DOCUMENT RESUME

EL 071 434 EM 010 678

AUTHOR	Palmer, Edward L.
TITLE	Formative Research in the Production of Television for Children.
INSTITUTION	Children's Television Workshop, New York, N. Y.
PUB DATE	73
NOTE	37p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Broadcast Television; Childhood Attitudes; *Child Psychology; Comprehension Development; Early Childhood; Educational Research; *Educational Television; *Formative Evaluation; Media Research; *Production Techniques; Research Design; Research Methodology; *Television Research; Television Viewing
IDENTIFIERS	Sesame Street; The Electric Company

ABSTRACT

The Children's Television Workshop (CTW), responsible for the programs "Sesame Street" and "The Electric Company", conducted extensive research into the systematic use of television to promote the social, emotional, and intellectual growth of young children. Working without precedents in the field of formative research practice, CTW was able to establish an operational model that included behavioral goals, competence testing, experimental production and the Writer's Notebook, a valuable asset emphasizing psychological processes, the child's experimental referents, and unbiased suggestions from the program itself. The subsequent Model for Research on Presentational Learning investigated the relationship between program attributes (appeal, comprehensibility, activity eliciting potential) and the internal compatibility of elements and "viewer outcomes". CTW found that this formative research approach was compatible with the trend toward explicit definition of instructional objectives followed by systematic trial and revision of instructional systems for achieving them. (MC)

FORMATIVE RESEARCH IN THE PRODUCTION OF TELEVISION FOR CHILDREN

Edward L. Palmer

Children's Television Workshop
1 Lincoln Plaza
New York, New York 10029

I. The CTW Operational Model

The principal activites undertaken in the production of "Sesame Street" have come to be viewed by CTW as a model, and this model was again applied in the production of "The Electric Company." If there is a single, most critical condition for rendering such a model of researcher-producer cooperation effective, it is that the researchers and the producers cannot be marching to different drummers. The model is essentially a model for production planning. More specifically, it is a model for planning the educational (as opposed to the dramatic) aspects of the production, and the formative research is an integral part of that process. In the case of "Sesame Street" and "The Electric Company," at least, it is hard to imagine that the formative research and curriculum planning could have been effective if carried out apart from overall production planning, either as an a priori process, or as an independent but simultaneous function.

The activities included in the model are presented below in their approximate chronological order of occurence.

A. Behavioral Goals

As the initial step toward establishing its educational goals, CTW, in the summer of 1968, conducted a series of five three-day seminars dealing with the following topics:

1. Social, Moral, and Affective Development
2. Language and Reading
3. Mathematical and Numerical Skills
4. Reasoning and Problem-Solving
5. Perception

Formative Research in the Production of Television for Children

Edward L. Palmer

There is currently in the United States unparalleled interest in the systematic use of broadcast television to promote the social, emotional, and intellectual growth of young children. Support for this movement lies in the recognition that television is ubiquitous, reaching into 97% of all U.S. households; that young children are exposed to upwards of thirty hours of television fare each week; that while they learn a great deal from what they watch, there have been far too few significant attempts to plan program content in order to address important areas of learning and development systematically; and that no other approach can promise to deliver so much to so many at so small a unit cost.[1] An important feature of this movement is its emphasis upon "formative" planning and research, whereby important objectives are first clearly identified and systematic audience tests are then carried out in order to evaluate progress toward their achievement during the actual course of a program's production. Formative research is typically contrasted with summative research, which is concerned with follow-up testing to determine the educational effect of new products and practices when actually put in to use.[2] What is to be presented here is a description of the approaches to formative planning and research taken by CTW in the

production of "Sesame Street" and "The Electric Company." Although an effort will be made throughout to discuss these approaches in ways that will suggest their potential usefulness in the development of other new educational products and practices, this paper is not a compendium of research results for the general guidance of producers of educational materials; it is, rather, a case study. As a case study, it will focus on the overall operational framework within which CTW's formative research proceeds, the strategies and rationale for the design of formative field research methods, organizational and interpersonal conditions, and similarities and contrasts between the functions and the methods of formative research on the one hand and those of more traditional research approaches on the other.

The Children's Television Workshop (CTW) was created in 1968 to produce a series of 130 hour-long broadcast television programs for preschool children, with special emphasis on the needs of the urban disadvantaged child. The result, the now well-known "Sesame Street" series, in June of 1972 completed its third broadcast season. The formative research methods developed and applied in the production of "Sesame Street" have been reapplied and extended by CTW in planning and producing its second major program series, "The Electric Company," designed to teach selected reading skills to children from seven through ten years of age.

