R E P O R T   R E S U M E S

ED 014 913   56   EM 006 019

AN INVESTIGATION OF THE RELATIVE EFFECTIVENESS OF CERTAIN SPECIFIC TV TECHNIQUES ON LEARNING. FINAL REPORT.

BY- SCHWARZWALDER, JOHN C.
TWIN CITY AREA EDUCATIONAL TV CORP., ST.PAUL, MINN.

REPORT NUMBER NDEA-VIIA-085-FR   PUB DATE   60
SAINT PAUL PUBLIC SCHOOLS, MINN.
EDRS PRICE MF-$0.25 HC-$2.32   56P.

DESCRIPTORS- *EDUCATIONAL TELEVISION, *VISUAL STIMULI, *AURAL STIMULI, *REINFORCEMENT, *ACADEMIC ABILITY, SCIENCE INSTRUCTION, ACADEMIC ACHIEVEMENT, ANALYSIS OF VARIANCE, TEACHING METHODS, ST. PAUL, MINNESOTA

GOAL OF THIS STUDY WAS TO ASSESS THE EFFECTS OF CERTAIN AUDITORY AND VISUAL STIMULI PRESENTED BY EDUCATIONAL TV. SPECIFIC QUESTIONS ASKED WHETHER VISUAL REINFORCEMENT, CONTINUITY, AND MANIPULATION TECHNIQUES INCREASE MASTERY OF SCIENCE INFORMATION BY FIFTH GRADERS. 72 TV LESSONS IN 9 AREAS OF SCIENCE, VARYING COMBINATIONS OF THE 3 TECHNIQUES AT EACH OF 2 LEVELS, WERE PRODUCED. CLASSROOM, INSTRUCTORS, EQUIPMENT, SELECTION, AND TESTING VARIABLES WERE CONTROLLED. A FACTORIAL DESIGN TO TEST EFFECTS OF THE TECHNIQUES AND THEIR INTERACTIONS BY ANALYSIS OF VARIANCE WAS EMPLOYED. FOR 40 CLASSES, PRE-TEST SCORES OF MENTAL ABILITY WERE USED AS COVARIATES WITH IMMEDIATE POST-TEST SCORES OF ACHIEVEMENT. THE AUTHORS USED THE .10 LEVEL OF SIGNIFICANCE BECAUSE EDUCATIONAL TV RESEARCH WAS A NEW AND UNREFINED FIELD. MISSING DATA WERE PREDICTED FROM A REGRESSION EQUATION. 22 OF THE 63 DIFFERENCES WERE REPORTED AS SIGNIFICANT. OF THESE 20 WERE SIGNIFICANT AT THE .05 LEVEL. MOST OF THESE DIFFERENCES IN ACHIEVEMENT RESULTED FROM INTERACTION OF TWO TECHNIQUES, SHOWING THAT THE TV APPEARED TO FACILITATE TEACHER EFFORTS. IMPLICATIONS FOR FURTHER RESEARCH ARE DISCUSSED. (LH)

A Final Report

TO

THE CHIEF, RESEARCH SECTION

EDUCATION MEDIA BRANCH

OFFICE OF EDUCATION

DEPARTMENT OF HEALTH, EDUCATION AND WELFARE

From

JOHN C. SCHWARZWALDER, PRINCIPAL INVESTIGATOR

For

A RESEARCH PROJECT ENTITLED

"AN INVESTIGATION OF THE RELATIVE EFFECTIVENESS OF CERTAIN SPECIFIC TV TECHNIQUES ON LEARNING"

A Final Report

TO

THE CHIEF, RESEARCH SECTION

EDUCATION MEDIA BRANCH

OFFICE OF EDUCATION

DEPARTMENT OF HEALTH, EDUCATION AND WELFARE

From

JOHN C. SCHWARZWALDER, PRINCIPAL INVESTIGATOR

For

A RESEARCH PROJECT ENTITLED

"AN INVESTIGATION OF THE RELATIVE EFFECTIVENESS OF CERTAIN SPECIFIC TV TECHNIQUES ON LEARNING"

This research was conducted by personnel from the St. Paul Public Schools and the Twin City Area Educational Television Corporation and was supported in large part with funds provided by The Department of Health, Education and Welfare under the provisions of Title VII of the National Defense Education Act of 1958.

TABLE OF CONTENTS

KTCA-TV

TWIN CITY AREA EDUCATIONAL TELEVISION CORPORATION

CHANNEL 2

MINNEAPOLIS • ST. PAUL

MINNESOTA

1640 Como Avenue
St. Paul 13, Minnesota

JOHN C. SCHWARZWALDER
General Manager

Telephone
Midway 5-5565

December 31, 1960

Letter of Transmittal

The Chief, Research Section
Educational Media Branch
Office of Education
Department of Health, Education
and Welfare
Washington 25, D.C.

Dear Sir,

This is the final report of a Research Project entitled "An Investigation of the Relative Effectiveness of Certain Specific TV Techniques on Learning" which was supported in large part with funds provided by the Department of Health, Education and Welfare under the provisions of the National Defense Education Act of 1958.

The names of those who were directly involved in the project and their capacities follow:

John C. Schwarzwalder	Principal Investigator
Nolan C. Kearney	Principal Investigator & Chief Consultant
Richard I. Evans	Psychological Consultant
Roland Buchman	Research Assistant
Carl H. Ruble	Producer-Director*
Paul H. Owen	Television Consultant
William Schrankler	Teacher-Presenter
Theodore Johnson	Science Advisor
Robert Shager	Lesson Plan Advisor
Thomas H. Engebretson	Statistical Aide
Guy W. Ueckert	Producer-Director **
Warren J. Panushka	Curriculum Advisor
Bette Herzog	Administrative Assistant
Arnold Thorson	Accountant

* During planning and inception of the research
** During actual filming of kinescopes

Test-Item Team

James E. Inskeep, Jr.
Lelon R. Capps
Val E. Arnsdorf
Ronald T. Lambert
Theodore Johnson
Roland Buchman

Kinescope-Rating Team

Lelon Capps
Edgar B. Williams
Helen C. Ambrose
Virginia Anderson
James J. Shannon
Kenneth F. Stillwell
Bruce H. Montgomery
Lilliam F. Wright
Karen Johnson
Evelyn Pearson
Fred B. Clark
Angela Burque
Leonard C. Larson
Warren J. Panushka
Robert Shager

IBM Analyst

Beno Ranweiler

Particular attention should be given to the work of Messrs. Kearney, Owen, Evans, Buchman, Schrankler, Johnson, Ruble and Ueckert. Dr. Kearney coordinated the project and did a major portion of the editing. Mr. Owen supervised the activities of Mr. Ruble and Mr. Ueckert who, in turn, effectively controlled a large television staff. Dr. Evans' critical evaluation and revision of parts of the penultimate form of the report resulted in a measureable improvement of the final product. The planning and execution of the design and analysis were almost entirely the work of Mr. Buchman. Few advisors are as helpful and useful as Mr. Johnson and no teachers are more devoted than Mr. Schrankler.

A number of statistical aspects of the results have been analyzed in addition to those mentioned in the original proposal. Particularly is this true as to whether results depended upon the intelligence of the viewer.

I believe that the results of this investigation will have the most far-reaching results in that they demonstrate a research method whereby the "no significant difference" pattern may be broken and new strata of research possibilities can be opened up.

I am grateful to the Office of Education for this opportunity and it is my hope that future investigators will share this gratitude and extend these results.

Sincerely,

John C. Schwarzwalder
Principal Investigator

CHAPTER I

THE PROBLEM

Statement of the Problem

The problem in this research is to assess the effectiveness of certain techniques of presentation in television programs designed for educational purposes. These techniques involve auditory and visual stimuli.

Ever since the advent of television, educators and others have regarded it with varying degrees of warmth as a potential educational tool. The problem is to measure the effect on the learning of pupils of various ways in which materials can be organized. Specific problems are to find out (1) whether or not the mastery of various types of science information and skill is increased by the use of visual reinforcement, (2) whether or not learning is affected by visual continuity, and (3) whether or not learning is affected by the manipulation of the visual presentations.

Stated in slightly different words, the problem is to throw some light on how best to use television as an instructional tool, and the research is restricted to a consideration of techniques in the arrangement and presentation of televised materials.

Local Background for the Study

During the year preceding the planning of this research project, KTCA-TV, Channel 2, had been cooperating with the St. Paul Public Schools in presenting (among many programs), science demonstrations at the fifth grade level. Staff members at KTCA and in the St. Paul Public Schools became interested during this time in developing new and improved techniques for presenting science demonstrations on television. In this work, there was need for information regarding the best way to handle demonstration materials and the various visual aids designed to emphasize aspects of the lessons that were necessary if pupils were to understand. Very little published material of a research nature was found concerning the best way to present demonstration programs on television. There were some discussions based on opinion and experience. There was a lack of information indicating whether or not the use of certain techniques actually resulted in increased learning on the part of students.

Thus there grew up a wish to undertake a controlled testing program to determine the relative effectiveness of certain techniques or factors. The resultant project does not attempt, of course, to investigate all of the factors upon which information is needed. For example, it does not include such things as the appropriate amount of time that a visual image should be kept on a TV screen, the timing of camera shots, or whether the shots should go on the air before the cue word is said or after the cue word is said.

Many factors need investigation besides those covered in the present research. Aside from the three mentioned above, there are such items as verbalization rate, appropriate size of letters on the TV screen, the number of lines and letters per line that should appear on the screen at any one time, the ideal length of a program period, plain background versus figured background, the degree of simplicity or ornateness of visual material, and the relative effectiveness of black on white background and white on black background.

Some General Considerations

There is a tendency in discussing educational television to compare it to instructional films, in some cases pointing out dissimilarities and in some cases pointing out similarities. So far we do not have much to go on in making such comparisons. It would seem reasonable that the two media have more similarities than differences as a tool to be used in stimulating learning. They both are moving pictures with synchronized sound effects. This is their great basic similarity. The differences are more numerous if not more compelling. The visible differences are in such things as size and photographic detail. The advent of video tape recording, which due to its novelty has received little attention as yet from researchers, suggests additional ramifications in this entire question.

The mechanical differences are numerous and pronounced. Films and video tape can be cut and changed after they have been filmed. By the same token, they sometimes lack the timeliness and spontaneity of "live" television programs. With films and video tape there may be, in some instances, latitude in planning and in adding cues and titles. Films and tapes can be interrupted while the classroom teacher presents supplementary material. They can be presented at the period and day where they fit best into the work of the class. They can be viewed by small groups with the use of earphones or loudspeakers. Educational films and video tapes can be evaluated before use. The narrator's script can be altered and edited. On the other hand, in "live" television, much of the editing must be done while the program is being broadcast. The "live" television broadcast may have certain hazards that are not possible in film presentations. For example, the television director must make instant and irrevocable decisions about how long a shot shall be on the air and when to shift from one camera to another. He cannot erase or alter unfortunate ad lib remarks or unforeseen departures from the program's outline. By the same token the vitality and spontaneity of an educational program may be entirely lost if the teacher knows that editing is possible.

The part played by the teacher in "live" television is thus quite different than the part played by the teacher in educational films. In educational television, the teacher knows he must do well, while in the film, the teacher's work, if poor, can be edited and changed. The incentive to fine teaching may, therefore, be lacking in film or video tape as compared to "live" television. These are merely illustrations of many differences.

