DOCUMENT RESUME

ED 089 715 IR 000 427

AUTHOR Schramm, Wilbur
TITLE Notes on Instructional Cross-Media Comparisons. A Working Paper Written by Wilbur Schramm and a Committee, and Bearing Some Sign of Its Authorship.
INSTITUTION Stanford Univ., Calif. Inst. for Communication Research.
PUB DATE Dec 71
NOTE 54p.

EDRS PRICE MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS Comparative Analysis; Control Groups; *Developing Nations; Educational Research; Experimental Groups; *Field Studies; Instructional Media; Instructional Systems; *Laboratory Experiments; *Media Research; *Research Methodology; State of the Art Reviews
IDENTIFIERS *Cross Media Studies

ABSTRACT

Instructional cross-media experiments are difficult to design because they involve complex instructional systems which include variables that are hard to control. Strict controls can be approached in the laboratory setting, but as the experimenter moves from the laboratory to the field study, especially in developing countries, concessions to realism must be made. In these situations, the experimenter finds his control over conditions eroded by geographic isolation and the lack of financing and facilities, by the degree of commitment and interference of administrators, and by the moral imperative for total reform which denies the opportunity to exclude control groups from innovation. Nevertheless, cross-media field studies can be useful in evaluating the success of new systems, for conducting pilot studies, and for planning purposes, for what they lose in controls they gain in realism. Such studies are particularly effective when the researcher is aware of necessary trade-offs, provides the client group with descriptive and cost data on a full range of dependent variables, develops local tests, and helps to designate the maximally effective instructional system under a given set of environmental conditions. (PB)

NOTES ON INSTRUCTIONAL CROSS-MEDIA COMPARISONS

A working paper
written by Wilbur Schramm
and a committee, and bearing some
sign of its authorship

Institute for Communication Research
Stanford University
December, 1971

NOTES ON INSTRUCTIONAL CROSS-MEDIA COMPARISONS

This paper will be concerned with some of the problems of performing instructional cross-media experiments, and in particular with how to make such comparisons maximally useful to educators and educational planners in developing countries.

The particular problems

Most of the problems of designing cross-media experiments are common to all experiments, whether done in a physical laboratory, a Skinner box, or a school. But in one respect cross-media studies present a challenge that most other experiments do not. This is because the message is inexorably interwoven with the medium.

In a sense, "cross-media" is a misleading term. The chief source of variance is not the medium, but what it carries. The medium itself is a fairly simple delivery system; the total treatment, however, is infinitely varied. And for this reason it is almost impossible to make an experimentally derived statement about the relative effectiveness of television (or radio, or print, or films, or any other medium) without qualifying it by some other statements about content and intent.

When cross-media instructional experiments are performed, the situation is further complicated. Real-life instruction is carried on not by media, but rather by systems. "Conventional classroom instruction" (whatever that is) with which media teaching is compared in the great majority of "cross-media" experiments can only be called a medium by stretching that word out of all meaning. It is really a system of knowledge sources and learning opportunities. Teaching by media such as television or radio or films is also a system. The teleteacher is supported by various visual and auditory teaching aids and uses certain instructional strategies to motivate and reinforce the learner, and to build his activities into non-television learning opportunities that are carefully provided in the classroom. There is usually a classroom teacher who is an important part of the system. There is classroom interaction of many kinds. There are textbooks, workbooks, visual aids, projects, problems, guided practice. A cross-media comparison involving, say, television and classroom instruction, is not at all a comparison of television with the classroom, but rather of two armies of activities that vary from place to place and time to time.

In other words, the system is the treatment. When an experimenter tries to work in this field, therefore, he has a frustrating problem trying to dissect the treatment and control as many as possible of its components in order to

know precisely what he is comparing.

This poses an uncomfortable choice. If he emphasizes internal validity and scientific reproducibility, then he runs the risk of squeezing the realism out of his experimental situation, and losing external validity. If he tries primarily to answer questions of practical policy, then he runs the risk of not meeting the requirements of experimental research, and, indeed, of having his results apply to only one place and one situation. Therefore, most cross-media studies represent a trade-off at some level between science and realism, although both those qualities are desperately needed.

A few examples will suggest the nature of these trade-offs. Suppose one wants to control the teacher variable, so as to focus on other elements. One of the usual tactics is to use the same teacher for both experimental and control treatments. For example, we are going to describe, a little later in this paper, a design in which the same teacher taught a class on television and then went to a classroom and taught the same course face-to-face. This is obviously better than having two teachers work from the same lesson plan, because it should eliminate much of the difference in teaching style and ability. But it does not assure that the teacher will give the course in precisely the same way in both conditions; if he is a good teacher he will probably feel that the classroom situation requires something different from the studio

situation. Suppose, then, that the experimenter requires this teacher to learn the same script and give it verbatim before the camera and before the class. Then he faces the nagging doubt as to whether one or both systems will be short-changed; is the performance equally suitable for studio teaching and classroom teaching?

Suppose he tries to control the teacher variable more fully by using the same performance. This is usually done by moving a class into the studio, and using that as the "face-to-face" class. Think of the problems of realism involved in that situation. Is it a representative class when camera men and studio crews move around the classroom, bright lights glare, and people hold up visuals? Is it a typical teaching performance when a teacher is torn between looking into the camera and looking at his students, between speaking to his audience close at hand and his audience in distant classrooms, between interacting freely with his studio class and filling up the time of his televised classes who are not able to interact so directly with him? Each of these controls sacrifices some degree of realism and consequently of practical usefulness.

Suppose that the experimenter is trying to compare a one-track with a two-track medium, say, radio with television or films. If possible he would like to keep the common sensory material identical in order to find out what is the result of introducing the other sensory track. Some schol-

ars have tried to do this by using the sound track of the television or film program to represent radio, then measuring what happens to learning when the sight track is added. But would any skillful broadcaster use the same spoken commentary for television as for radio? Recall the difference between what a sports announcer says when he is broadcasting a football game by radio and by television. If he were to put the radio broadcast on television, viewers would say he is talking too much; if he were to put the television sound track on radio, they would say he is not telling them enough. So here again, an experimenter has to decide what trade-off he is willing to accept between control and realism.

Suppose that an experimenter wants to separate out the effect of what happens as a result of the television from what happens as a result of the surrounding activities. He can, of course, compare a lecture on television with a lecture in the classroom, and if he controls the content carefully enough he will probably have a clean experiment. But in most cases he would be limited to studying one performance and following it immediately with a test, because it would be rather extraordinary in anything except home study if there were no discussion, no follow-up activities after such a lecture. But again, how realistic is this? Most schools, where television is used to teach the core of a course, precede and follow the television program with motivation, review, practice, related readings, and exercises in applica-

tion. The most effective television teaching, so far as we can identify it, lays great emphasis on building a bridge to those related learning activities. In other words, instructional television really does function as part of a system. Take it out of the system and one has something that is not quite real and probably not very effective.

