may try to find out for what kinds of tasks these are most appropriate.

Imagine a program which can be divided into a number of different tasks based on some learning hierarchy, taxonomy or a task analysis. Assume also that a number of general and specific aptitude measures of learners are taken. The researcher prepares then a number of alternative ways of teaching each of the program's tasks such that each task (chapter, topic, or any other discrete component) is taught to another group of learners using a different medium or technology. Each medium prepared to teach the material in one of the task units is so structured as to capitalize on the special attributes of that medium.

Comparable groups of learners, preferably in their natural learning habitats, are taught the same program. However, each group is exposed to different task/medium compositions.

The design (labeled as the Rotation Design for our present purposes) is shown in Figure 3.

Figure 3 about here

For illustrative purposes we have put into the design four student groups, four media and a four-task learning program. Other combinations are of course possible. As it can be seen, each group learns the whole program with a different combination of media. This enables us to compare during-the-study behaviors as well as posttest performance within each row separately, that is, within one task and across media. Given that we employ measures which are common to all tasks (e.g. curiosity), it becomes possible for us to compare results within one medium and across tasks. Finally, aptitude-treatment-interactions can be studied within each row, thus showing whether learning of a task by means of one medium benefits certain learners unlike learning the task by means of another medium which may benefit other learners. The same analysis can be carried out within one medium and across tasks and groups.

In spite of its appearance, this is not a factorial design. No row or column main effects are sought after. Indeed, what could a comparison between the columns yield? And how instructive would a row comparison be? The inter-column comparison would be like a formative evaluation study, but then -- it was not our intention to test the overall effectiveness of an instructional package with this design.

Each row in the design represents one learning task, topic or period of any desired duration and complexity. Within the row, a one-way analysis of variance, to test media effects, becomes possible. This could be done with each row separately. However, since learning of the program is hierarchical or at least cumulative, the rows should not be analyzed independently of each other. Thus, from the second row on, analyses of covariance (or multiple regressions) are called for so that previous learning -- the covariates -- are taken into consideration. This design could have been used by Allen & Weintraub (1968) rather than having three independent experiments, each dealing with a different task. Since, however, they were concerned with the interaction between filmed motion and tasks, outside the realms of a real instructional program (thus not considering the gradual accumulation of knowledge), their study lacks representativeness. The Rotation Design could overcome this shortcoming.

An Example

Samuels, Biesbrock & Terry (1974) wished to determine whether pictorial illustrations would influence beginning readers' attitudes toward stories they read. Some of the psychological effects of illustrations when used in primary readers were investigated earlier, indicating strong interference effects (Samuels, 1970). Thus, the present study was concerned mainly with affective effects and their instructional utility. Using a Graeco Latin Square Repeated Measures Design, the researchers assigned students to one of three groups. Each group read one story each day for three days. Each story was accompanied by a different type of illustrations. Thus, no two groups read the same story with the same illustrations, nor did two groups read the same story on the same day.

The design used by Samuels et al. differs only slightly from the Rotation Design. The major difference lies in the fact that no particular order of story-presentation was needed, whereas the Rotation Design is better suited to deal with curricula in which chapter or topic order is given.

The Graeco-Latin Square design as used by Samuels et al. does not consider interactions with individual differences. In the Rotation Design this is a critical component. Yet, in spite of these differences the two designs appear to be rather similar inasmuch as both deal with all possible media/task combinations and both can be carried out in natural settings.

Summary

Research on Media and Technology in Education appears in many forms and styles. However, its yield in terms of understanding media, guiding their utilization, or improving education was, and still is, quite disappointing. A number of critics have shown the inadequacy of gross media comparisons, and have pointed to the difference between research on media and research with media. Media attributes, rather than media per se, were dealt with and a strong need for conceptualizing these attributes in terms of their effects and functions was suggested. This has led to more specific and sophisticated research questions, concerned with media x task x learner interactions, to replace the questions about the universal merits of media.

However, the more specific the questions became, the more the research moved away from the real world of education, in spite of the promise for better conceptualization and theory. The research became highly specific, neat, even sterile to a degree, thus increasingly losing representativeness.