It is often said that the demand for great teachers with television skills exceeds the supply. Some people assume that a certain type of personality projects itself with a greater resultant degree of learning on the part of the pupil than do others. Reference is often made to people who made a reputation for themselves as television teachers. However, there are seldom any references to research that indicates the excellence of such teachers in terms of the amount of learning of the pupils. In fact such research has simply not been undertaken. It is often pointed out that when the classroom teacher appears on television, he brings with him his classroom techniques and his classroom habits. These are sometimes ill-fitted for television presentations, particularly as they involve the use of charts, blackboards, and demonstrations.

This whole problem of the evaluation of the teaching personality is important enough to justify a few words here. We sometimes hear it said that the whole problem of educational television would be a simple one if we could find teachers of high caliber and put them before the TV cameras. Examples of such teaching are often cited. It cannot be reiterated too often that there is no evidence that such programs were exceptionally fruitful in the learning that resulted from them. Neither is there evidence that this is not the case. We need further research in this area. We know from research in other areas that certain types of teachers, certain types of discipline, certain types of deportment on the part of the teacher gain almost universal adult approval but do not always show themselves positively in the learning of the children. We cannot assume that because a program is artistic, startling, logical to adult minds, or presented by handsome, personable people who have well modulated voices, that it is necessarily the kind of a program from which children will learn. Neither, at this time, can we assume the contrary. We need conclusive evidence about such things based on careful research. The fact that an educational television broadcast attracts an increasingly wide audience is no sure indication that the program educates. Yet, a large audience is no criterion that a program lacks educational value. There are a host of hypotheses that the doubter may adopt. However, the justification for an educational television program is that people learn from viewing it. This outcome must be measured. When we can show that the results of learning by television significantly exceed the results of learning in other contexts we will be justified in using such phrases as "the full utilization of the medium for educational purposes". At the moment too little research has been done to either prove or disprove the necessity for such full use. The possibility of "volume education" by television, however, remains as a tremendous force for further experimentation.

Closely related to comparisons of the relative merits of educational films and educational television (and of the limitations of both and the strengths of both) is the comparison of television presentations to the efforts of live teachers in classroom situations. The teacher on television is very restricted by the nature of his medium. The same, of course, is true of the teacher in the classroom. The television teacher may not be able to receive questions from members of his audience. He may not be able to know them well and to make personal remarks to them. He may not be able to adjust his materials to the individual needs of various children, throwing in little asides here and there that are meant for just one or two. He cannot move about freely to see how the work progresses. On the other hand, the classroom teacher lacks the mechanical assistance that can be given over television. Television helps to concentrate the attention of the learners upon one or two objects at the proper time. It flashes names across the screen right over the picture in order to emphasize an important word at an important time. Television teaching gives the teacher the advantage of a program director with more than one camera at his command, who can select from among many possibilities the one that is most effective at the moment. It allows every student a "front seat". Educational television allows closely integrated teamwork between the producer, the performers, and all the other persons behind the scenes. This obviously will be superior to the unaided efforts of many teachers.

The Science Area Was Selected

An early step in planning our research was to decide upon the subject matter that was to be used to enable the techniques of presentation to be investigated. It was decided, finally to do the research in the area of fifth grade science. Science, with its many demonstration problems, would provide many ready opportunities to measure the effectiveness of various demonstration techniques. In science, materials for presentation could be selected that would be new enough to fifth grade children so that the factor of previous knowledge would not be an upsetting one when it came to measuring the learning of the pupils. In addition, there was great need in St. Paul (as there is all over the country) to strengthen the science curriculum at the elementary level. In past generations, elementary teachers have not commonly had broad preparation for teaching science. In the second half of the twentieth century, it became much more important than ever before that <u>all</u> people have at least an elementary background in science.

Educational Television Seeks its Place

Educational television is now entering the realistic phase of its development. In its early days, many people were amazed at what television was able to do as an educational tool, and many others were highly impatient, feeling that television could do much more than anyone had given it opportunity to do. Some people went so far as to believe that through the use of television, the teacher shortage could be overcome, class size could be increased, much school work could be done at home, subjects could be taught in school for which the teachers themselves were not qualified, and master teachers could have their work broadcast over great areas of territory. The conclusions as to the practicability of all this are not, as yet, sufficiently documented.

Almost all educators, however, are agreed that the immediate problems of educational television have to do with improving the techniques that must be developed regardless of how wide the ultimate use of the medium may prove to be. Almost all are agreed that educational television is and will remain a powerful educational tool and that attention should be given to perfecting it. By the same token, effort must not be wasted in striving to develop television practices that are not justified by the results. Perhaps the most helpful of all the numerous impacts of educational television to date is that it has forced educators to re-evaluate their traditional teaching methods and to examine how few really <u>valid</u> insights they have in respect to their present methods.

CHAPTER II

REVIEW OF SOME OF THE LITERATURE

A review of the literature in any area of study must take into account the predispositions of the authors of that literature. With reference to television, some will be colored in favor of the new medium while some will be antagonistic. Research on the subject will vary from careful, excellent work, to work that is poorly conceived and controlled. Educational television is new and careful research in th area is relatively scarce.

There is, however, a considerable amount of research in the area of educational films. Some of this would seem to have applications and implications for educational television. Much would not. Indeed, much of this latter research is not all that might be desired. Our effort will be to make selections from the literature, including the research, in such a way as to suggest possibilities rather than to arrive at established principles.

Some studies in the middle fifties compared one television format or method of presentation with another. Brandon found no significant differences among television methods that involved (1) lectures, (2) interviews, and (3) discussions. Ulrich tested children to see if they retained more information from observing a kinescope recording of a lecture (1) without visual aids, (2) with aids handled by the lecturer, or (3) with aids that were merely flashed on the screen. He found, for his eighth grade pupils, that more information was retained when the recording of the lecturer was supported by poster-type visual aids. He found that it was not important whether the visual aids were handled by the instructor or whether aids were flashed on the screen. He found, too, that after thirty days, the advantage of lecturers employing visual aids over those not employing them tended to disappear. Klapper compared a highly visualized version of a television lesson with a blackboard or "normal" classroom presentation of the same material on television. The students who saw both versions favored the heavy use of production devices in a ratio of 9 to 1. However, Klapper points out that we have not yet learned which kinds of visual aids and audio devices are most effective in aiding the student to learn specific kinds of information. Kumata reported a study in a Chicago suburb which indicated that for classroom use, the television screen should be between four and one-half and six feet from the floor, and staggered rows of chairs should be used. However, these results were in terms of the preferences of the viewers rather than in terms of their learning. Carpenter and Greenhill in their study of closed-circuit television for teaching university courses found no significant differences between the mean scores made by three types of teaching procedures on three tests in a psychology course. The students showed a small preference for the greater use of two cameras so that one could be used for the close viewing of small equipment and demonstrations.

Some research regarding the use of educational films should be briefly mentioned. Nelson and VanderMeer emphasized the importance of both picture and sound in presenting figures and new terminology. They found that the commentary should direct attention to specific elements in the pictorial aspect of the film and should give a brief repetition of important facts at the end. They found that vivid descriptive terms were effective.

Jaspen found that films designed to teach an assembly task were more effective when they had a slow rate of development, showed possible errors, and included four repetitions of the assembly procedure. This research was done with Navy men about materials that would be necessary in emergency situations. It was, of course, desirable to pace the instructional tool so that even the least able would gain some mastery of the material. This is a different learning situation than that in the ordinary classroom.

Barrow and Westley, in TV Effects, propose that clarity of a TV instructional program is a nonlinear function of a number of cues in the message, but that there is a saturation point at which added cues may even diminish comprehensibility. They also propose that there is an ideal or optimum rate at which materials can be presented but they are not explicit as to how this rate may be determined. They believe that interference with learning by television may result from the choice of words and from the grammatical construction. iewers tend to maintain their attention when the material is familiar, meaningful, and related. The authors' suggestions are in the nature of hypotheses and should be accepted as such.

Rimland, McIntyre, and Sterk point out that perceptual motor tasks should be provided in a context as similar as possible to that in which performance will take place. Thus, the camera angle should be the angle from which the task would be approached in actual life. This is called a zero degree camera angle. Roshal also found consistent differences in favor of the zero degree camera angle. He also found that films in which no hands were seen (manipulating the knots that viewers were to learn to tie) were significantly more effective than other films with the hands showing. Perhaps the hands served to clutter up the perceptual field.

Zorbach, experimenting with courses on the college level, found that students were overwhelmingly of the opinion that courses were more interesting and that they learned more from them when they were specifically planned for television rather than when they were made by photographing teachers in regular classroom situations. Mr. Zorbach reports, however, that tests did not show that the students learned more. Zorbach sees so many shortcomings in his own research that his results cannot be accepted as conclusive.

In the Department of the Army technical manual TM11-491, Training by Television, April, 1959, page 17, the point is made that television can present a lesson in shorter time than it can be presented in the classroom because time is saved in the television performance that is used by the instructor in the classroom performance to write on the blackboard and move back and forth while pointing to a chart or equipment. As much as five or ten minutes can be saved from one class period.

The question remains, of course, as to what is "saved". It may be that classroom activity serves to "peg" learning so that it is more readily learned and retained. Again, we need evidence about resultant learning. Saved time is not important unless related to learning.

Along another line, Barrow and Westley demonstrated that a televised newscast was more effective than an equivalent radio newscast when the testing was for immediate recall. The differences, however, though consistent were small.

It can be said without objection that television as a teaching medium is as effective as direct teaching in certain content areas. Several studies showed that it was more effective than conventional instructional methods. These studies, however, are far from conclusive. In many cases, effectiveness is measured by asking the learners which types of classroom experiences they prefer. Many of the studies were based upon television lessons and did not take into account the full potential of television instruction. In truth, also, it must be said that many of the comparisons were with classroom situations where the methods of instruction were typical of the past rather than of the best present practise of pedagogical art.

CHAPTER III

METHOD

The broad problem was how to use television most effectively as a teaching tool. It was necessary, of course, to limit the initial study. This meant that it was necessary to select from among many, the variables which were to be included in the study. Then, the main concern of the study became that of measuring the effectiveness of the selected variables.

Television as a medium of communication and as an educational tool is not a simple thing. A brief review of the research indicates that it is made up of a variety of practices and techniques that may or may not be significant. In other words, television is made up of many factors that contribute to the communication and to the education that may result from its use. Many or all of these factors may be active or interacting at the same time.

Careful attention was given to the selection of the practices and techniques that were to be investigated. The literature and the research were carefully reviewed. Particular attention was paid to the suggestions of those who had used television for educational purposes. It was found that producers of educational films face many of the same problems that are faced in educational television. It became increasingly clear that television is not a one-factor problem. It derives its characteristics from a combination of many factors.

It was finally decided to study three techniques (factors) which shall be referred to as visual continuity, (factor A); visual reinforcement, (factor B); and visual manipulation, (factor C). Each of these techniques has two levels. The first technique (A), on one level is planned and on another unplanned. In planned continuity, the demonstration materials are arranged for camera shots in logical order, in developmental order, or in some other order, while in the unplanned continuity lessons, the demonstration materials are in random order. In the latter instance, the various materials are placed by the teacher without considering the camera or the educational usefulness of the ordered presentation that is made possible by the television camera. The second technique (B) is also on two levels. On the first level, "supering" is made use of, while on the second level, no "supering" is used. Supering is the technique by which a word or phrase or other visual image is superimposed over the picture of the lecturer talking or of the model that is being exhibited. The third technique (C) is visual manipulation. "Television visuals" is a term referring to such things as graphs, charts, and pictures prepared for television presentation. On one level we will have visuals manipulated by the teacher in much the same way that the teacher in the classroom uses such instructional aids. On the other level, visual manipulation will be the result of (1) careful planning in the preparation of the television presentation and (2) judgement on the part of the cameramen and the director.