Highly abstracted cross-media comparisons typically use the very simplest subject matter in order to control variation. Travers' excellent experiments (1966) comparing learning from audio and visual and audiovisual tracks chiefly used nonsense syllables. Hartman (1961) found a highly ingenious way to use the two tracks of a film to study recognition through different "media" channels. On the upper half of the visual track he put pictures of 75 individuals identically dressed. On the lower half of the visual track he printed a name to be associated with each picture. On the sound track he had an announcer speak the name to be associated with the teacher. Thus by turning the sound track on or off, turning the visual track on or off, or covering half the visual track, he could arrive at seven different treatments: picture alone, picture plus print, picture plus sound, picture plus print plus sound, sound alone, sound plus print, or print alone. Each exposure was five seconds. The test was to look at or listen to 25 individual pictures or names, and decide whether each of them had been among the 75 individuals in the first presentation. Of course, these are laboratory experiments

intended to contribute to theory, and several levels of translation must occur before they can be directly applied in policy decisions.

For laboratory or quasi-laboratory experiments, the problem of controlling treatments is the most troublesome one. (We shall see that as experimenters move into the field, and most particularly when they work in developing countries, there are frequent problems also in controlling the experimental subjects.) But there is a particular problem of measurement that often plagues cross-media experiments, whether in the laboratory or the field, in economically advanced or newly developing countries.

For example, Hartman could not have tested persons who had seen only the pictures, by giving them a list of names, spoken or printed. Obviously, that would have been ridiculous. But something very like that occurs in even the most ordinary cross-media comparisons. For example, it is common practice to test achievement learning by paper and pencil examinations, usually multiple choice questions. Is this entirely fair, if the comparison is, let us say, between radio and television, or film and print, or television and classroom teaching? If the pupils have learned through one modality, is it entirely fair to test them on another? Suppose, for example, that in a television print or television-classroom comparison, the test were on visual discrimination or visual recognition; would the result be the same as

if the test were verbal and printed, and required verbal and written answers? So this is another problem that requires either a design complication or a compromise.

The field of cross-media studies reaches all the way from experiments like those of Travers and Hartman, intended to contribute to theory, to field studies intended to tell one particular school system how its own particular media-centered system compares with its own particular non-media system, and not claiming any generality. In this paper we are going to try to work between these two extremes, and concentrate particularly on the problems of field experiments where laboratory controls are not possible. Let us first look at some actual experiments to see how the problems of cross-media comparison were handled.

A model experiment

We shall begin with a very well-designed experiment, one of the ten which Stickell (1963) pronounced "interpretable" out of approximately 250 that he examined. This experiment, by Carpenter and Greenhill (1955), was one of a series done at Penn State in the 1950's to compare classroom instruction with closed circuit television instruction. A number of college-level courses were studied, but we are here going to report only on a study of the course in General Psychology.

Penn State was fortunate, at the time of this study, to have two buildings designed for trial use of closed-circuit television. The Sparks Building, which was used for the Psychology study, had a control room, a studio originating

room that was large enough for a studio class as well as performers and studio crew, and three classrooms each with two receivers.

The Psychology course was taught three hours a week in sections at different hours. Students who registered for the course were assigned randomly to one of the two hours, and to a television receiving room, the studio room, or a room in another building where the same class was taught without television. The random assignment was counted on to even out the inequalities in the different groups.

In order to equate the quality of instruction as fully as possible, two experienced teachers were assigned to the course. While teacher A taught the course on television at 8 o'clock Monday morning, teacher B taught it face-to-face in the other building. When it came time for the 1 o'clock section, teacher B taught the course on television and teacher A went to the other building and taught face-to-face. Thus on every class day each of the two instructors taught 40 students in the studio, 80 students by television, and 40 students face-to-face. There were no discussion sections; it was entirely a lecture course. Furthermore, there was no discussion or practice in the classroom. Except for the textbook and the examinations, television carried all the instructional load for the television sections.

This design helps, of course, to make the treatments comparable. If there had been discussion sections conducted

by graduate teaching assistants, as most lecture courses have, or if television had filled only part of the period with the remainder used by the classroom teacher, as typically happens when television is used for core teaching, then it would have been much more difficult to equate the instructional treatments.

Every effort was made to control the content of the teaching. The two instructors decided on the textbook, on the objectives of each day's teaching, and on the examinations, which were made carefully so as to include only what both instructors had taught. Beyond that, they were free to make the best use they could of the situation in which they taught. That is, there was nothing to insure that the television in the first section would be identical with (or as effective as) television in the later section; or that the early morning face-to-face class would be taught as effectively as the afternoon class. It might have been that one of the instructors had a special flare for television, or was especially effective in the atmosphere of the small classroom. If so, this might have biased the results, but the cross-over design would have helped to eliminate bias, and special interactions between teacher and medium would have been revealed by statistical analysis.

Looking back at what was done, it seems as though the situation, if it were biased at all, was biased for the face-to-face classroom. The situation for the studio class

must have been particularly awkward because those 40 students sat in an auditorium intended for 400, with the entire front of the room filled with a studio set and television gear: two cameras, microphones, monitor, four 750 watt spotlights, two 1500 watt and one 500 watt scoop lamps, giving 175 foot candles of illumination. The studio class was permitted to ask questions, but when a student did venture a question much of the sound was lost in the huge auditorium. The instructor must have been torn between his obligations to the class in front of him and to the TV cameras that were broadcasting to the other classrooms. The television teacher was allowed no rehearsal time; although he was eoncouraged to make such use of the medium as he saw fit there was no major effort to maximize the effect of the teleclass or make him an expert teleteacher. In fact, the experimenters noted that in front of the cameras the teachers did "what comes naturally"--taught about the same way as they had been teaching for a number of years in the classroom.

The television classes sat in ordinary classrooms equipped with two 24-inch television receivers each. An "observer" was present to keep order; otherwise, the class was on its own. The face-to-face classes also were in ordinary classrooms, and had the obvious advantage of being able to interact as fully as they wished with the instructor.

If the cards were slightly stacked against television, this was not completely undesirable because the purpose of

the experiment was to find out whether students taught by television in a relatively inexpensive way would learn as much as from the kind of classroom teaching they were accustomed to in college.

What were the results? Three term examinations and a final examination were given, as well as a number of attitude measures. In the second examination, the face-to-face class did significantly better than either of the others. In general, the studio class and the television-only classes tended to drag slightly behind the face-to-face class, but at the end there were no statistically significant differences. There was a significant difference in how much was learned from one teacher as compared with the other. The attitude measures were slightly less favorable to the course among the television sections than the classroom sections, but there were no significant differences in attitudes toward psychology or willingness to register for another course in the subject.

A field researcher from a developing country, looking at this experiment, could not help being impressed by two things. One is the amount of control it was possible to exert over the participants and conditions. The psychology department at Penn State had a number of students whom it could randomly assign without difficulty or objection and thus equate subjects. It could make a schedule in such a way as to allow two teachers to teach all of the experimental

conditions on the same day. It could secure the agreement of these teachers on what was to be taught and what the examinations were to cover, and could limit the class to a lecture course, so as to help in equating treatments. And it had an experimental television classroom building which made it easy to carry out the programs and distribute students among experimental conditions. This degree of control is not common in field situations, and especially in developing countries.