A clear conflict between better control over specific variables and representativeness (that is, relevance to the real world of education) could be felt. A number of solutions were suggested. These pertained particularly to two instances: first, to studies conducted in natural settings on innovative media programs which usually suffer from the absence of internal validity, and to well-controlled experiments suffering, as indicated, from poor external validity or representativeness.

As for the field studies two designs were discussed. The first, the Statistical Pseudo-Experimental Design, is based on the use of multiple regressions. It enables us to conduct large-scale research on the effects of a program, without control groups, while still making it possible to estimate effects, interactions and causes for those. In spite of the absence of experimental controls, this design enables us, given certain provisions, to maintain satisfactory internal validity. Second was the Staged Innovation Design, which allows inter-group comparisons when a program is introduced in stages.

Two designs were also suggested for experimental research; the Rotation Design was offered to supplement other newly developed multivariate designs discussed in the current literature. That design enables us to study the interactions of media, tasks and learners in real life settings, taking into consideration the complexity of the educational setting as well as the accumulation of knowledge. The Ecological Design was offered as a method of including a wide range of situational and content variables, and a similarly wide range of media attributes, in an experimental structure.

The designs discussed herein somewhat blur the distinction between the traditional types of media research mentioned in the opening section. For instance, evaluation studies can be so conceptualized and designed as to answer also questions pertaining to the effects, and not only the effectiveness, of a media program. Similarly, experimental studies, originally designed to deal with psychological effects of specific media attributes, can be made to deal also with questions of instructional effectiveness in the real world of education.

The major theme thus is, that by moving media research out of its traditional either-internal-or-external-validity conflict, better and more fruitful research could emerge.

References

Allen, W. H. Instructional media research: past, present and future. AV Communication Review, 1971, 19, 1, pp. 9-18.

Allen, W. H. Intellectual abilities and instructional media design. Prepublication draft. ERIC Clearinghouse, Stanford University, 1974.

Allen, W. H., & Weintraub, R. The motion variables in film presentations: Final report. University of Southern California, 1968.

Ball, S., & Bogatz, G. A. The First Year of Sesame Street: An Evaluation, Princeton: Educational Testing Service, 1970.

Berlyne, D. E. Structure and direction in thinking. New York: Wiley, 1965.

Bogatz, G. A., & Ball, S. The Second Year of Sesame Street: A Continuing Evaluation. Princeton: Educational Testing Service, 1971.

Bracht, G. H., & Glass, G. V. The external validity of comparative experiments in education and the social sciences. American Educational Research Journal, 1968, 5, pp. 437-474.

Buss, R. A. A general developmental model for interindividual differences, intraindividual differences, and intraindividual changes. Developmental Psychology, 1974, 10, 1, pp. 70-78.

Campbell, D. T. Reforms as experiments. American Psychologist, 1969, 24, 4, pp. 409-429.

Clark, R. E. Taxonomizing media attributes for research purposes, prepublication draft. ERIC Clearinghouse, Stanford University, 1974.

Cohen, J. Multiple regression as a general data-analytic system. Psychological Bulletin, 1968, 10, 6, pp. 426-443.

Conway, J. K. Information presentation, information processing, and the sign vehicle. AV Communication Review, 1968, 16, 4, pp. 403-414.

Cronbach, L. J. The role of the university in improving education. Phi Delta Kappan, June 1966, pp. 539-545.

Ebel, R. L. Some limitations of basic research in education. Phi Delta Kappan, 1967, pp. 81-84.

Foxall, A. Television and radiovision in the teaching of modern mathematics: A comparative study. British Journal of Educational Technology, 1972, 3, 3, pp. 236-244.

Frederiksen, N. Toward a taxonomy of situations. American Psychologist, 1972, 27, 2, pp. 114-123.

Freeman, F. N. (ed.), Visual Education. Chicago: University of Chicago Press, 1924.

Fuchs, D. A., & Lyle, T. Mass media portrayal--sex and violence, in F. G. Kline and P. J. Tichenor (eds.) Perspectives in Mass Communication Research. Beverly Hills: Sage Publications, 1972.

Gagne, R. M. Educational technology and the learning process. Educational Researcher, 1974, 3, 1, pp. 3-9.

Gay, L. R. Delay of feedback and its effects on the retention of three types of computer-assisted mathematics instruction. Florida State University, 1972.