Treatments

The three techniques (continuity, visual reinforcement, and the manipulation of visuals) can be tested alone or in combination in the same lesson. The total of the techniques used in any television lesson is a treatment.

Table I summarizes the eight treatments with their combinations of techniques. For example, television lesson No. 1 (treatment No. 1) has unplanned visual continuity, no visual reinforcement, and the visuals are manipulated by the teacher. Contrast this with treatment No. 8 where there is planned visual continuity and visual reinforcement in the form of supering and where the manipulation of the visuals is done by TV team effort. Treatments No. 2 through 7 include all other combinations of the three techniques.

TABLE I

EIGHT TREATMENTS*

Treatment	Order of Presentation Systematically Planned or Unplanned	Visual Supering	Visuals - Order of Production Presentation versus Normal Classroom Delivery
1	unplanned	none	Teacher manipulated
2	unplanned	supering	Teacher manipulated
3	unplanned	none	TV team effort
4	unplanned	supering	TV team effort
5	planned	none	Teacher manipulated
6	planned	supering	Teacher manipulated
7	planned	none	TV team effort
8	planned	supering	TV team effort

*The total of all the techniques in any one video recording is a treatment

Selection of Science Areas

An attempt was made to select science content areas which were appropriate to our purposes. For example, presentation of a lesson in electricity adapts itself much better to the use of three dimensional materials than would some other lessons, since electricity can be explained verbally and also can be demonstrated visually through the use of magnets, windings, generators, and other equipment. The presentation of a lesson in weather might make less use of three dimensional materials, depending instead on drawings, films, and explanations by the teacher.

It was finally decided to present lessons in nine areas. These were lessons on plants, mammals, heat, machines, telegraph, wind, ground water, lenses, and longitude and latitude.

Hypotheses

A review of the materials presented so far with particular attention to the techniques and the treatments, provide a basis for the formulation of three hypotheses. These will be stated in the form of null hypotheses. A null hypothesis is a hypothesis that the true difference is zero. The three are as follows:

1. The effectiveness of a science demonstration does not depend on how demonstration materials are arranged for camera shots.
2. The effectiveness of a science demonstration does not depend on how much use is made of visual reinforcement.
3. The effectiveness of a science demonstration does not depend on how visual materials are incorporated.

Lesson Plans and Treatments

Previous experience in St. Paul with the Exploring Science series for fifth grade pupils over KTCA provided the basis for many of our decisions in preparing lesson plans. For example, the KTCA staff and the science teacher had good background to call upon in selecting principles that might best be explained orally or that might best be supplemented by visual presentations. Likewise, they had ideas and they had a basis in experience for feeling that some visual presentations might interfere with the teacher's oral presentation. Suspicions of this kind led to the decision to use a number of content areas in which to test the three hypotheses. They knew that certain factors might distract attention from the message or mask it. They knew, too, that these factors might emphasize and strengthen it.

Subject Matter Material

The production staff regarded the organization of the content material as quite critical. The material used had to be of such a nature that it could successfully fulfill the following criteria: (1) It must be normal in the experience of fifth grade viewers; (2) it must be broad enough to provide ample material for testing; (3) it must be flexible enough to provide for the testing of the various techniques under study; (4) it must be within the reach of the fifth grade pupils without depending on their previous learning; (5) it must be amenable to visual presentation; and (6) it must be presentable in adequate fashion within a twenty-five minute period. Some details concerning the particular content of the topics in each of the nine science areas are given below.

1. Mammals. The main groups of Minnesota mammals (rodents, flesh-eaters, hoofed-mammals, etc.) were identified by a study of their tooth structure. Other characteristics of each group were also surveyed. To a small extent, the range and adaptations of all mammals were discussed.

2. Telegraph. This lesson was devoted to some of the underlying principles of electro-magnetism, to the men who discovered them, and to the invention of the telegraph. This lesson attempted to show that inventions are not the work of one man alone, but depend upon a vast number of previous discoveries.

3. Heat. This lesson gave a molecular definition of heat and taught the three means of heat transmission (conduction, convection, and radiation). Practical applications of heat transmissions were stated.

4. Ground Water. Ground water is important to man, both for survival and pleasure. The television treatments put emphasis on underground structures and their relation to the presence and movement of ground water.

5. Plants. This topic dealt with the great variety of plants, but emphasized seed plants, because of their importance to man. The processes of seed dispersal, osmosis, photo-synthesis, capillarity and function of plant parts were discussed.

6. Longitude and latitude. This lesson tried to give an understanding of longitude and latitude as a system of grids. Angles and degrees were stressed in contrast to miles and other more "constant" measures.

7. Machines. In the main, this lesson was meant to teach the names of the parts of a lever, and to show how the placement of these parts would result in the 'three classes' of levers.

8. Lenses. The properties of various lenses and their effect on light as it passes through them provided the subject matter for this lesson. The practical uses of the various types of lenses were considered.

9. Weather. This lesson examined the prevailing global wind belts of both hemispheres. Stress was placed on some of the possible causes of these belts.

In setting forth the content of the nine science areas in each of which a lesson was to be written, care was taken to stress content only. No procedures were outlined by which to motivate the children or to provide introductory material, and so forth. Everyone who became involved with the selection of content for the lessons regarded it only as a vehicle to use in testing the hypotheses. The final form of each individual lesson depended very largely upon the particular combinations of techniques that were being tested. The extent of the project, then, involved eight treatments for each of nine lessons, with each treatment containing a combination of three techniques. Thus, it was necessary to plan for the actual production of 72 television presentations.

Television Teacher

It was clear that much desirable uniformity (necessary, in fact) was to be obtained by using one teacher in all 72 presentations. The teacher was selected who had previously taught the fifth grade science program designed for the St. Paul Schools and broadcast over KTCA. He was accustomed to working in television, he was an experienced classroom teacher, he held a Master's degree in curriculum and instruction, and he had been active on the elementary science curriculum committee for the St. Paul Public Schools.

Throughout his work with the 72 lessons, the teacher-presenter made every possible attempt to present similar treatments in similar ways in the various content areas. He tried to maintain the same enthusiasm and interest for each lesson.

Studio Personnel

The studio personnel selected for all 72 programs was the same throughout. There was a director, an audio operator, a floor director, and a corps of cameramen. In order to avoid the possibility that the cameramen might increase the quality of production as they became increasingly familiar with the science areas and the camera patterns, a system of rotation was set up for the cameramen so as to secure every possible degree of uniform production quality.

Construction of Visuals

It was necessary to construct visual materials for use in presenting lessons that would test the third hypothesis. In one version of each lesson, all visuals were prepared by the TV teacher himself, and all vocabulary terms were written on the blackboard by the teacher if they were to be emphasized visually at all. Some flannelboard material was prepared by the teacher for his own use. Other visual materials used by the teacher were those readily at hand in the textbooks and other source materials. On the other hand, there were visuals that were prepared for use by the whole television production team, working together. These were carefully prepared. Pictures were drawn by people especially retained for such purposes. In some cases visuals were reduced to gray scale, and every effort was made to enable easy identification by students.

Demonstration Materials

After final plans had been made for all 72 programs, carefull attention was paid to the demonstration materials to be used in each lesson plan. It was necessary to adapt some materials to television use. This was done, for example, by painting parts of the machines in gray scale so that they would reproduce most efficiently on television. The demonstration materials used were secured from reliable equipment firms.

Other Decisions

Numerous other decisions were necessary. It was necessary to decide on the length of the various lessons. There was some attraction in the use of short lessons, but it was finally decided to present lessons twenty-five minutes in length. They were to be presented to fifth grade students in St. Paul and for the period of time necessary for the presentation no other science instruction was to be received in school by the students who were receiving the TV lesson. It was decided to broadcast the twenty-five minute programs to the students so that they would receive the lessons directly from the television screen. It was decided to use every precaution to control the experiment so that the measurement of the variables would not be confused. In an effort to impose certain controls, the seating arrangement for each of the classes participating in the experiment was charted. The placement of the TV sets and the arrangement of the seats were uniform among the classrooms.*

*Television placement and seating arrangements were those recommended in the South Shore Study.

In a further effort to secure uniformity, all TV sets were checked by an independent engineering firm so that there would be little deviation in picture quality from classroom to classroom. TV sets that were found deficient in any way were repaired so as to conform to established standards. Some classrooms were rejected because the audience situation was unsatisfactory.

All programs were scheduled in the morning so as to eliminate any possible advantages resulting from the time of day that a program might be presented. No programs were scheduled in the week prior to the Thanksgiving and Christmas vacations in order to eliminate any lack of attention due to the approaching holidays. One science area was presented in each of nine weeks and one extra week was set aside in case of scheduling difficulties or transmitter failure. The lessons were scheduled between October 26 and February 19 (1959 and 1960).

It was decided to use kinescopes rather than live programs. These kinescopes were produced during the summer of 1959 between June 1 and September 1.

Development of Scripts

The scripts in which the visual continuity was unplanned were prepared by the teacher and the science advisor. First, they went through the lesson plans for each of the nine content areas and developed TV scripts that would best approximate the visual continuity that characterizes classroom teaching as opposed to the highly structured nature of the visual continuity in film and television presentations. Then, under the producer-director, scripts were developed which incorporated versions of the various lessons where the visual continuity was planned according to the standards of TV usage.

The same thing was done in incorporating the other two techniques into the various lesson plans. Thus, the characteristics of the eight scripts that were used in each of the nine content areas can be reviewed briefly. Script No. 1 provided an unplanned use of camera shots, the traditional classroom visuals, and no supering of terminology. Lesson No. 2 included no planning of camera shots or classroom visuals, but supering was developed and used for camera shots with particular emphasis on the vocabulary terms. Lesson No. 3 included no planning of camera shots and no use of supers, but TV visuals were used. Lesson No. 4 had unplanned visual continuity, but it made use of supers and television visuals. Beginning with Lesson No. 5, visual continuity was planned. In addition, Lesson 5 made no use of visual reinforcement and the visuals were of the regular classroom kind. Lesson No. 6 incorporated the planning of camera shots with the supering of terminology, but used classroom visuals. Lesson No. 7 included the planning of camera shots and the use of TV visuals, but made no use of supers. Lesson No. 8 included planned visual continuity, supering, and TV visuals.

Needless to say, the mechanics of combining these techniques in various lessons raised some problems, but they did not prove to be insurmountable. For example, in the lessons where the visual continuity was unplanned, the teacher prepared himself in advance, but the director had no idea how the visual materials were to be handled or how the demonstration materials were to be incorporated. These lessons were done in an ad lib, unrehearsed fashion. Since Lessons No. 2 and No. 4 called for supering, it was necessary for the director to do the best he could to superimpose the word when it was introduced by the teacher without prior warning or consultation. However, in the lessons where visual continuity was planned, considerable time was allowed in the studio before the broadcast for the director and the teacher to plan the placement of demonstration materials, the camera movements, and the exact timing that seemed desirable in incorporating the visual and demonstration materials. Many other examples could be used to illustrate how the various techniques were interwoven.