A second quality of this experiment that impresses one is the degree of realism that has been retained while maintaining careful control over experimental requirements. Random assignment of students is often very difficult in the field; here it was handled naturally by assigning students to sections of a very large undergraduate course. It is often hard to equate treatments, either in the field or in a quasi-laboratory situation like the one at Penn State. Here it was handled, like the random assignment, with a minimum of unnaturalness. True, not many new countries are willing to turn over all the teaching of a course to television. True, if a class is going to be taught in a studio it is probably there to help the teacher rather than the students. But any instructional experiment across-media is almost certain to require a trade-off between realism and control, and this one has reached an uncommonly successful balance.

What conclusions can be drawn from the results of this experiment?

It says that Penn State has proved it can build a Psychology lecture course around television, in which the kind of students who are at Penn State, when examined by paper and pencil tests, will probably score as high as they will in face-to-face classroom courses covering the same subject matter and taught by the same teachers.

Note that it does not say that instructional television is as effective as classroom teaching. That is a higher order of generality. However, an experiment done as carefully as this one should encourage another university, if it has some reason to use instructional television, to believe that it could do with ITV what Penn State was able to do.

A field experiment in a developing country

Let us now turn from the quasi-laboratory conditions under which the Penn State experiment was conducted to the mountains of Ecuador where a cross-media project was done under field conditions (Spector, et al., 1963).

This was a large and well-financed project, with adults for its experimental and control groups, conducted in a remote and isolated section of the lower Andes. Thus the researchers were able to avoid two of the problems that often plague instructional research in developing countries: lack of resources to do the job, and the need to change an old

and entrenched institution like traditional schools. That does not mean, however, that there were no problems.

The goal was to compare media for their ability to persuade people in rural communities to adopt new practices which were important because of their relationship to the development plan. The particular innovations chosen were building latrines, building stoves, canning marmalade, and submitting to vaccination for smallpox. The medium of greatest interest to the researchers and their sponsor was radio, because of its ability to overleap literacy and its relatively low cost. One experimental treatment, therefore, was to be radio, directed at persuading listeners to adopt the four practices listed above. A second treatment was to be a combination of audiovisual media: motion pictures, photographs, posters, slides, public speeches, and so forth. A third treatment was a combination of the two: radio plus audiovisual media. Each of these was to be compared also with a control group.

By this time the reader of this paper is probably wondering how the experimenters in Ecuador were able to control the content of the "audiovisual" treatment so that it was comparable, within experimental limits, to the content of the radio treatment. The answer is that they did not try to do so. Rather they outlined the broad arguments to be used and the principal information to be conveyed. Then they put the task into the hands of competent radio and audiovisual men,

and told them to make the best and most convincing programs they could. Competition was encouraged among the different production groups, in order to maximize the effect of each media system, rather than restricting it in any way by controls from doing the best job it could. In other words, the experiment was designed to compare the best versions of the different media systems that the research team could produce with the resources at hand.

The total production was about 250 hours of radio broadcasting and 250 hours of audiovisual entertainment and persuasion, during seven weeks. The radio time included a great variety of materials--spot announcements, jingles, music, a serial drama, several local series, news, and instruction. Radio receivers were handed out, and a radio transmitter was brought in especially for the project. The audiovisual treatment blanketed the experimental village with posters, photographic exhibits, illustrated bulletins, and so forth. One to three times a week, a "show" was held, including an entertainment film, a movie or slides on the recommended innovations, and a talk.

One experimental village was to hear only the radio. One was to get only the "audiovisual" treatment. A third received both radio and audiovisual treatments. And a fourth, the control, was to receive none of the treatments.

When one works in villages, rather than schools, it becomes more difficult to separate treatment groups. This

is especially noticeable when one is working with public media. In a school one can put a screen or a receiver in one room, assign experimental subjects to that room, and isolate their experience. But an outdoor movie or poster or public meeting is available to a whole village. A radio broadcast covers many villages and is available to anyone who has a chance to listen to a receiver. Consequently, there is a real problem in trying to assign individuals or groups to different treatments randomly. In an experiment like this one it seemed impossible to randomize subjects within villages, and even to sample by villages, because radio could not be excluded from any village within its coverage. There were enough receivers in any village to contaminate the results.

The experimenters in Ecuador did not find any magic way to solve this problem. They chose three isolated villages, some distance apart, and a fourth village--in another province--as a control. They did what they could to match the villages in size, and on certain demographic indices. Another "matching" item, they reported, was the willingness of the village government to join in the project--suggesting another problem of field work in a developing culture.

This purposive sampling, and the nature of the treatments, left little possibility for speaking with any great generality about the results of the experiment. In place of this, the experimenters tried to maintain the highest pos-

sible degree of realism in all the treatments.

The chief dependent variable was the number of people in each village who actually adopted or accepted one of the innovations--i.e., built a latrine, permitted themselves to be vaccinated for smallpox, and so forth. In addition to this, a number of interviews were conducted to find out what sources the people in the villages had found most persuasive or most helpful to them, and their reasons for adopting or not adopting. The design forced the researchers to use Chi-square for a significance test, rather than a more powerful statistic.

What was the finding? In general, radio seemed the most effective of the three treatments in bringing new practices into use, but the audiovisual elements were more effective in conveying detailed instructions.

The study is richer than our description may have made it sound. The post-experimental survey dredged up a good deal of information about how the subjects used the different media. It is difficult to generalize upon the results because one is not entirely sure what the experimental population represented, or even to speak very precisely about what was being compared--especially because of the miscellaneous quality of the audiovisual treatment. But the point is that, even without the tight experimental controls we should like to see in it, a study of this kind can produce some useful information. One thing this particular study

did was to provide additional evidence that radio--used under the conditions in which it would realistically have to operate in Ecuador--can be quite effective in development campaigns, and indeed compare quite favorably with a more complex and more expensive multi-media campaign. Thus the study doubtless encouraged other development planners to make use of television's less glamorous, less costly sister.

From Laboratory to field

An experimental study in the field, like the one we have just been talking about, operates under far more difficult conditions than a laboratory experiment. These are field problems rather than developing country problems, although location in a developing country makes its own contribution to the difficulty. To illustrate the effect of moving farther into a field setting, we have put together the following chart which inserts six cross-media comparisons between the Penn State and the Ecuador studies. All these studies were done in the United States. They were chosen not quite at random and arranged in rough order of progression from quasi-laboratory conditions to field conditions. In each case we have tried to indicate how the experimenters handled two of the most challenging problems of designs--how to control the subjects, and how to control the treatments.