Gombrich, E. H. The visual image. In Olson, D. R. (ed.) Media and Symbols: The Forms of Expression, Communication, and Education. The Yearbook of the NSSE. Chicago: The University of Chicago Press, 1974.

Gordon, G. N. The Languages of Communication. New York: Hastings House, 1969.

Gordon, G. N. Classroom Television: New Frontiers in ITV. New York: Hastings House, 1970.

Guttentag, M. Models and methods in evaluation research. Journal for the Theory of Social Behaviour, 1971, 1, 1, pp. 75-93.

Jamison, D., Suppes, P., & Wells, S. The effectiveness of alternative instructional media: A survey. Review of Educational Research, 1974, 44, pp. 1-69.

Knowlton, J. Q. A socio-linguistic theory of pictorial communication. AV center, Indiana University, 1964.

Levie, W. H., & Dickie, K. E. The analysis and application of media. In R. M. W. Travers (ed.) The Second Handbook of Research on Teaching, Chicago: Rand McNally, 1972.

Liebert, R. M., Neale, J. M., & Davidson, E. S. The Early Window: Effect of TV on Children and Youth, New York: Pergamon, 1973.

Lumsdaine, A. A. Instruments and media of instruction. In N. L. Gage (ed.) Handbook of Research on Teaching. Chicago: Rand McNally & Company, 1963.

Mielke, K. W. Asking the right ETV Questions. Educational Broadcasting Review, 1968, 2, 6.

Mielke, K. W. Evaluation of learning from televised instruction. In R. C. Burke (ed.) Instructional Television. Bloomington, Indiana University Press, 1971.

Miller, R., & Hess, R. D. The effects upon students' motivation of fit between student ability and level of difficulty of CAI programs. Stanford University, 1972.

Olson, D. R. Introduction. In D. R. Olson (ed.) Media and Symbols: The Forms of Expression, Communication and Education. The NSSE Yearbook, Chicago University Press, 1974.

O'Neil, H. F. Effects of stress on state anxiety and performance in computer assisted learning. Florida State University, 1970.

Paivio, A. Imagery and language. In S. J. Segal (ed.) Imagery. New York: Academic Press, 1971.

Perchum, A. C. Some fundamental problems in experimental psychology: An Overview. Psychological Reports, 1971, 28, pp. 439-455.

Salomon, G. Summary and Comments. In M. L. Fleming (ed.), Aspects of Perception - Their Relation to Instructional Communications. Viewpoints, 1971, 47, 4, pp. 207-214.

Salomon, G. Can we affect cognitive skills through visual media? An hypothesis and initial finding. AV Communication Review, 1972, 20, 4, pp. 401-422.

Salomon, G. What is learned and how it is taught: The interaction between media, message, task and learner. In Olson (ed.) op. cit. 1974 (a).

Salomon, G. Sesame Street in Israel: Its instructional and psychological effects on children. New York: Children's Television Workshop, 1974b (in press).

Salomon, G., & Snow, R. E. The specification of film attributes for psychological and educational research purposes. AV Communication Review, 1968, 16, 3, pp. 225-244.

Samuels, S. J. Attentional processes in reading: The effects of pictures on the acquisition of reading responses. Journal of Educational Psychology, 1967, 58, pp. 337-342.

Samuels, S. J. Effects of pictures on learning to read, comprehension and attitudes. Review of Educational Research, 1970, 40, pp. 397-409.

Samuels, S. J., Biesbrock, E. & Terry, P. A. The effects of pictures on children's attitudes toward presented stories. The Journal of Educational Research, 1974, 67, pp. 243-246.

Saettler, P. Design and selection factors. Review of Educational Research, 1968, 38, 2, pp. 115-128.

Shulman, L. S. Reconstruction of educational research. Review of Educational Research, 1970, 40, 3, pp. 371-396.

Snow, R. E. Research on media and aptitudes. In G. Salomon & R. E. Snow (eds.), 1970, op. cit.

Snow, R. E. Representative and Quasi-representative designs for research on teaching. Paper read at the annual meeting of the American Educational Research Association, New Orleans, February 1973.

Snow, R. E., & Salomon, G. Aptitudes and instructional media. AV Communication Review, 1968, 16, 4, pp. 341-357.