Controls

As many controls as possible were built into the project. The same teacher taught all the lessons; the same director and the same cameramen worked on the programs; viewing was under uniform conditions as to seating arrangement, distance from the screen, brightness of the screen, and so forth. In the production of visuals, there was uniform letter height, width, and spacing. All demonstration materials were produced by the teacher and used in all the lessons. These included slides, camera cards, mechanical devices, and so forth. All broadcasts lasted twenty-five minutes. The content of each version of a particular lesson was the same. The only variations in lessons were those necessitated by the particular techniques that were being used as variables. All the kinescopes were produced during the summer of 1959.

All kinescopes were reviewed by a panel from the staff of the St. Paul Public Schools (teachers, supervisors, and principals). The members were chosen because of their experience in teaching elementary school children and because they were familiar with the science areas. Supers (where used) were rated for location on screen, timing, and length of exposure. The individual camera shots and sequences of camera shots were checked. Visual materials were rated for clarity, legibility and appropriateness. The teacher's enthusiasm, vocabulary, general speech patterns and rate of oral presentation were rated. Finally, the use of review and the use of repetition were rated. Each aspect was rated on the basis of what the rater considered desirable for learning.

Supering, camera shots and visual materials are related to the treatment in a given kinescope. Distinctions were made among the kinescopes on these bases since the study was a good one to the extent that the only difference among kinescopes for a given science area had to do with three variables. The ratings indicated that the teacher's performance for the eight kinescopes in each science area was very consistent and was generally rated good or very good. This aspect of all kinescopes had to be quite similar since the teacher was not one of the factors under study.

Pupils Available

Available for the study were 67 fifth grade classes in the St. Paul Public Schools. This was all of the fifth grade classes in the school system with some specified exceptions. Excepted were fifth grades that were in combination with sixth grades or fourth grades. Also excepted were fifth grades in schools where the viewing facilities were entirely inadequate. Classes whose teachers had not used television programs prior to the study were not used in the study. Another reason for not using a class was scheduling conflicts. One class had a religious release period that conflicted.

Some pupils were not included in the study because they had been recommended for remedial reading and others were not included because they did not have available a recent IQ test score (not since the first grade). It was felt that the poorest readers would not be able to indicate what they learned since the test was a written test. Since some of the analyses involved the use of intelligence quotients, it was necessary to eliminate a few pupils who did not have IQ test scores.

By use of random numbers, 40 classes were selected to participate. The remaining classes were needed for use in the selection of the items for the lesson tests. The 40 classes were sufficient for five observations of each of the 72 kinescopes. This was possible because with 40 classes and eight treatments, there were five classes per treatment. The classes and the treatments were matched by a system of random numbers.

After the method of lesson preparation was determined for each of the 40 classes, the order of broadcasting was randomly assigned. Two programs were presented each morning, with one week being available for eight broadcasts of each lesson.

Preparing Tests by Which to Measure Outcomes

The mental ability of each child was measured by the Otis Quick Scoring Mental Ability (Beta CM) Test given at the fourth grade level. This mental ability score was used as a covariant.

Objective paper and pencil tests were prepared for each lesson. The items were multiple choice or true-false. Items covered vocabulary, facts, concepts, and the applications of science to everyday living.

In making the tests, approximately one hundred items were prepared for each lesson. In the main, the items were prepared by four graduate students from the University of Minnesota who were hired to prepare them. Each student was considered competent in the area of elementary education. The four met with the producer and the teacher to learn the purpose of the study and the structure of the study. Each graduate student was provided with a lesson plan and two TV scripts for each lesson. One TV script was the presentation of a lesson as it might have been prepared by a classroom teacher entirely unfamiliar with television. The other script was a lesson utilizing maximum coordination between teacher and producer. The one hundred items prepared for each lesson were screened by the science advisor and by the research assistant. They decided that some otherwise appropriate items needed editing. In some cases, new items were constructed to fill gaps where objectives did not seem to be adequately covered.

Item Analysis

Two forms of a test were made for each of the nine lessons. Between twenty-five and thirty items were included in each form. An effort was made to balance the two forms as to the number of items and their difficulties.

These forms were administered to fifth grade classes that were not to be used any place else in the study. First, the children viewed the kinescopes projected on a screen. Eight classes were used for each lesson and one class for each particular kinescope for the lesson. Every second pupil in class took one form of the test. The remainder of the pupils took the second form. About 240 pupils (120 for each form of the test) participated in this testing. The two forms of the test were then scored.

Some answer sheets from some classes were eliminated such as those that showed clearly erratic behavior in the selection of answers and those where five or six items were omitted.

A frequency distribution of scores for the remainder of the tests was made for each form of the test. Pupils with the highest percent of correct answers (25 to 35%) were grouped and called the upper group. Pupils with the lowest percent of correct answers (25 to 35%) were called the lower group. These two groups were always separated by at least four points. Table II shows the number of pupils having usable tests for each form in each science area and the number of pupils in the upper and lower groups for each test.

TABLE II

NUMBER OF FIFTH GRADE PUPILS USED IN THE ITEM ANALYSIS FOR FORMS 1 AND 2 IN EACH SCIENCE AREA

FORM I

Science Area	No. in Upper group	No. in Lower Group	No. of Usable Tests
Plants	25	29	96
Ground Water	27	28	84
Mammals	33	33	108
Machines	34	30	112
Heat	33	32	120
Winds	28	25	105
Telegraph	32	34	96
Lenses	29	29	90
Longitude & Latitude	32	31	118

FORM II

Plants	35	32	136
Ground Water	26	28	95
Mammals	32	31	114
Machines	33	36	108
Heat	32	31	118
Winds	25	27	94
Telegraph	26	28	90
Lenses	29	25	84
Longitude & Latitude	23	25	105

Two indices were obtained for each test item. One was the difficulty of the item. For this, the percent of the upper group answering correctly and the percent of the lower group answering correctly were averaged. The difficulty index is the average percent answering the item correctly.

The other index was the discrimination of the item. For this, the percent of the lower group answering the item correctly was subtracted from the percent of the upper group answering the item correctly. Desirable items should have a positive discrimination index. That is, it should be answered correctly more often by the upper group than by the lower group.

Final Test

All items on the two forms of a test were considered in making up each final test. A summary of the difficulty and the discrimination of each item for each test appears in Table III. There were thirty items for each test except the test on lenses which had twenty-nine.

TABLE III

THE DIFFICULTY AND THE DISCRIMINATION OF THE ITEMS SELECTED FOR EACH TEST

	Plants		Ground Water		Mammals		Machines		Heat	
Percent	Disc.	Diff.	Disc.	Diff.	Disc.	Diff.	Disc.	Diff.	Disc.	Diff.
0-10				1			1			
11-20	3		1	2	3		7		2	
21-30	10	2	5	3	5	2	4		7	
31-40	7	2	9	5	9		7	4	10	
41-50	7	7	5	4	9	5	7	7	8	4
51-60	1	6	4	4	1	8	3	9	1	7
61-70	2	4	4	5	2	9	1	3	2	3
71-80		6	2	4	1	4		6		13
81-90		3		2		1		1		3
91-100						1				

TABLE III (CONT.)

THE DIFFICULTY AND THE DISCRIMINATION OF THE ITEMS SELECTED FOR EACH TEST

Percent	Winds Disc.	Winds Diff.	Telegraph Disc.	Telegraph Diff.	Lenses Disc.	Lenses Diff.	Longitude & Latitude Disc.	Longitude & Latitude Diff.
0-10					1	1	1	
11-20	3				2		3	
21-30	4	1	6		5		6	
31-40	9	4	7		8	1	9	5
41-50	6	7	7	3	5	5	8	5
51-60	2	7	8	7	5	1	2	7
61-70	4	6	2	9	2	7	1	5
71-80	1	5		11	1	7		7
81-90	1					6		1
91-100						1		

The items in each final test covered the content of that lesson. Both true-false and multiple choice items were used. All items that were used discriminated positively though there was a range in difficulty.

The reliability of each test was found through analysis of variance by a formula developed by Hoyt.[1] In this procedure, the reliability coefficient is the proportion of the variance among scores of individuals taking the test that is related to true differences in performance. The error variance is subtracted from the variance among individuals and this difference is divided by the variance among individuals. A summary of the reliability coefficients is given in Table IV.

TABLE IV

RELIABILITY OF THE FINAL TESTS BASED ON HOYT'S ANALYSIS OF VARIANCE

Science Area	Number of Tests Used	Reliability
Plants	92	.58
Ground Water	101	.79
Mammals	112	.71
Machines	109	.66
Heat	104	.65
Winds	94	.74
Telegraph	96	.82
Lenses	90	.63
Longitude and Latitude	92	.64

[1]Cyril J. Hoyt, "Test Reliability Estimated by Analysis of Variance," Psychometrika, Vol. 6 (1941), pp. 153-163

Most of the reliabilities were lower than expected since the items in each test had discriminated well in the item analysis. Some discrimination coefficients were based on small groups and may have been unreliable for that reason. Removing the test items of students with reading problems may have reduced the reliability slightly. Higher reliabilities might have made the statistical analysis slightly more sensitive.

Each test measured sufficiently accurately to differentiate among individuals. This, too, was tested by the analysis of variance.

Subtests

The producer and the teacher-presenter went through the items of each test and chose those items most directly stressed in the lesson. The number of such items found was not the same for each of the lessons. The subtest on plants had eleven items; mammals, nine items; longitude and latitude, eight items; heat, seven items; machines, twelve items; telegraph, thirteen items; wind, twelve items; ground water, eleven items; and lenses, eight items.

Testing

The teacher of the class and the principal of the school knew when the class was to view the lesson on television. Tests and instruction for giving the tests were provided the teacher before the lesson. The class viewed the lesson and returned to the classroom. Within five or ten minutes after viewing, the pupils took the test. They were given twenty minutes to take the thirty-item test. The teacher timed the testing and returned both tests and answer sheets.

The teacher had a set of instructions to guide him in testing his pupils after the lesson. He had not been informed ahead of time about the lesson or the test items so he wasn't tempted to prepare his pupils. This was to reduce the teacher's concern about having his teaching ability reflected in the test scores.

Experimental Design

The problem was to test each of the three hypotheses in a single experiment. Of interest was the effect of each technique by itself and in relation to the other two techniques. The factorial experiment provided an evaluation of single factors and their interactions.

Since all classes were not expected to have the same mean IQ score and since IQ would be related to achievement, the design had to provide for making adjustments for differences among classes in ability. This led to the choice of the factorial design and a covariance analysis.

The analysis of covariance in Table V reveals the skeleton of the experimental design.

TABLE V

ANALYSIS OF COVARIANCE

Source of Variation	Degree of Freedom
Visual Continuity (A)	1
Visual Reinforcement (B)	1
Type of Visual (C)	1
Interaction AB	1
Interaction AC	1
Interaction BC	1
Interaction ABC	1
Error Term	N* - 9
TOTAL	N - 2

*N - number of classes

This made it possible to study each technique as thoroughly as though the whole study were devoted exclusively to that technique. Also, this provided a study of each technique under four different conditions corresponding to the four combinations of the other two techniques.