SETTING	WHAT WAS COMPARED	HOW SUBJECTS WERE CONTROLLED	HOW TREATMENTS WERE CONTROLLED
Carpenter and Greenhill (1955) University department of psychology	TV vs classroom teaching--psychology course	Randomized assignment	Same teachers for both treatments
Seibert and Homig (1959) University department of chemistry	TV vs classroom, one session on laboratory techniques	Randomized assignment	Different teachers, but effort to teach same content
Deutschmann et al. (1961) University department of engineering	Film vs classroom, one session on laboratory techniques	Subjects selected own groups, but were compared on earlier records	Same teaching outline but different teachers, and no control over what students did in laboratory
Westley and Barrow (1959) Four primary schools near university	Radio vs TV, 6th grade course in news interpretation	Randomized assignment within each of four schools	Same script writer, same actors, but tried to maximize effect of each medium in its own way
Erickson and Chausow (1960) Home-bound students and students on city junior college campus	Junior college by ITV at home vs TV in classroom vs classroom without TV	Home students essentially different	Same curriculum, same tests, but different teachers
Rock et al. (1951) Nine military bases	Military reserve classes by TV, kinescope, and face-to-face teaching	No randomization	Kinescope merely film version of TV; TV and face-to-face made from same lesson plans, but different teachers, and no control on performance
Bryan (1961) Small high schools scattered throughout an entire state	Science teaching by TV plus correspondence vs TV plus visitations from college science majors vs TV plus correspondence plus visitations	Random assignment of schools to treatments, but high drop-out rate from experiment	Local teachers and visitors not controlled, and different graders on correspondence papers
Spector et al. (1963) Four isolated villages in Ecuador	Radio vs combined audiovisual campaign vs both of these, for innovation	Villages purposively selected, matched in certain characteristics	Agreement on factual content, but each treatment in its own best way

We are not going to describe these studies in detail; many of them will doubtless be familiar to readers of this paper, in any case. But the chart makes it apparent that control problems do indeed become more elusive as one moves farther from the quasi-laboratory conditions under which Carpenter and Greenhill worked toward the remote and isolated field setting in which Spector and his colleagues worked.

None of these other experiments is controlled quite so neatly as the Carpenter-Greenhill study. Seibert and Homig, working under conditions much like those at Penn State, were unable to use the same teachers for the different treatments. Deutschmann felt that he should let students select their own groups, as they would in a completely naturalistic situation; and he also used different teachers and found no realistic way to control closely what was actually done in the laboratory. Westley and Barrow tried to maximize the effectiveness of each treatment, using the two media in whatever way seemed best. Erickson and Chausow had a situation in which randomization was almost impossible, and in certain ways the home-bound students were essentially different from those on campus. Rock selected his sites purposively, and apparently did not have the right to randomize at each location; furthermore, because he was operating at widely separated points he could not use the same teachers or maintain any tight control over content of the different treatments. And Bryan, who worked in a number of small high schools, with volunteer students

(many of whom dropped out of the program) and with college students as visiting tutors, faced very difficult problems of controlling both subject groups and treatments.

This does not mean that media comparisons should not be made in the field, because unless they are made under field conditions they will always lack a certain degree of realism. Nor does it mean that a field study, made under difficult conditions with less than perfect controls will nevertheless not produce useful results. Rather, it implies that an experimenter should be aware of the difficulties he faces and make special efforts to overcome them or adopt the best possible design under the circumstances.

Therefore, it might be useful at this point to ask why such studies are so difficult under field conditions, and especially in developing countries.

Some special problems of instructional media comparisons in a developing country

Every experienced researcher knows it is harder to do research in a developing country or a remote area because of difficulties in travel and communication, lack of records, scarcity of local people trained in research skills, scarcity of even such supporting equipment as typewriters and calculators (not to mention computers), and the alienness of the whole idea of studying educational results in the spirit of science. Beyond these, however, there are certain special problems, of which we can give some examples:

1. The commitment of administrators

Campbell speaks sympathetically, as he should, of what he calls the "trapped" administrator. This is the man who has committed himself and his system to a program of educational reform that may well represent a major part of his country's entire expenditure on development. The reform is usually planned before any research goes in, and when the researcher enters, only the smaller details of the planning remain to be done; in many cases the reform is already in operation. In Samoa, for example, a broad educational reform built around intensive use of television was in operation for some years before there was much interest in studying it. During that time the entire energies and resources of the school system were required simply to institute the educational changes and keep the system operating.

In that kind of situation, an administrator, if he is not too fearful of the results, wants to know "how the reform is going," and often has some obligation to report that to a donor agency. But he is not anxious to hear any bad news, and above all--having committed his resources and system fully--is not much interested in setting up a rival plan to compare with it. For example, having committed his system to instructional television and spent a large amount of money on training staff, supplying transmitters, studios, receivers, and curricular materials to accompany television, he is not inclined to go to the same trouble to create a comparable

system of instructional radio. He is willing to have his new system compared with what is left of the old system, but even this presents problems, as we shall point out later.

Campbell speaks of the blessing of being able to work with an "experimental" administrator, one who believes in testing alternative decisions before he takes them. These administrators are rare. When one finds them, they usually do not have the money to put adequate experimental studies in the field before the major funds are obtained for the reform itself. But the best time to make cross-media system comparisons in a developing country is before the reform gets started. This suggests that donor agencies and donor countries would be well advised to encourage testing before granting large sums that will commit an educational system to a certain policy for years. They seldom do.

2. The moral imperative of reform

Once the educational reform is under way, there is an almost irresistible pressure to extend it to all pupils. The pressure is both political and moral: if the changes are worth making, should not their benefits be spread equally? It is very hard to answer that question in the negative. This was one of the reasons--in addition to administrative impatience--why Samoa decided to begin ITV in all grades at once, rather than a grade at a time. It is one of the chief reasons why most developing countries are uneasy about es-

tablishing control groups where a certain number of children will not have access to what they consider an "improved" kind of teaching and materials. In most educational reforms it would be useful to maintain such control groups for several years, so that comparative records of progress can be maintained beyond the first bright year of change. But this is extremely hard for a public school system to do. The result is usually that it is not done at all, or that the comparison is contaminated by gradually introducing the reform to the control groups.

3. The cost of comparative treatments

In a certain Latin American country while we were studying an educational reform built around ITV in the middle school (7th, 8th, 9th) grades, a decision was taken to introduce a new curriculum throughout the primary schools. The primary teachers needed help, and a series of television programs was planned for the in-service training of teachers in use of the new curriculum. But the question arose, what could radio do for the primary teachers--at a much lower unit cost?

It seemed like a very useful media system comparison. But it proved not to be feasible because it would have required making a parallel set of treatments to compare with the television treatments. Television teams and studios were on hand; a radio production team would have had to be put together, a studio obtained for at least a short time, and some radio receivers would have had to be purchased over and above

the cost of television receivers. The administrators did not feel able to make such a commitment of funds, time, and human resources. And thus a very good opportunity for comparative research was lost.

This situation is typical. In a major educational reform within a developing country, time, funds, and talents are always in short supply for what has to be done. The situation will never be any different until the spirit of experimentation spreads through both the donor agencies and the developing countries, and a certain part of the available resources are committed in advance to trying out alternatives before committing the whole country to one of them.