Van de Bogart, E. North of the Namaskeag: A case study in Viewer-Active television. Maine Public Broadcasting Network, Orono, 1972.

TABLE 1

Amount of Post-Viewing Variance Accounted For By

Background, Initial Achievement and Exposure

(After Salomon, 1974b)

Variance accounted for on test of...	Source of variance	All background variables R^2	All pre-viewing tests $+R^2$	Total R^2	Exposure $+R^2$	Exposure F
Letter matching	Lower class	26.7%	21.1%	47.8%	16.3%	5.40*
	Middle class	14.8	36.0	50.8	4.3	1.96
Parts of the whole	Lower class	20.9	27.6	48.5	6.6	3.60
	Middle class	10.0	17.8	36.8	18.3	6.90*

* $P < .05$

Figure 1 : THE STAGED INNOVATION DESIGN
(After Campbell, 1969)

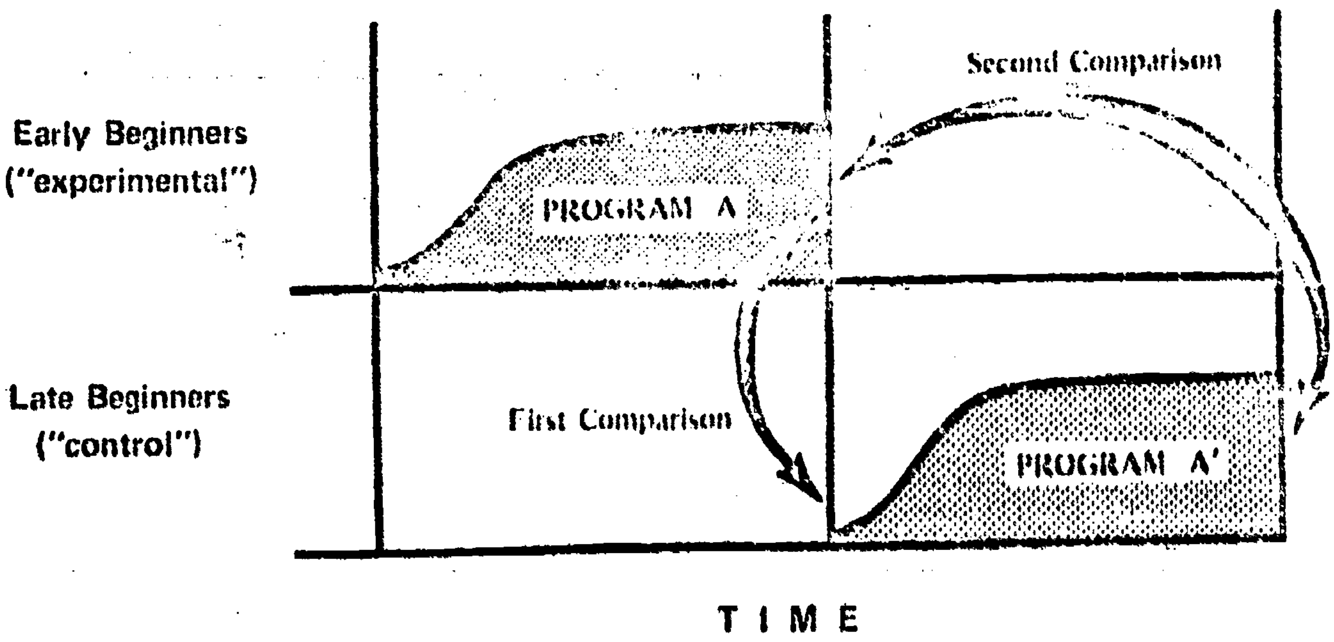

Figure 2 : THE ECOLOGICAL DESIGN
(based on Snow, 1973)