There were eight treatment combinations in all. The number of classes needed for equal numbers to view each treatment combination had to be a multiple of eight. From the outset, the goal was 40 classes or five for each treatment combination. This was met with several classes to spare.

? It was hoped that each class would view their lesson for each area. Whenever a class missed a lesson, the score for that class was estimated and used in the analysis. The relationship (a regression equation) between IQ and test score was found for the classes viewing the presentation if one or more classes missed the lesson. The mean IQ of the class missing the lesson was inserted into this regression equation and the test score estimated. Proper adjustments were made in the degrees of freedom for the error term.

Hypothesis Testing

The hypothesis that was tested is stated earlier. Hypothesis testing can err in one of two ways (not both ways for a single hypothesis). If a hypothesis is rejected, there is error only if the hypothesis is true. The probability of making this kind of error is fixed in this study. When a hypothesis is accepted, there is error only if the hypothesis is false. The power is the probability that the hypothesis will be rejected if it is false. Ordinarily, power should be considered to be as important as the level of significance. The power was not fixed in advance.

It is essential to know the power of each analysis, especially for those having values nearing significance. If the hypothesis is accepted, though the result is very near being significant and the power is very low, there is high probability of being wrong.

The difficulty in setting the desired probability of each kind of error in any study stems from assigning a value to the risk involved. Risk refers to the risk of either a monetary loss or a loss of information if the wrong decision is made.

What are the risks in this study? If a hypothesis is rejected that is actually true, more time and effort may be put into a TV production than would normally be needed. If a hypothesis is accepted and the hypothesis is actually false, then teaching by television may not be as effective as it should be, and all students viewing TV lessons will suffer. Since this field of research is new and unrefined, and since many later studies will consider the leads of prior studies (hopefully, such as this), the level of significance should not, perhaps, be set too high. Nor should it be assumed that this study is the final word. The loss of information to students may weigh much more heavily than the increased time or the monetary equivalent of the time of a few teachers and producers.

Level of Significance

Thus, the .10 level of significance was considered to be a reasonable level of significance for this study since there is a desire to defer final decisions until more evidence is available. In other words, a difference was declared significant if it was of such a magnitude that sampling errors could account for the difference less than ten times in 100.

In this study, the power indicates the probability of detecting a five percent increase in learning if this increased learning actually results. Dixon and Massey[2] present tables entitled "Power of the Analysis-of-Variance Test." These tables were used to get approximate values of the power for each of our nine analyses. This appears in the analysis chapter.

The probability of rejecting a true hypothesis was ten in 100. In order to be as careful as possible, however, whenever a hypothesis was accepted, the probability of accepting a false hypothesis was determined.

Interpretation of Significant Differences

Whenever a difference was significant, the means for treatments had to be calculated and compared to see which was the highest. If only two means were contained in the comparison, the higher of the two identifies the treatment giving the better results. If there were four means, as in the interactions, a further analysis was made to find which differences among the four were significant.

The following steps were used in determining which treatment or treatments resulted in the highest class means:

1. The mean score for classes within a treatment was adjusted for the mean IQ score for those classes. This gave the achievement for the classes under each treatment after the IQs had been equated by mathematical formula.

2. If A, B, or C proved significant, there were only two adjusted means to consider. The higher mean pointed out the better treatment and the analysis was completed. If AB, AC or BC were significant, there were four adjusted means which led to further analysis. The four adjusted means were compared in pairs using Duncan's multiple range test.[3] In this, the least significant difference is found for each pair of means within the set of means, and the actual difference is compared to the least significant difference. If the actual difference is greater than the least significant difference, the two means differ significantly. Again, the higher mean of a pair where the pair differed significantly pointed out the best treatment.

If A, B, or C was significant, the two levels of the particular treatment did not result in the same amount of learning. One level was better than the other, on the average.

A significant interaction (AB, AC or BC) showed that the result of one treatment was related to the level of the other treatments. As an example, suppose supering, when used with classroom visuals, increased learning. When used with TV visuals, it failed to improve scores. The value of supering, then, depends on whether the classroom visuals or TV visuals are used.

Final interpretation depended on which sources of variation in the analysis of covariance table were significant and the particular means within a source of variation that differed.

The Effect of High and Low Intelligence

At the point where the analysis was practically completed, it seemed advisable to run another analysis to see if the procedures used were more discriminating with the more intelligent children rather than the less intelligent children, or vice versa. For example, it was thought that the use of visuals and supering and planned continuity might be more effective either with slow children or with intelligent children (or again vice versa) because such techniques might prove, for example, to be distracting or to be valuable reinforcements. In other words, it was thought that the value of a particular technique might be related to the ability of the group that was being taught or that it might be related to a subgroup within the group. In order to get at this question, the tests for two science areas were analyzed. The test on heat, one of the easier tests, and the tests on machines, one of the more difficult tests, were used.

---

[2] W.J. Dixon, and F.J. Massey, Jr., Introduction to Statistical Analysis. New York: McGraw-Hill Book Company, Inc., 1957

[3] David B. Duncan, "Multiple Range and Multiple F Tests," Biometrics, Vol. 11, No. 1, 1955, pp. 1-42.

Three ability groups were formed: IQ less than 100, IQ greater than 110 and those in between 99 and 111. Again the analysis of variance was used. None of the techniques, singly or in pairs, were interacting with IQ. The groups with the highest IQ did best, and the group with the lowest IQ the poorest. This was expected.

It is possible that the first analysis (covariance) wouldn't show significance for different techniques even though something was there. It seemed possible that one level of a technique may have helped the brighter pupils and the other level may have helped one of the other ability groups. The analysis of the two tests showed that this wasn't true for heat or machines, and in the absence of contrary evidence, the assumption was that it wasn't true for the other science areas.

CHAPTER IV

ANALYSIS

The experimental design provided for an analysis in each science area. First, the hypotheses were tested by the analysis of variance. If this procedure indicated that one of the sources of variation was significant, second and third steps were pursued. In the second step, the test scores were adjusted for the IQ scores. If the class had an IQ above the mean IQ for all 40 classes, the test score was adjusted downward. If the opposite were true, the test score was adjusted upward. The third step was to compare the adjusted means. With only two means for a significant source of variation, the technique having the higher mean was the better. With four adjusted means in the comparison, (i.e., interactions were significant), differences among the four were tested for significance using Duncan's multiple range test.

The tests, in their entirety, were handled this way first. This was followed by the analysis of the subtests; each was composed of a set of selected items.

After completing the analysis of variance and the study of the significant sources of variation, the power of the test was determined for the main effects that were not significant. This revealed the probability of detecting a certain difference in the treatment effects, if such differences actually occurred. The probability of detecting a 5% increase in learning was selected as was mentioned earlier.

Tables are given below under each area which indicate sources of variation, degrees of freedom, mean squares, significance or lack of significance of the differences, and significance of the subtest. Under "source of variation" in each analysis of variance table, A refers to visual continuity at two levels--planned or unplanned. B refers to visual reinforcement at two levels--supering or no supering. C refers to the types of visuals at two levels--classroom visuals or TV visuals, as described in Chapter III. Two or more letters in combination refer to interaction terms.

Plants

Thirty-seven classes viewed the lesson on plants. Test scores for the three classes that missed the lesson were estimated (a regression estimate), using the relationship between IQ and test score for the presentation they were to have viewed. The analysis of variance test of the hypothesis appears in Table VI.

TABLE VI

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON PLANTS (30 Item Test and 11 Item Subtest)

Source of Variation	Degrees of Freedom	Whole Test: Mean Square	Whole Test: Differences	Subtest: Differences
A	1	0.92	NS	S
B	1	1.61	NS	S
C	1	20.13	S	S
AB	1	0.01	NS	NS
AC	1	0.34	NS	NS
BC	1	0.03	NS	NS
ABC	1	0.03	NS	NS
error	28	1.10		
Total	35			

The mean test score, after adjusting for class IQ, for all classes viewing kinescopes having classroom visuals was 18.2. The adjusted mean for classes viewing kinescopes having TV visuals was 16.8. Since those classes viewing the kinescopes with classroom visuals had the higher mean, classroom visuals were better than TV visuals in this instance. (For contrast, see longitude and latitude below.)

The remainder of the hypotheses under test were accepted. The power of the analysis for the accepted hypotheses was .82. This indicates that we can be quite confident that there wasn't a 5% or larger treatment difference where the hypotheses were accepted.

The subtest on plants contained 11 items. The analysis of the 11 selected items is given in the table. All three main effects were significant with the selected items in the subtest. The mean test scores, after adjusting for class IQ, for treatment means were 5.7 for planned visual continuity and 5.5 for unplanned visual continuity; 5.7 for visual reinforcement and 5.5 for no visual reinforcement; 5.8 for classroom visuals and 5.4 for TV visuals.

Thus, the adjusted mean for classroom visuals was better than the mean for TV visuals, as was true for the whole test. In addition, visual reinforcement was better than no visual reinforcement and planned visual continuity was better than unplanned visual continuity.

Mammals

All 40 classes viewed the lesson on mammals. The analysis of variance test of the hypotheses appears in Table VII.

TABLE VII

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON MAMMALS
(30 Item Test and 9 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degrees of Freedom	Mean Square	Differences	Differences
A	1	0.01	NS	NS
B	1	3.16	NS	S
C	1	0.38	NS	NS
AB	1	0.45	NS	NS
AC	1	0.25	NS	NS
BC	1	15.88	S	S
ABC	1	3.46	NS	NS
Error	31	1.68		
Total	38			

The table reveals that none of the factors by themselves were significant but the BC interaction was significant. The adjusted test scores for the BC interaction were:

BC Combination	Mean Test Score
Visual reinforcement, classroom visuals	18.4
Visual reinforcement, TV visuals	17.4
No visual reinforcement, classroom visuals	16.5
No visual reinforcement, TV visuals	18.0

Of these means, 18.4 was significantly higher than 16.5 or 17.4 which suggests the conclusion that visual reinforcement with classroom visuals is better than visual reinforcement with TV visuals and also better than no visual reinforcement with classroom visuals. Also, 18.0 was significantly higher than 16.5. This suggests the conclusion that TV visuals with no visual reinforcement was better than classroom visuals with no visual reinforcement.

The power of the test for hypotheses accepted was .68. In this case, accepting the hypothesis doesn't offer much evidence that the hypothesis is true. We can say that the sample doesn't refute certain hypotheses.

The analysis of the subtest shows that Factor B and the BC interaction were significant. The mean test score, after adjusting for class IQ, for treatment means were 5.0 for visual reinforcement and 4.7 for no visual reinforcement; 5.1 for visual reinforcement, classroom visuals and 4.8 for visual reinforcement, TV visuals; 4.6 for no visual reinforcement, classroom visuals and 4.9 for no visual reinforcement, TV visuals. This result was slightly different from the results of the 30-item test. Visual reinforcement was better than no visual reinforcement. More specifically, since the only significant difference among the four interaction means was that 5.1 was larger than 4.6, visual reinforcement exceeded no visual reinforcement when classroom visuals were used.

Longitude and Latitude

All 40 classes viewed the lesson on longitude and latitude. Table VIII contains the analysis of the hypotheses for the 30 item test and the 8 item subtest.