4. Difficulties in controlling conditions

The researcher in a developing country really has no control over the experimental conditions. He must always work through administrators who can exert control. These administrators have a number of good reasons for not wanting to alter the naturalistic situation. Some of these are political, some cultural, some geographic, some merely inertial.

On one occasion we had been able to set up a rather careful experimental design in a number of schools, each of which had two classes at the same grade level. We got permission to assign pupils randomly between each pair of classes, to control materials and procedures. The chief remaining source of probable variation was the classroom teach-

ers. We hoped to have them, too, assigned randomly, or at least to have them exchange classes in mid-year. This proved to be impossible, inasmuch as a rather delicate situation existed between teachers union and schools, and the administrators did not feel able to push the point. Looking back on the experiment later, we concluded that differences in the quality of the teachers did indeed substantially affect the results.

This is only one example. It is difficult to randomize students within one-room schools, or even schools with one class in each grade. One can randomize schools, but this too may cause problems. If one of the treatments includes broadcasting, there is a special problem. To withhold ITV or instructional radio from schools within the coverage area is to risk political repercussions and raise ethical and moral questions. Even when these questions are surmounted, there is always the chance of contamination by unintended reception. Beyond this there is the problem of trying to control treatments when broadcast crews or film makers are still learning their jobs, still harried by time schedules, and unaccustomed to being restricted in the way they teach.

These are a few examples. Any field researcher can add to them.

Some suggested strategies

The question is, what advice can be given to a donor agency, a field researcher, or an administering agency as to

how to use media comparisons as effectively as possible within such limitations as have been described. We are going to divide our suggestions into (a) research strategies and (b) research tactics, by which we mean (a) when a certain kind of media-comparison can best be done and (b) some considerations of doing it as well as possible.

For the present at least it is unlikely that many field experiments are going to be performed in developing countries per se--without reference to some ongoing or anticipated change. Therefore we can expect that the researcher will come into one of three situations:

1. When the commitment to a certain change, such as an educational reform, has already been made.

That is, he comes in to estimate the "success" of the new system. The donor, if there is an outside donor, wants to have the effect of his investment evaluated; he likes to see his judgment vindicated. The local administrator wants to know how things are going, although chiefly he seeks confirmation, and the changes he is prepared to make are small rather than major ones. The comparisons the researcher is going to be able to make are probably between the new media system and the previous one, as long as that persists.

This is a perfectly reasonable use of research, and yet our experience leads us to believe that if the activity is to be useful as possible inside and outside the experimental country much of the research emphasis might well be on some other matters than the comparison of new and old. For example:

a. Criterion testing of the new system. It would be surprising if the amount of effort put into a major educational reform did not produce higher achievement scores than the system it was intended to replace. More useful in many cases might be the act of stating very precise criteria for success of the new system, and measuring results against those. These might be behavioral objectives of the new curriculum; they might be standardized norms of achievement, by grade levels; they might be retention of pupils in school or performance after graduation from a certain level of school. It hardly needs saying that most educational plans and new curricula do not have such criterion objectives, but the act of making them seems to us an important part of reform planning. And if research can be used to test results against criteria, then it can contribute in a very important way to the function which we can call:

b. Maximizing the effectiveness of the new system. Comparing a new system with an old one does not usually tell one very clearly how to improve the new system. Furthermore, inasmuch as standardized and normed tests usually do not exist in developing countries, it is always difficult to tell <u>how</u> important a gain is. However, if a significant part of the research effort can be put on pretesting of materials, feedback research from the classrooms related directly to learning from elements of the treatment, and at least a limited number of comparison tests of materials, then results of that kind can be applied directly to improving the new system.

Before this can be done effectively, a spirit of experimentation and self-assessment must be built into the production teams and all the rest of the staff concerned with the teaching that goes into the classroom. This spirit has to begin with the donor agency's willingness to assign research funds to such less spectacular research, and it must be adopted by the administrators of the program so that they can convince the writers, producers, teleteachers, and others that science as well as art can contribute to their effectiveness, and that they can and should make use of research results. Furthermore, time must be left in the production schedule to try out, to remake, to review topics when research shows that they have not been successful. And finally, the research team itself must accept this less glamorous kind of research as a part of their task.

c. One of the ways that research can be most helpful outside the experimental country is through an ongoing study of the problems that occur in making the necessary educational changes, and how these problems were solved; the distribution of resources; the costs of the program; the steps involved in introducing and expanding the new program; the opposition to it, if any, and how this was met--in other words, policy and administrative questions. These are all problems that another country is likely to face when it, too, sets about making a major educational change.

2. When a country wants to extend its educational system or its out-of-school development services in a substantial way.

This situation offers the researcher a great deal more freedom than the previous one. Furthermore, he has much more chance to have an impact on policy. A pilot experiment comparing two systems might save a great deal of money later when the services are extended widely. If that is not feasible, there is always the performance standard of the existing system with which to compare the performance of one or more pilot projects in the field. For example, Mexico is beginning to offer secondary school by television to any village that is willing to provide a room and a teacher; the achievement in that system can be compared with that of the residence high schools in the cities. Mexico is also trying out a plan to extend three-year primary schools to six-year schools, by running double shifts and teaching most of the 4th, 5th, and 6th grades by radio. These situations are made to order for comparisons, which would have great practical import both in Mexico and elsewhere.

One of the chief attractions in both these Mexican situations is the possibility of saving substantially in costs over having to build and staff new schools. Therefore cost becomes itself a criterion, and the question, how much can a system get for what investment, becomes an experimental question. Here again one may use the cost and achievement of

comparable residence schools as a standard. For example, if Bavaria finds that it can teach a full-time equivalent student by television and correspondence in upper secondary for $300 per year whereas a residence student costs $600, that is a contribution to policy making. In a case where the rural curriculum will be different from existing curricula, one can test against specified goals of achievement and specified targets of cost. In other words, if cost is important in the kind of assessment required by condition 1, it is crucial in condition 2.

3. When a country is still in the planning stage of education change, far enough from the point of decision to be able to study alternatives.

This is the ideal time for cross-media experimental comparisons. Unfortunately the researcher is not often involved at this point in planning. Typically a country feels it must be sure of the resources before committing itself either to a major change or to substantial research looking toward major change. Therefore, it spends its time drawing up a plan for presentation to a donor agency, and the agency typically sends out a team to make a feasibility study, rather than ordering field research. When the feasibility study is made, if the agency is convinced, it makes the grant or the loan, By that time, it is too late for preliminary research to do much good.

If there is going to be any significant change in this

situation, then, there will have to be a change of attitude on the part of funding agencies. Feasibility studies are very useful, drawing as they do upon the experience and the intuitions of able persons; but they are no substitute for field research under realistic conditions. It would be rather refreshing if a funding agency were to make a grant to an educational system to find whether a or b or c is the most cost-effective way to meet a certain educational goal, with the understanding that if a certain standard can be reached funds will be made available for a major reform based on that finding.