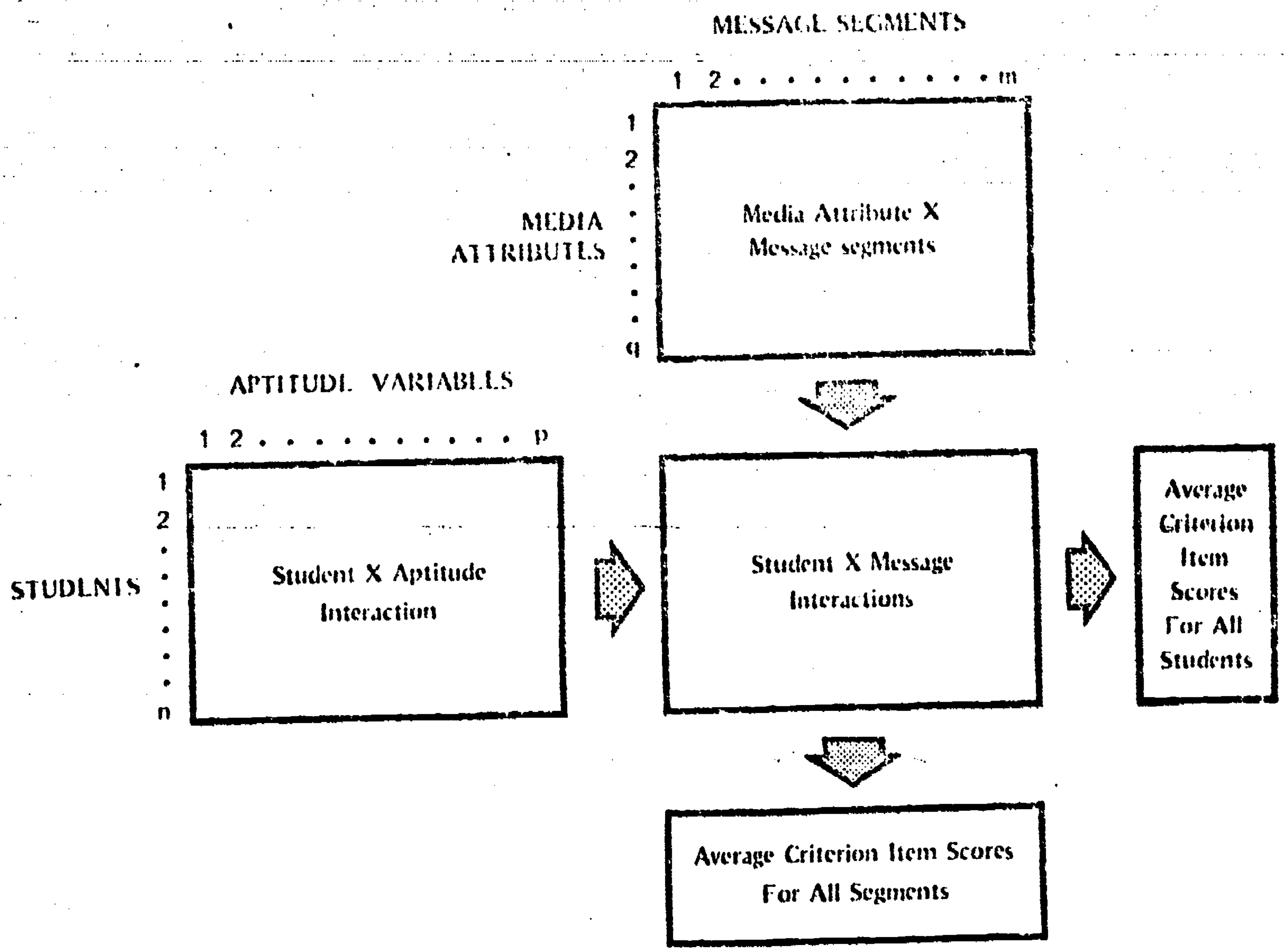

Figure 3. THE ROTATION DESIGN
(task x medium x learners experiment)

GROUPS

LEARNING TASKS	G_1	G_2	G_3	G_4
T_1	G_1 A* T_1	G_2 D T_1	G_3 C T_1	G_4 B T_1
T_2	G_1 B T_2	G_2 A T_2	G_3 D T_2	G_4 C T_2
T_3	G_1 C T_3	G_2 B T_3	G_3 A T_3	G_4 D T_3
T_4	G_1 D T_4	G_2 C T_4	G_3 B T_4	G_4 A T_4

* A, B, C, D, . . . - Different media attributes, technologies or different versions of the same medium.