TABLE VIII

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON LONGITUDE AND LATITUDE (30 Item Test and 8 Item Subtest)

Source of Variation	Degrees of Freedom	Whole Test Mean Square	Whole Test Differences	Subtest Differences
A	1	2.16	NS	NS
B	1	0.85	NS	NS
C	1	13.00	S	S
AB	1	2.69	NS	NS
AC	1	0.07	NS	NS
BC	1	8.32	S	S
ABC	1	1.83	NS	NS
Error	31	1.56		
Total	38			

The table shows that factor C and the BC interaction were both significant. The mean test scores, after adjusting for class IQ, for treatment means included in the significant sources of variation were 15.9 for classroom visuals and 17.0 for TV visuals; 16.1 for visual reinforcement, classroom visuals and 16.4 for visual reinforcement, TV visuals; 15.5 for no visual reinforcement, classroom visuals and 17.6 for no visual reinforcement, TV visuals.

Classes viewing kinescopes with TV visuals had a higher mean than those viewing kinescopes with classroom visuals. Comparisons among the four means in the BC interaction shows that 17.6 is significantly higher than 15.5, 16.1, and 16.4. Thus, TV visuals were better than classroom visuals and TV visuals were at their best with no visual reinforcement.

The power of the test for the hypotheses that were accepted was .70. Again this wasn't conclusive evidence that these hypotheses were true, but they were not refuted by the results of this lesson.

Analysis of the eight selected items revealed the identical pattern of significant and non-significant differences.

Heat

Thirty-eight classes viewed the lesson on heat. The analysis of the hypotheses for the 30 item test appears in Table IX.

TABLE IX

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON HEAT (30 Item Test and 7 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degrees of Freedom	Mean Square	Difference	Difference
A	1	0.01	NS	NS
B	1	8.62	S	NS
C	1	0.32	NS	NS
AB	1	0.[illegible]	NS	NS
AC	1	3.50	NS	NS
BC	1	3.77	NS	NS
ABC	1	4.95	NS	NS
Error	29	1.73		
Total	36			

The mean test score, after adjustment, for those classes viewing the kinescopes that didn't have supering was 18.2 and with supering, the mean was 19.1. Supering resulted in the higher of the two adjusted means and, for this lesson, was better than no supering.

All other hypotheses under test were accepted for this lesson. The power of the test for the hypotheses that were accepted was .65. These samples do not refute these hypotheses, though the results don't justify confidence that the hypotheses are true.

Analysis of the 7 selected items revealed no significant differences among the treatments. This differed from the results of the whole test where the B factor was significant.

Machines

Thirty-eight classes viewed the lesson on Machines. The analysis of the hypotheses for the 30 item test appears in Table X.

TABLE X

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON MACHINES (30 Item Test and 8 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degrees of Freedom	Mean Square	Difference	Difference
A	1	0.55	NS	NS
B	1	0.67	NS	NS
C	1	0.08	NS	NS
AB	1	2.23	NS	NS
AC	1	1.35	NS	NS
BC	1	2.24	NS	S
ABC	1	0.22	NS	NS
Error	29	1.29		
Total	36			

None of the hypotheses were rejected. The power of the test for all hypotheses, since all were accepted, was .77. This serves as some evidence, though not conclusive, that we haven't missed rejecting false hypotheses.

On the subtests, the BC interaction was significant.

The mean test score, after adjusting for class IQ, for treatments included in this interaction were 6.2 for visual reinforcement, classroom visuals, and 5.6 for visual reinforcement, TV visuals; 5.9 for no visual reinforcement, classroom visuals, and 6.1 for no visual reinforcement, TV visuals.

Both 6.2 and 6.1 are significantly larger than 5.6. That is both visual reinforcement with classroom visuals and no visual reinforcement with TV visuals were better than visual reinforcement with TV visuals.

Telegraph

Thirty-eight classes viewed the telegraph lesson. The analysis of the hypotheses for the 30 item test appears in Table XI.

TABLE XI

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON TELEGRAPH (30 Item Test and 13 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degrees of Freedom	Mean Square	Difference	Difference
A	1	1.45	NS	NS
B	1	0.69	NS	NS
C	1	0.06	NS	NS
AB	1	3.67	NS	NS
AC	1	24,87	S	S
BC	1	2.64	NS	NS
ABC	1	0.35	NS	NS
Error	29	3.32		
Total	36			

The AC interaction was significant. After adjusting for class IQ, the mean test score for treatments included in this interaction were 18.6 for planned visual continuity, classroom visuals, and 17.1 for planned visual continuity, TV visuals; 17.4 for unplanned visual continuity, classroom visuals, and 19.1 for unplanned visual continuity, TV visuals.

Here the two larger means, 19.1 and 18.6, did not differ significantly. Neither did the two smaller means, 17.4 and 17.1. However, the two larger means differed significantly from the two smaller means. In other words, (a) planned visual continuity with classroom visuals does not differ from unplanned visual continuity with TV visuals; and (b) unplanned visual continuity with classroom visuals does not differ significantly from planned visual continuity with TV visuals. Treatments under (a) were significantly better than those under (b).

The power of this test for the hypotheses that are accepted is .44. Over half the time this analysis would not reject hypotheses that were false. In this case, accepting the hypothesis does not indicate that it is true.

Analysis of the 13 selected items revealed the same pattern as found for the items of the whole test.

Winds

Thirty-eight classes viewed the lesson on winds. The analysis of the hypothesis for the 30 item test appears in Table XII.

TABLE XII

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON WINDS (30 Item Test and 12 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degree of Freedom	Mean Square	Difference	Difference
A	1	0.57	NS	NS
B	1	3.74	NS	NS
C	1	0.48	NS	NS
AB	1	20.12	S	S
AC	1	5.13	NS	NS
BC	1	0.31	NS	NS
ABC	1	1.41	NS	NS
Error	29	1.97		
Total	36			

The AB interaction was significant. After adjusting for class IQ, the mean test score for treatments included in this interaction were 18.5 for planned continuity, no visual reinforcement, and 17.6 for planned continuity, visual reinforcement; 16.9 for unplanned continuity, no visual reinforcement and 18.8 unplanned continuity, visual reinforcement.

Here 18.8 was significantly higher than 16.9 and 17.6. Also 18.5 was significantly higher than 16.9. So, visual reinforcement with unplanned continuity was better than either visual reinforcement with planned continuity or no visual reinforcement with unplanned continuity. In addition, planned visual continuity with no visual reinforcement was better than unplanned visual continuity with no visual reinforcement.

The power of the test with respect to the hypotheses that were accepted was .61. Again this doesn't indicate strong evidence that the accepted hypotheses are true.

Analysis of the 12 selected items showed the same results as were found in the analysis of all 30 items.

Ground Water

Thirty-eight classes viewed the lesson on ground water. The analysis of the hypotheses for the 30 item test appear in Table XIII.

TABLE XIII

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON GROUND WATER (30 Item Test and 11 Item Subtest)

		Whole Test		Subtest
Source of Variation	Degrees of Freedom	Mean Square	Difference	Difference
A	1	0.58	NS	NS
B	1	0.05	NS	NS
C	1	0.71	NS	NS
AB	1	19.70	S	S
AC	1	1.58	NS	NS
BC	1	0.12	NS	NS
ABC	1	0.93	NS	NS
Error	29	1.59		
Total	36			

The AB interaction was significant. After adjusting for class IQ, the mean test score for treatments included in this interaction were

19.1 planned visual continuity, no visual reinforcement, and 17.6 for planned visual continuity, visual reinforcement; 17.9 for unplanned visual continuity, no visual reinforcement, and 19.3 for unplanned visual continuity, visual reinforcement.

Here 19.3 and 19.1 did not differ significantly. Also 17.9 and 17.6 did not differ significantly. However, 19.3 and 19.1 were significantly higher than 17.9 and 17.6.

Therefore, (a) visual reinforcement with unplanned visual continuity did not differ from no visual reinforcement with planned continuity, and (b) planned visual continuity with visual reinforcement did not differ from unplanned visual continuity with no visual reinforcement. Treatments under (a) were significantly better than those under (b)

The power of the test for those hypotheses that were accepted was .88. This offers good evidence that false hypotheses were not missed.

Analysis of the 11 selected items lead to the same conclusions as found with the 30 item test.

Lenses

All forty classes viewed the lesson on Lenses. The analysis of the hypotheses for the 30 item test appears in Table XIV.

TABLE XIV

ANALYSIS OF DIFFERENCES AMONG TREATMENT MEANS ADJUSTED FOR IQ DIFFERENCES FOR THE LESSON ON LENSES (30 Item Test and 8 Item Subtest)

Source of Variation	Degrees of Freedom	Whole Test: Mean Square	Whole Test: Difference	Subtest: Difference
A	1	3.31	NS	S
B	1	6.79	S	NS
C	1	0.04	NS	NS
AB	1	3.10	NS	NS
AC	1	5.50	NS	S
BC	1	0.51	NS	NS
ABC	1	2.17	NS	NS
Error	31	1.95		
Total	38			

The B factor was significant. The mean scores for each of the treatments, visual reinforcement and no visual reinforcement, were adjusted for IQ differences. After adjustment, the mean for visual reinforcement was 17.4 and for no visual reinforcement was 16.5. Therefore, visual reinforcement was better than no visual reinforcement.

The power of the test for those hypotheses that were accepted was .62. Again this offered little evidence that the accepted hypotheses were true. The sample did not refute the hypotheses.

On the subtest, factor A and the interaction AC were significant. After adjusting for class IQ, the mean test score for treatments included in these sources of variation were 4.7 for planned visual continuity and 4.2 for unplanned visual continuity; 4.9 for planned visual continuity, classroom visuals and 4.4 for planned visual continuity, TV visuals; 4.2 for unplanned visual continuity, classroom visuals, and 4.4 for unplanned visual continuity, TV visuals.

The mean for planned visual continuity was 4.7 and for unplanned continuity was 4.2. Therefore, planned continuity was better than unplanned continuity. For the means in the AC interaction, the only significant difference was between 4.9 and 4.2. Planned visual continuity is better if it is used with classroom visuals rather than TV visuals.

Summary

At least one of the hypotheses was rejected for eight of the science areas using the whole test. All hypotheses were accepted for machines. At least one of the hypotheses was rejected for eight of the science areas using the subtests. Whole tests and subtest analyses gave similar, though not identical, results.

All but two of the 22 differences reported as significant would have been detected as significant had the .05 level been used instead of the .10 level. (See Table XV).

The majority of the significant differences were caused by the interaction of two factors. In every case, the best treatment is a combination of one technique using the team approach (planned continuity or supering or TV visuals) and a second technique using the teacher-only approach (unplanned camera continuity or no supering or classroom visuals).

The remainder of the significant differences were caused by individual factors. Some of these duplicated information found in the interaction such as with longitude and latitude lesson. Here the single factor C shows that TV visuals were better than classroom visuals while the BC interaction shows, in addition, that TV visuals were at their best with no visual reinforcement.

The only negative finding was in the lesson on plants. Here, classroom visuals were superior to TV visuals. In all others, a technique emphasizing the team approach singly or in combination with another technique gave the best results.

While each hypothesis was rejected at least several times, no pattern developed throughout the science areas. Supering showed to advantage in several science areas but not throughout all. The same is true of the other two techniques.

The power of the analyses ranged from .44 for the telegraph lesson to .88 for ground water. In fact, only ground water had a small probability of error where the hypothesis was accepted.