This would challenge researchers to maintain the highest possible scientific quality and a maximum of realism. Indeed, here is the place for carefully controlled experiments. It would encourage factorial designs, amongst others, that would permit decisions among different combinations of treatments. Of course, we are not implying that once a successful pilot study has been done, there are no remaining problems in expanding the pilot nation-wide. That clearly is not the case; the attempt to expand the pilot radio rural forum over India showed what can go wrong. But there is no doubt that a procedure like this, if understandable impatience could be restrained, would start an educational reform on much sounder footing and with greater confidence, and might save much money and frustration.

If a different attitude on the part of funding agencies

is called for, so is a change of attitude on the part of many political administrators. As Campbell says, reforms are typically advocated as though they are sure to succeed. Consequently it becomes unthinkable, unacceptable, that they should not succeed. But one shift in political posture--quoting Campbell again--a shift from advocating a specific reform (say, instructional television) to advocating the seriousness of the problem and hence the importance of finding an adequate solution, would immediately take a great deal of the pressure off an administrator. Suppose, for example, he would say, this is a crucial problem, but we don't know whether solution A, solution B, solution C is the best answer to it. We are going to try them out on a small scale and decide which one to adopt. Or suppose he would say, this is a crucial problem, and we are going to try solution A (which seems highly promising). If there has been no significant improvement, we are going to change to solution B or something more promising that appears by that time. That kind of attitude would take much of the pressure off administrators and open the door to hard-headed research and honest findings.

Some suggestions on tactics

We have no intention of trying to tell an experienced researcher what to do in a field situation. For researchers less experienced in developing countries, however, these few suggestions may be of some interest.

1. The trade-off.

One seldom appreciates the necessary trade-off between realism and science, mentioned earlier in this paper, until he works at designing experiments, and especially media-comparison experiments, for developing countries. In this situation the great pressure on him is to produce a finding useful for policy making or educational practice. The person looking over his shoulder is not one of his colleagues but an educational administrator or a representative of the funding agency. Thus he is constrained to define his problem in practical and realistic terms, but he is also keenly aware that his findings are not going to remain on data sheets or go only into a report for other scholars; they are being depended on by administrators and planners as a guide in making policy decisions. Therefore, he needs both realism and science. He must design his research in terms relevant to the changes that are being made or contemplated in the system where he is working, and he must be as confident as possible in the results he obtains.

In that kind of situation he is likely to shy away, for example, from using a television sound track to represent radio in a media comparison. He is likely to compare systems, even though they are harder to control. He will sympathize (though perhaps not agree) with local administrators in their reluctance to withhold the reform from some students but not others, and in the difficulties they have in providing

alternative treatments, control groups, and randomized assignments. But in order to maximize the reliability and validity of his findings, he is likely to call on all his resources to patch up a design, to find quasi-experimental designs that are adequate to the purpose, to look at the entire problem of gathering and weighing evidence from a somewhat freer viewpoint. And this is what field research in developing countries is all about.

2. The need to describe

For his own sake and the sake of his readers, the researcher who conducts media comparison research in developing countries has a rather special responsibility to describe the elements with which he is working. The laboratory experimenter also feels such an obligation, of course, but it is rather because he typically works with abstracted and "operationalized" elements--something less varied than life--whereas the field researcher works with life in all its confusing variety.

The fact that Country A has been able to attain a certain degree of success with a certain project is encouraging both to Country A and to Country B, but before Country B can estimate the usefulness of this finding, it must know about it in considerable detail: what kinds of students were studied, under what conditions, and precisely what was the treatment they were given. Beyond the always pertinent information on what the project cost, what the money bought, and what kinds of problems had to be faced, the most useful

quantitative information Country B can derive from Country A's media project is its effect on a given kind of student. For example, what effect did Country A's treatment have on the rural schools, on the low socio-economic students, on the students who come to school originally speaking an indigenous language, on the students who are highly motivated or very little motivated? For example, the finding we reported earlier that ITV worked very effectively for highly motivated students at home is important information, even without an equivalent control group, for a country that also needs to teach students at home.

Therefore, the demographic and educational background of a student, some measure of his relative ability to handle the language, to read, to learn, some indication of his family background and his degree of motivation to study, are more important for a field researcher than a laboratory researcher to report. He needs to be able to describe the learning environment in some detail: the teacher's preparation and experience; the class size, the nature of the school, what went on in the classroom beside the principal component of the experimental treatment; attitudes and motivations; possibility of a "Hawthorne effect," and so forth.

If the treatment is a system, it is especially important to be able to describe that in detail. In this respect we are surprisingly weak because the theory does not yet exist to permit us to describe in shorthand, say, the content of

instructional television. That being the case, we must do as much as possible to describe the content in longhand. If the medium is ITV, what kind of ITV? How is it taught? What instructional strategies are used? Is it aimed at keeping the student active or passive? How much practice is given? There is just as much need to be able to describe what goes on in the classroom around the broadcast, and here too there is an opportunity to make a practical contribution to a field that is somewhat heavy with elaboration. How can we describe, simply, the teacher's way of teaching, the interaction that takes place between teacher and student, and among students? What kinds of questions are asked, what kind of discussion takes place, what sorts of projects are undertaken, how much work is done on individual initiative? It hardly needs saying that what happens in the classroom can be both an independent and a dependent variable; for if it helps to determine how much and what is learned, so also a change in the media system should have an effect on teaching and interaction in the classroom.

3. The dependent variables

Achievement tests and attitude scales are necessary but hardly sufficient for the best media comparisons. There is a challenge in media studies to advance the state of the art in measuring effects of instruction. What else beside what we measure by achievement scores and attitude responses is happening to the student as a result of being taught in dif-

ferent ways? What cognitive skills is he learning? Is he becoming skillful in dealing with abstractions, in his ability to discriminate among concepts and to apply a principle to a problem? What happens to his social behavior and his interest in school? What changes take place in his education and career aspirations? What does television do to his ability to learn from visual models? We are not suggesting that every cross-media study should answer all these questions, but rather that a series of such studies should progressively contribute to our knowledge of effects. We suspect that, as this work goes on, unobtrusive tests such as attendance, dropouts, post-school career, observation of social behavior, group relationships, will be used more than they have been. But it is clear that if media and media systems in education are to be satisfactorily compared, a broader view of the effects of instructional media is called for.

4. Testing

In view of the need just expressed, it is well to say a word about the special need in media comparison projects for competence in making tests. Few developing countries have standardized tests. Where they have ability tests, these are usually translations or adaptations of tests made for other cultures. They have few if any tests aimed at measuring things like cognitive effects. Therefore, the ability to make criterion tests and to adapt or construct tests for

special purposes is a matter of real concern. Criterion tests, in particular, are best made within the system if they are to reflect system goals. Other tests also are better made within the system if they are to reflect the constraints of the culture. Consequently providing test competence only in a visiting research team is not the best way to meet the problem. The training or hiring of an adequate staff, within the local table of organization and closely related to curricular changes and preparation of materials, should be one of the early objectives of an instructional research program in a developing country.

5. Studying cost

At the moment, cost is such an important element in the policy making of new countries that it can hardly be left out of a study that reports on the relative effectiveness of instructional media. "Effectiveness at what cost?" the planner or administrator asks. What did they buy for what they spent? Strangely enough, relatively few cross-media studies have included this kind of information. Only one of the eight studies described earlier in this paper had anything to say about cost. It is interesting to consider the possibility that cost might even enter into the design as a sort of control. For example, a treatment might represent the best course on general science, taught by means of a system built around instructional radio or television,

that Country X could produce on a given budget. That would give more information to Country Y as to whether Country X's accomplishment has any relevance for Y.

6. The "maximum effectiveness" comparison

The practice of comparing media instructional systems that have been made as effective as possible, within the constraints of local resources, is likely to recur more and more often in cross-media comparisons, especially in field situations. This is the question that really concerns policy makers. Mexico wants to know how the best course they can build around television (within their resources) for village students, compares with the best classroom course they have been able to develop for students who are in school; and how the best 4th, 5th, and 6th grades they can build around radio compare with the 4th, 5th, and 6th grades they have been able to develop in the classroom. They do not care particularly whether the same teacher is on television and in the classroom, nor whether the material on television is identical or almost identical with the exposition in the classroom. They want to know whether the system they can construct will do a good job of meeting their curricular needs and what it will cost to do that in comparison to what it would cost to do it in school.

This challenges a researcher to find out what makes a maximally effective system. By factorial designs, he can often say something about the results of different combina-

tions. By entering into what we might call production research, he can test different kinds of programs, different kinds of classroom activities, different combinations of learning experiences. We regard this as a potentially important contribution to the development of both theory and practice in the next decade. By testing single elements and combinations, on limited samples, a researcher can made use of some experimental designs that are denied him in field research. By studying the effectiveness of different programs or materials he is likely to advance knowledge of content variables. By working intensively on program components as well as extensively on field data, he is in position to balance a research program in a way that most such programs have not been balanced.

If this is to be the shape of future research programs connected, for example, with educational reforms, then funding agencies must recognize that it will cost more, and require more, and perhaps different kinds of, researchers. But all our experience indicates that an emphasis on the testing and improvement of combinations, preliminary to their use in an overall instructional system, would be an extraordinarily good investment in quality and in knowledge.

7. Non-quantitative results

In our preoccupation with experiments we have neglected to point out that some of the most useful and vaulable results of attaching a research team to a project may well turn

out to be non-quantitative ones. We are referring to the recording and analysis of problems that occur in the introduction and operation of the project; the arrangements that prove to be necessary for training personnel, for organizing a staff, for maintaining two-way contact with the schools or groups that are served, for installing and maintaining the technical equipment; and the scheduling of all the preparations and activities that enter into a major educational change. These are problems that any other administrator must face; he can benefit from knowing about them in advance, and learning how they were handled in another country. Assigning a research team, in part, to this kind of work will require some redefinition of task on the part of both the funding agency and many research organizations, and the inclusion in the team of social researchers of different kinds than those who might be sent to do experimental research.

8. Repetitions

Repetitions of instructional media field experiments are relatively few (except the long series of television or film vs face-to-face studies, most of which have found n.s.d.). The reasons are not hard to see. Such research is expensive. Projects are not alike. "System" comparisons introduce so much variability that one is often unsure just what is being replicated.

This is in many respects unfortunate. It contrasts unpleasantly with experimentation in the physical sciences and in a field like animal learning, where important experiments are likely to be repeated dozens, even hundreds of times, with a resulting increase in confidence in the results. The question is, what can be done considering the difficulties and variabilities in the kind of research we have been talking about?

It seems to us that repetition built into research projects is relatively more important in this field than in most others. For example, if in three successive years we get corresponding results in the same grade with the same or equivalent materials, we have more confidencein that finding. If we find that Treatment T works well on samples A, B, D, and E, but not on C, then we are in better position to say where and when we can expect good results from the treatment. If we find, in one year, that a certain learning experience is notably more effective with one kind of sub-group than others, then it may pay to test that again in the following year.

As this kind of study becomes more common and better known, experimenters should try to build into their designs a replication of one or more interesting findings from another project or another design. And in the meantime, we have a few comparable projects in different countries that could even now be analyzed jointly with profit, and lead to

further testing of key points. For example, in at least seven countries the combination of correspondence study with radio or television, outside school, has been tried. We do not have experimental results on many of these projects, but we do have a number of cost figures, enrollment and retention data, student grades, records of student background, and the like. This combination has greatly interested many countries, who see in it a possible way to avoid some of the expenses of building more campuses and schools. A comparative analysis of findings from these projects might well lead, not only to a better understanding of the correspondence-broadcast system, but also to experimental research that would help to clear up questions of effectiveness which are not answered clearly by the existing data.

The usefulness of quasi-experimental designs

Kurt Lewin once said, with tongue in cheek, that whatever was socially interesting could not be measured reliably, whereas whatever could be measured reliably was likely to be of little social interest. This is not a viewpoint we are prepared to espouse, and yet it is one that will strike a responsive chord in any field researcher working on media comparisons.

In a situation where true experimental controls are simply not available to the researcher, and especially where he finds it impossible to randomize his whole population of subjects, quasi-experimental designs have much to commend

them. As a matter of fact, most of the better media comparisons in developing countries have been done with quasi-experimental designs, although they are not always so advertised. The true experiment is seldom seen in the field, unless the field is defined to include the agriculture experiment station.

For this reason, a paper like the chapter by Campbell and Stanley in the Handbook of Research on Teaching (1963), and Campbell's article on "Reforms as Experiments," in the American Psychologist (1969), are priority reading for anyone contemplating instructional media studies in developing countries. These authors suggest a number of quasi-experimental designs that do not require randomization, and nevertheless have considerable promise. Furthermore, the defects of these designs are known and can to a certain extent be compensated for.

The great majority of media-comparison experiments in developing countries are done with non-equivalent control groups, "equated" by some kind of statistical treatment or "matched" on the basis of a pretest or other previous scores. For example, when we studied the effect of introducing ITV into the 7th grade in El Salvador, we could not randomly select treatment and control groups. The reform was being introduced into one part of the country, not into others; even in schools where there were two 7th grade classes, it was morally and politically unacceptable to give one randomly selected group of students the new system, and withhold

these improvements from the other group. Consequently we used covariance analysis to "adjust" the figures on the basis of differences in ability and previous performance between the groups. This is not a satisfactory substitute for random selection, but it was the best alternative open to us.

Campbell and Stanley point out that any acceptable kind of control group will strengthen a design like this one. One non-equivalent control group design used the second year class at Annapolis as experimental group, and third year class as control. This at least helped to eliminate the possibility that whatever learning or maturation that goes on during the second year at Annapolis was responsible, rather than the treatment, for the results obtained.