TABLE XV

SUMMARY OF DISCOVERED SIGNIFICANCES
(Tests and Subtests)

TESTS	Visual Continuity Factor A	Visual Reinforcement Factor B	Teacher vs. TV Visuals Factor C	Combination of Factors AB	AC	BC	ABC
Plants			S				
Mammals						S	
Longitude & Latitude			S			S	
Heat		S					
Machines							
Telegraph					S		
Winds				S			
Ground Water				S			
Lenses		S					

SUBTESTS	Visual Continuity Factor A	Visual Reinforcement Factor B	Teacher Vs. TV Visuals Factor C	Combination of Factors AB	AC	BC	ABC
Plants	S	S	S				
Mammals		S				S	
Longitude & Latitude			S			S	
Heat							
Machines						S	
Telegraph					S		
Winds				S			
Ground Water				S			
Lenses	S				S		

CHAPTER V

DISCUSSION AND CONCLUSIONS

First to review the specific findings. The present study revealed that at least one statistical difference favorable to television production techniques based on whole test scores was found for all of the various science areas with the exception of the area, "Machines". Likewise, at least one statistical difference favorable to television production techniques based on sub-test scores were found in all of the science areas with the exception of the area, "Heat". On the whole, patterns of results based on calculations from whole test results generally coincided with sub-test results.

The majority of the differences favorable to television production - team-inspired-teaching techniques resulted from the interaction of two of the factors examined in the present study. In most instances, the highest student achievement resulted from a combination of one technique using the television team production efforts (planned continuity or supering or television visuals) and a second technique using only the teacher's efforts (unplanned camera continuity or no supering or classroom visuals). In other words, the television devices appeared to facilitate the effects of the teacher's efforts.

The remainder of the significant differences in support of the null hypotheses may have resulted from artifacts of course content. For example, some of the technique interaction situations involved duplication of content (e.g., longitude and latitude lesson).

The only finding in the study that to a significant degree favored teacher-developed techniques by themselves was based on the lesson dealing with plants. Here, classroom visuals were apparently more effective in terms of student achievement than television visuals. In all other instances, however, where the null hypothesis was rejected, a technique resulting from the television team approach singly or in combination with another technique precipitated the highest achievement among the students.

To recapitulate, with a somewhat different perspective, the general hypotheses and methodology in the present study and speculate beyond the specific findings, there appears to be an opportunity because of the many provocative findings in the present study, to look in depth at some of the operational variables which are seldom, if ever, precisely examined in the studies which report "no significant differences" when traditional classroom vs. television instruction are compared.

In most of the studies which report "no significant differences," as pointed out in the review of the literature, television presentation is nothing more or less than placing a professor in front of a camera. Traditional classroom presentation is usually the same teacher operating in about the same way without the television camera. The present study affords an opportunity to look more carefully at presentation techniques which are unique to the television medium. This aspect of the present study allows us to challenge research reporting "no significant differences" on the basis of the possibility that such studies by "over-control" do not involve TV as a unique medium with unique potential for developing particularly effective visual instructional devices. For example, it may be inferred from our present findings, that if any instructor improves and systematizes his order of presentation of demonstration materials and is able to use something at least like a superimposition technique, and is able to use maximally a variety of visuals the student achievement will probably increase whether the presentation is by television or is not. However, if television becomes an important stimulator to this kind of re-thinking of course presentation methodology, and in addition has peculiar characteristics which allow the maximum utilization of such devices as superimposition and various visuals designed to bring in sharp focus course content ideas, so much the better.

At an even more practical level, as follows, are some broad areas of questions concerning television instruction often voiced by classroom teachers which in the writer's opinion cannot be dealt with on the basis of studies reporting "no significant difference" findings, but can by implication be dealt with from the findings in the present study:

1. What can a professional television production staff do in terms of the organization of demonstration materials to facilitate learning of students that a teacher himself is unable to do?

Through this question there is implicitly expressed great reservation on the part of the teacher as to whether or not systematic organizations of demonstration by television materials specialists are very important. The results of the present study suggest that in many instances the stimulus toward organization of demonstration materials which results from a television production team, results in more effective achievement, as measured by typical achievement tests, particularly when this is done in combination with other techniques unique to television.

2. Why not allow the camera to simply record the classroom teacher in action, since the very spontaneity of this situation will somehow capture the maximum effort of the given teacher without the use of specialized television "gimmicks" such as superimposition?

Many television instructors in effect essentially call on television production staffs to do just this. Again, the present study gives us a lead which suggests that contrived television-centered techniques such as superimposition _in themselves_ in many instances do _add materially_ to student learning, and particularly so _in conjunction_ with other relevant production techniques.

3. Aren't intuitive, unstructured presentations of visual aids based on the professional competence of the teacher somehow more academically sound and more effective than the use of visual aids tailored for or made possible through the television medium?

Again, the present study suggests that this belief, at least as far as student achievement is concerned, is probably unfounded. As used in the present study, visual aids developed in cooperation with television production staffs contributed as much as and in many instances more to student learning than did visuals prepared by the teacher alone.

Returning to the specific findings in the present study, it is very clear that systematic efforts to study the relative impact on learning of three presentation variables unique to television in their utilization, order of presentation of demonstration materials, superimposition, and an array of additional visual aids, cannot be accomplished with perfect precision. It is therefore not surprising that attempts to isolate each of these variables, and demonstrate their relative strength in a total teaching situation failed to yield clear-cut results favoring their use. But the fact that every one of these factors, either alone or in some combination did yield results under at least some conditions indicating statistically significant greater achievement is very encouraging. It may be recalled that in our review of the literature it was pointed out that some of the researchers speculated that over-saturation in the use of these techniques may soon undermine their effectiveness. It is altogether possible that in the present study where differences in effectiveness between teacher- and TV-centered techniques were not significant might very well have resulted from the fact that the "point of no return" with various forms of visual manipulation in the present study was often present. If, for example, all of these special presentation methods were pitted against the virtual absence of their use, as is the case in much traditional classroom instruction, one might hypothesize that their contribution to higher student achievement would be even more dramatically apparent.

Looking now at one area of the results particularly supportive of television team preparation, combining methods, suggests an interesting psychological analogy. The television presentation on the screen at any given moment may be conceived of in the sense of a Gestalt psychology figure-ground relationship. If we can conceptualize this situation as the concept or idea to be learned as the figure and any supporting techniques of presentation as the ground or background, we have another framework for the analysis of our results. It might be hypothesized, for example, that in the case of the organization of demonstration materials, where the only significant differences were found in the case of two sub-test results, that organization *per se*, since it is not in evidence on the screen at any given moment, is a more subtle cognitive reinforcing agent than, say, the use of superimposition or visual aids which are usually on the screen at any given moment. However, when organization of demonstration materials is systematized and used in combination with the other techniques the latent effects of the more systematic organization maximizes the opportunity for the principal idea being taught to be truly reinforced. In other words, organization in itself only tends to be effective when its accumulated effects can serve as a background of a "figure", in this instance a visually reinforced idea being taught.

In conclusion, then, the following observations pertaining to the present study might be made:

1. Using the content of a 5th grade science class as a model, a television production team's assistance in tailoring for television the order of presentation of visual materials, superimposition, and visual aids resulted in no less student achievement than the efforts of the teacher himself.

2. Significant differences in student achievement on whole or subtests favoring television production team help over the sole efforts of a teacher, however, were found in at least one instance in each of nine science areas included in the course content.

3. Combining television production team assistance with the use of one technique with either of the other techniques of this type, or with teacher techniques developed by the teacher himself, appeared to result in the most frequent evidence of significantly higher student achievement.

4. A significant possibility of a break-through in the research methodology which has heretofore often led to "no significant difference" findings in instructional television investigations may have been demonstrated, in the sense that course demonstration devices were successfully isolated and manipulated with relative effects on student achievement determined.

BIBLIOGRAPHY

Barrow, L.C., and B.H. Westley, TV Effects, Research Bulletin No. 9. Madison, Wisconsin: Wisconsin University Television Laboratory, 1958.

Brandon, James R., The Relative Effectiveness of Lecture, Interview, and Discussion Methods of Presenting Factual Information on Television, unpublished PhD dissertation, University of Wisconsin, 1955.

Carpenter, C.R. and L.P. Greenhill, An Investigation of Closed Circuit Television for Teaching University Courses, Instructional Television Research Project No.1. University Park, Pa.: Pennsylvania State College,1955.

Dixon, W.J. and F.J. Massey, Jr., Introduction to Statistical Analysis. New York: McGraw-Hill Book Company, Inc., 1957.

Duncan, David B., "Multiple Range and Multiple F Tests," Biometrics, Vol. 11, No. 1, 1955.

Hoyt, Cyril J., "Test Reliability Estimated by Analysis of Variance," Psychometrika, Vol. 6, 1941.

Jaspen, Nathan, Effects on Training of Experimental Film Variables, Study I; Verbalization, Rate of Development, Nomenclature, Errors, How-It-Works," Repetition, Technical Report SDC 269-7-17. University Park, Pa.: Instructional Film Research Program, Pennsylvania State College, 1950.

Johnson, P.O. and R. .B. Jackson, Modern Statistical Methods: Descriptive and Inductive. Chicago: Rand McNally & Company, 1959.

Klapper, Hope Lunin, "A Prospective of Televised Instruction," in College Teaching by Television. Washington, D.C.: American Council on Education, 1958.

Kumata, Hideya, An Inventory of Instructional Television Research. Ann Arbor, Michigan: University of Michigan, 1956.

Nelson, H.E. and A.W. Vander Meer, The Relative Effectiveness of Differing Commentaries in an Animated Film on Elementary Meteorology, Technical Report SDC 269-7-43. University Park, Pa.: Instructional Film Research Program, Pennsylvania State College, 1955.

Rimland, Bernard, Charles J. McIntyre, and H. Dennis Sterk, Effectiveness of Several Methods of Repetition of Films, Technical Report SDC 269-7-45. University Park, Pa.: Instructional Film Research Program, Pennsylvania State College, 1955.

Roshal, Sol M., Effects of Learner Representation in Film-Mediated Perceptual Motor Learning, Technical Report SDC 269-7-5. University Park, Pa.: Instructional Film Research Program, Pennsylvania State College, 1949.

Ulrich, John H., An Experimental Study of the Acquisition of Information from Three Types of Recorded Television Presentations, NAEB Fact Sheet, Series II, Production Principles. National Association of Educational Broadcasters, Urbana, Illinois, March 13, 1957.

Zorbaugh, Harvey as a panelist in College Teaching By Television. Washington, D.C.: American Council on Education, 1958.

APPENDIX

This appendix is meant to provide background to the study. It contains (A) excerpts from the application for HEW funds, (B) an outline of the content of one of the nine lessons (each one was presented in eight different ways), (C) a copy of one of the 72 scripts that were prepared for the kinescopes, and (D) a copy of the test for one of the nine lessons.

APPENDIX A

EXCERPTS FROM APPLICATION TO H.E.W.

APPLICATION TO THE COMMISSIONER OF EDUCATION, U.S. OFFICE OF EDUCATION DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE, FOR FUNDS TO SUPPORT A PROJECT UNDER THE PROVISIONS OF TITLE VII OF THE NATIONAL DEFENSE EDUCATION ACT OF 1958. (P.L. 85-864)

<u>Submitted by:</u> Twin City Area Educational Television Corporation (KTCA-TV)
Loring M. Staples, President
Dr. John C. Schwarzwalder, General Manager

<u>Principal Investigator:</u> Nolan C. Kearney, Assistant Superintendent of the St. Paul Public Schools for Research and Curriculum

<u>Title:</u> An Investigation of the Relative Effectiveness of Certain Specific TV Techniques on Learning

<u>Problem:</u> The use of television as an education tool has grown tremendously in recent years but few educators would claim that they have been able to use the medium to the limit of its potentialities. One of the areas that is almost completely unexplored is how television techniques involved in the presentation of a teaching program, affect learning.