In any use of non-randomized subjects, however, the researcher has a considerable responsibility to consider what causes, other than the one studied, might have entered into the effects. For example, in a non-equivalent control group design

O X O
O O

one must ask whether the groups differed significantly in ability, previous knowledge, age, or the kind of differential incentive offered in some cultures by different sex roles, socio-economic status, family background, etc. If the control over time is not perfect, one must ask also whether anything could have happened during the course of the experiment to affect the results in one group, but not in the

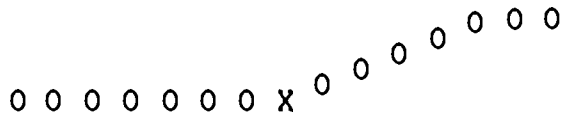
other. For example, in one study where we were able to randomize the subjects, we suspected nevertheless that the teachers infused a competitive spirit into one of the groups that motivated them more than the other groups. Any experiment that lasts through a term or a school year is likely to have some contaminating elements infused into it by things that happen in one group but not in others. The lesson is, not that one should avoid long-term experiments, but that when complete controls are not possible, the experimenter had better know a great deal about his subjects and what happens to them.

In certain situations, other quasi-experimental designs, like Campbell and Stanley's Time Series (number 8), or the Multiple Time Series (number 14), are especially attractive. These consist essentially of measuring performance on some scale at regular time intervals. Somewhere in the series the experimental treatment is introduced, and the experimenters want to know what happens to the series of measurements after that.

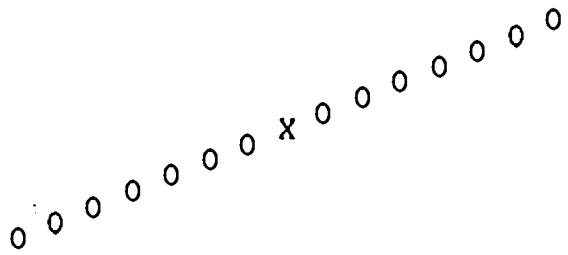
For example, consider a study of family planning that has been under way in Iran. The dependent variable was the number of persons who applied each month to the family planning clinics in one province. A mass media campaign in support of family planning was introduced midway in the series of measurements. Thus the design was

O O O O O O O X O O O O O O O

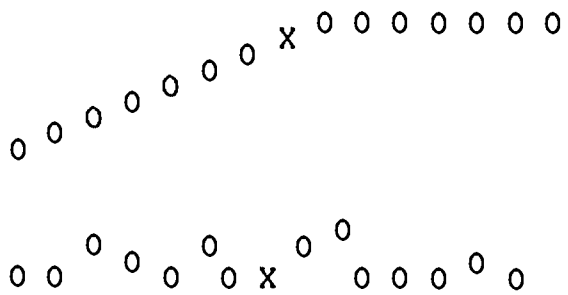
Now suppose the measurements looked like this:



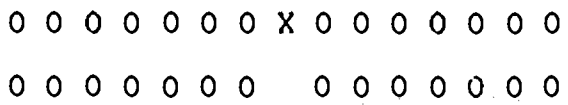
In that case, there would be considerable confidence that the experimental treatment did indeed have an effect. But if it looked this way:



Then the effect would have to be regarded as not proved, and similarly in either of these cases:



Now if a control group could have been added--for example if simultaneous measurements could have been made in a comparable area of the country where no new campaign was being offered--then the experimenters could feel somewhat more confident that what happened in the experimental population was indeed the result of the campaign. The design would have looked like this:



Another interesting design, usable when only one experimental group is available, but when one can control what is presented to them and the times of presentation, is what Campbell and Stanley call Equivalent Time Samples (number 8). This was used by Kerr to examine the effects of music on industrial productivity. He played music during a large sample of randomly selected days, and compared productivity on those days with productivity on an equivalent sample when music was not played. Of course, the result generalized only to the particular group of workers studied. It could have been generalized further by replicating the experiment on other groups. A similar experiment could be conducted in a classroom where oral practice during a television presentation could be required on a randomly selected series of days, and not on other days, the results measured each time.

A related design is what Campbell and Stanley call the Equivalent Materials design (number 9) in which two kinds of material can be compared with a single group on randomly selected occasions.

Finally, the so-called "counterbalanced" design (Campbell and Stanley number 11) has certain attractions. This is, in effect a Latin square design and can be diagrammed in this way (X_1, X_2, etc., representing different treatments):

	Time 1	Time 2	Time 3	Time 4
Group A	X_1O	X_2O	X_3O	X_4O
Group B	X_2O	X_4O	X_1O	X_3O
Group C	X_3O	X_1O	X_4O	X_2O
Group D	X_4O	X_3O	X_2O	X_1O

These designs are discussed in a highly sophisticated way in the chapter we have cited, and that, rather than further comment by us, is what a media-comparison researcher should read. The principle we should like to recommend is simply that a researcher in a developing country, facing restrictions on random selection and control of treatments, should neither give up the task nor be content with dirty research, but rather should use less than ultimate designs as long as he is aware what he is doing, is willing to patch them up by supplementary work and additional analysis, and is scrupulous in not claiming more than he has found. Given the best designs he can find for his purpose, given ingenuity, he can still make highly useful contributions to policy and practice.

BIBLIOGRAPHY

Bryan, E. F. A comparative study in the teaching of high school chemistry and physics. Oklahoma City: Oklahoma State Department of Education, September, 1961.

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

Campbell, D. T., and Stanley, J. Experiments and quasi-experimental designs for research on teaching. In Gage, N. L., ed. Handbook of research on teaching. Chicago: Rand McNally, 1963. pp. 171-246.

Carpenter, C. R., and Greenhill, L. P. An investigation of closed-circuit television for teaching university courses. University Park, Pennsylvania: The Pennsylvania State University, June 31, 1955.

Deutschmann, P. J., Barrow, L. C., jr., and McMillan, A. The efficiency of different modes of communication. Audiovisual Communication Review, November-December, 1961, 9, 6, 263-270.

Erickson, C. G., and Chausow, H. M. Chicago's TV college; final report of a three year experiment. Chicago: Chicago City Junior College, August, 1960.

Hartman, F. R. Recognition learning under multiple channel presentation. Audiovisual Communication Review, January-February, 1961, 9, 1, 24-43.

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

Rock, R. T., jr, Duva, J. S., and Murray, J. E. The effectiveness of television instruction in training naval air reserves. Final project report, Project NR 781-007. New York: Fordham University, April, 1951.

Seibert, W. F., and Honig, J. M. A brief study of televised laboratory instruction. TVPR Report, No. 8, February, 1959. West Lafayette, Indiana: Purdue University.

Spector, P., Torres, A., Lichtenstein, S., and Preston, H. O. Communication and motivation in community development: an experiment. Report of Phase I, submitted to the Agency for International Development, November, 1963.

Stickell, D. W. A critical review of the methodology and results of research comparing televised and face-to-face instruction. Thesis for the degree of Doctor of Education, Pennsylvania State University, June, 1963.

Travers, R. M. W., et al. Research and theory related to audiovisual information transmission. Salt Lake City: University of Utah, Bureau of Educational Research, 1966.

Westley, B. H., and Barrow, L. C., jr. Exploring the news: a comparative study of the teaching effectiveness of radio and television. Madison, Wisconsin: University of Wisconsin Television Laboratory, Research Report No. 12, May, 1959.