Teachers, producers and directors of instructional programs <u>do not know</u> how to present material most effectively on television.

There are many subject areas in which an investigation of appropriate techniques might be studied. None, however, is more important in the present world context than science. Science in the elementary school is an especially critical area.

Obviously the discovery of these most effective ways is dependent upon a clear understanding of the relative effectiveness of various TV techniques. The value of this proposed research is that it would determine the effectiveness of specific techniques.

Objectives: Since the principal objective of this research is to determine the effectiveness of certain television techniques which would serve as a guide to teachers, producers and directors of Educational Television programs, we need first to state our objectives in general hypotheses that, if proven or disproved, would serve as general guides to persons involved in such programs.

Hypothesis I: It makes no difference in a science demonstration program whether visual continuity is planned for or not.

Hypothesis II: It makes no difference whether visual reinforcement is used in a science demonstration or not.

Hypothesis III: It makes no difference whether visual material is handled by the teacher or the TV team in a science demonstration program.

These hypotheses and all techniques associated with them will be paired with nine science content areas.

Procedure: The experimental design is a 2x2x2 factorial with 5 replications. All programs will be kinescope recordings and will be televised to classes of 5th graders.

Controls: Same teacher, director, cameramen for all programs. Viewing will be under uniform conditions. Production of visuals will be under uniform conditions. All programs will last 25 minutes. The content for each program for a particular lesson will be the same. The only variations in programs will be those necessitated by the particular technique used. To assure uniformity of presentation, all programs will be kinescoped during the summer of 1959. Each kinescope will be reviewed by a panel of experts to see if the specified techniques have been demonstrated.

The population will be the 67 fifth grade classes in the St. Paul Public Schools. Of those eligible for the study, 40 will be selected, by use of random numbers to participate. The remaining classes will be needed in the final selection of items for the lesson tests.

A special test will be designated for each lesson to test learning.

Time Schedule: July 1, 1959 -- October 31, 1960

Budget:	Total cost of Project	$51,019.00
	Total Federal Funds requested	$51,019.00

APPENDIX B

Content for Kinescopes on

HOW DOES A TELEGRAPH WORK?

A telegraph is a device which can send and receive code messages. It accomplishes this through the use of magnetism produced by electricity.

The principle of producing magnetism by electricity was first discovered by Hans C. Oersted, a Danish scientist. Oersted found that whenever electricity flows through a wire, magnetism is present. Magnetism is a force which among other things can be used to attract iron. It is this effect of magnetism which is used to operate a telegraph.

Demonstration of Oersted's discovery

1. It will be shown that when electricity flows through a wire a compass needle is attracted.

2. It will also be shown that the wire attracts iron filings but will not attract various other materials.

Additional things we should know about magnets produced by electricity.

1. Magnetic fields

   The space around a current carrying wire in which magnetism is present is called a "field."

2. We can control the field.

   a. Make it stronger by coiling the wire.

   b. Make it stronger by placing an iron rod (core) in the coil.

   c. Make it stronger by sending more electricity through the wire or coil.

   d. We can turn the field on or off at will by switching the current on or off.

We will show now how these principles of electro-magnetism are used in a telegraph. The application of electro-magnetism was first used by S.B. Morse in his invention of the electric telegraph. In his telegraph an electro-magnet was used to attract a piece of iron (armature) in such a way as to produce audible "clicks." These "clicks" would be under the control of a distant operator. The combination of "clicks" were used to send messages according to a code devised by Morse.

We will now demonstrate the steps which Morse may have tried in the development of his telegraph.

1. Strip of iron (armature) and single wire.

   Armature is not drawn down

2. Strip of iron (armature) and coil of wire.

   Armature may move but not sufficiently to produce "click."

3. Strip of iron (armature) and coil of several layers which has had a soft iron core inserted into it.

   Field is now strong enough to attract armature which produces a "click." More batteries will send additional current through the coil and produce a louder click.

4. We substitute a spring type of switch for the knife switch. This makes possible a more rapid interruption of the flow of electricity. This spring switch is called the sending key.

In essence then, the telegraph consists of two main parts -- the sounder, which changes the interruption of electricity into meaningful clicks and the sending key, which interrupts the flow of electricity.

Vocabulary:

telegraph	battery	field	attract
sending key	core	sounder	Morse code
magnetism	coil	switch	S. F. B. Morse
electro-magnetism	iron filings	armature	H. C. Oersted

APPENDIX C

Script for Lesson Six on

TELEGRAPH

LESSON 6 (Planned, supering, Teacher Manipulated - see page 9)	24 minutes, 55 seconds
T/S teacher at desk. Cut to T/S telegraph set in center of desk.	Ad lib telegraph. Discuss discovery of magnetism.
(1) T/S Teacher. Super op. Hans Oersted.	Discuss Oersted as discoverer.
(2) (1) C/U demo of dry cell, compass needle and wire. Super animated op. with lighting.	Teacher discusses relationship between electricity and magnetism.
T/S teacher.	
(2) (1) C/U demo of dry cell, wire ar iron filings - super animated ops. of lighting.	Teacher demos relationship between electricity and magnetism.
T/S teacher.	Ad libs other things to know about electro-magnetism.
C/U wire batteries iron filings on desk.	Discusses field.
(1) T/S teacher. Super op. field.	
T/S teacher.	Ad lib controlling fields.
C/U demo area - hold C/U for all four examples: a. coiling wire b. adding core c. increase current d. turning current on and off. Super op.	Ad lib controlling fields.
T/S teacher. Cut to Chalkboard	Review 4 steps.
(2) DI and pan to T/S - teacher at desk. Holds up book at angle for T/S Morse. Super op. Sam F. B. Morse.	Ad libs Morse.
(1) T/S teacher - super op. armature.	Ad lib armature.

	C/U demo area - strip of iron and single wire.	Explain armature is not drawn down.
(2)	Super animated ops. of current.	
(1)	T/S teacher.	
	P/B to 2/S teacher and chalkboard.	Explains 1st method.
	DI to T/S teacher at desk.	Discusses coils.
(2)	C/U demo of teacher coiling wire.	Teacher explains coil not strong enough.
(1)	Super animated ops. of current in coil.	
	T/S teacher.	Explains second method of Morse.
	P/B to 2/S as teacher walks to chalkboard. T/S coil and iron strip drawing.	
	DI and pan to T/S teacher at desk.	Discusses adding core.
(2)	D/U demo of coil and core inserted.	Discusses adding core.
(1)	Super animated ops. of current flowing - bigger arrows for more current.	
	P/B and pan to 2/S of teacher at chalkboard. T/S drawing.	Explains 3rd step.
(1)	DI and pan to T/S teacher at desk. Super op. key.	Discusses use of key.
(2)	C/U demo of knife switch and key.	Explains use of key.
(1)	Super animated ops (2 sets) to show speed of current flow.	
	P/B and pan to 2/S of teacher and chalkboard. T/S drawing. Erase board.	Explains step 4.
	2/S teacher. T/S of drawings as used.	Teacher reviews 4 steps.
(2)	DI and pan to T/S teacher at desk. Super op. Morse code, PB to include book and teacher.	Ad lib Morse code.
	T/S key and sounder. Super op. Morse code.	Sound of telegraph from TV tape recorder.
	DI to T/S teacher.	Teacher closes.

APPENDIX D

Examination for Lesson on

TELEGRAPH

Make sure that your name and the name of the school is on the answer sheet.

These sentences are either right or wrong. If a sentence is right, mark in the space numbered 1 opposite the number of the question on the answer sheet. If a sentence is wrong, mark the space numbered 2 oppositethe number of the question on the answer sheet.

1. We start and stop the magnetic force by means of the sending key.

2. If we construct a strong enough magnetic force, the sounder will make a click when the current is shut off.

3. The telegraph is a device which can send and receive voice messages.

4. A telegraph uses the principle of magnetism.

5. The magnetic force attracts the armature when the current is on.

6. Instead of an electro-magnet, we could use a bar magnet in our telegraph.

7. We take out the soft iron core and replace it with a core of copper. This will increase the magnetic force.

8. The Morse code uses dots and dashes or clicks to represent letters.

9. An electro-magnet can be made from a battery, nail and wire.

10. Magnetism is present whenever electrical current flows through a wire.

11. The magnetic field may be strengthened by placing a glass core in the coil of wire.

12. The sending key is actually a switch.

13. If the flow of electricity increases, the strength of the magnetic field decreases.

14. The telegraph was invented by the Danish scientist, Oersted.

There are three or four answers listed for each of the following items. You are to pick the correct answer, find the number on the answer sheet which corresponds to your answer, and black in this space.

15. The magnetic field of the electric telegraph is turned off when
    1. the current is turned off
    2. the sounder clicks
    3. the sending key closes the circuit
    4. more batteries are added

16. Increasing the number of batteries in a telegraph will
    1. make more rapid clicks
    2. make louder clicks
    3. make slower clicks
    4. make softer clicks

17. The clicks in a telegraph result when the
    1. electro-magnet attracts the armature
    2. the key is left up
    3. magnetic field is weakened
    4. armature is removed

18. An electro-magnet is made stronger by
    1. using wire with thicker insulation
    2. increasing the number of turns of wire on the coil
    3. reversing the current
    4. using smaller wire

19. When the armature does not move, the reason may be
    1. the current is too weak
    2. the armature is made of iron
    3. a spring-type switch is used
    4. too much insulation is used

20. A compass needle moves when a wire carrying electricity is placed near it. This shows
    1. that the needle is not very magnetic
    2. that the telegraph needs electricity
    3. that all magnets repel
    4. that an electric current produces magnetism

21. The sending key of a telegraph
    1. increases the amount of electricity
    2. decreases the amount of electricity
    3. interrupts and completes the flow of electricity
    4. takes the place of the battery

22. Which rod would make the best core for a telegraph set?
    1. plastic
    2. copper
    3. hard rubber
    4. iron

23. The most essential parts of a telegraph are the
    1. sending key and electro-magnet
    2. sending key and battery
    3. sending key and coil
    4. sending key, sounder, and battery

24. The principle of electro-magnetism was first discovered by
1. Benjamin Franklin
2. Hans Oersted
3. Samuel Morse
4. Thomas Edison

25. When the current in a telegraph set is on, the magnetic field attracts the
1. core
2. armature
3. key
4. switch

26. In the telegraph, an armature is used
1. as the key
2. as the sounder
3. as the iron core
4. as the switch

27. A telegraph uses the principles of
1. permanent magnets
2. batteries
3. electro-magnetism
4. nuclear energy

28. In the diagram* of a simple telegraph set, which letter locates the core?
1. a
2. b
3. c
4. d

29. Which part is missing in this diagram* of a telegraph set?
1. core
2. battery
3. key
4. armature

30. What could be done to this telegraph (shown in diagram*) to make it work better?
1. take out one of the batteries
2. increase the number of windings
3. make the armature bigger
4. shorten the core

*Diagram omitted in this reproduction