At the beginning of the "Sesame Street" project, the functions formative research could serve and the field methods it could apply were not at all clear. There were no precedents of sufficient scope and generality, either from the field of educational television or from the field of educational planning and research in general, to provide clear

guidelines. What has been learned about the formative planning and research process at CTW has come about under quite unusual circumstances, and since it is unlikely that these conditions will ever be duplicated in substantial detail, it remains to be seen what sorts of new or modified approaches will be required in different situations. What is presented here certainly cannot be construed as a dependable recipe that will assure the success of other like ventures.

Among the unusual circumstances associated with the Workshop's productions, some, no doubt, had quite a direct bearing on the effectiveness of the formative research. For instance, the two Workshop projects were well funded, each budgeted in its first season at upwards of seven million dollars for production, research, and related activities. This level of support made it possible to utilize high-level production talent and resources and to make extensive use of expert educational advisors and consultants. In addition, both projects enjoyed unusually long periods of time -- in each case, approximately eighteen months -- for pre-broadcast planning and research. Time and resources were available to plan their respective curricula carefully and to state their educational objectives in very explicit terms, so that producers and researchers alike, as well as the independent evaluators who were carrying out pre- and post-season achievement testing projects, could proceed without ambiguity of purpose and in a coordinated fashion. Had there been ambiguity, either in terms of the particular objectives to be addressed or in terms of the commitment of the producers to direct each segment toward the achievement of one or more of those objectives, the formative research could not have been useful, for there would have been no clear criteria against which to evaluate a program segment's

effectiveness.

Also unusual in the CTW case were the organizational and interpersonal relationships between the in-house research and production staffs, and the policy followed in production recruiting. All of the key producers came from commercial production backgrounds. None had formal professional training in education or experience in educational television production. Yet, they were given the responsibility for final production decisions. They did not work under the researchers, nor did the researchers work under them. The intended function of the formative research was to provide information which the producers would find useful in making program-design decisions, relative to both appeal and educational effect.

To the extent that the formative research worked, it worked in large measure because of the attitudes taken toward it by producers and researchers alike. The producers were committed to experimenting with the cyclic process of empirical evaluation and production revision, and tended to have the creative ability not only to see the implications of the research, but to carry these implications through into the form of new and revised production approaches. Accordingly, the usefulness of CTW's formative research has depended not only upon the qualities of the research itself, but also upon the talents of those who put its results to use. Moreover, the producers never expected the research to yield full-blown decisions; they recognized that its function was to provide one more source of information among many. From the research side, because the responsibility for final production decisions resided with the producers, it was necessary to develop and apply only methods which producers themselves found useful. Accordingly, the producers were involved from the outset in all research

planning. No observational method was ever persistently applied, and no specific study was ever taken into the field without their participation.

At this time, there is no tradition of accumulated knowledge in the area of formative research practice. This is partly because so little research of this type has been done, but it is even more a result of the fact that it has only recently come to be recognized as a distinct and systematic field of endeavor. Early contributions to the explicit and systematic conceptualization of the field have been made by Cronbach,[3] Hastings,[4] and Scriven.[5] Formative research studies associated with specific product improvements are reported for various media, content areas, and student levels: e.g., for televised instruction, by Gropper and Lumsdaine;[6] for programmed instruction, by Dick;[7] for kindergarten instruction in conceptual skills and for art, by Scott.[8,9] However, the present scope and depth of the formative research literature is in no way commensurate with its promise for education. The promise of the approach is that it will provide designers of educational products and practices with empirical data far more directly pertinent to their respective media, materials, and learning conditions than are the results of traditional, more basic research. For the field of educational television in particular, it offers ways to help bring about planned effects. With mass broadcast distribution making it possible to reach hundreds of thousands or even millions of viewers (in the case of "Sesame Street," the weekly audience is estimated to include more than eight million different children), it behooves the producers to employ every reasonable means for helping to ensure in advance that the programs will achieve their objectives.

I. The CTW Operational Model

The principal activites undertaken in the production of "Sesame Street" have come to be viewed by CTW as a model, and this model was again applied in the production of "The Electric Company." If there is a single, most critical condition for rendering such a model of researcher-producer cooperation effective, it is that the researchers and the producers cannot be marching to different drummers. The model is essentially a model for production planning. More specifically, it is a model for planning the educational (as opposed to the dramatic) aspects of the production, and the formative research is an integral part of that process. In the case of "Sesame Street" and "The Electric Company," at least, it is hard to imagine that the formative research and curriculum planning could have been effective if carried out apart from overall production planning, either as an a priori process, or as an independent but simultaneous function.

The activities included in the model are presented below in their approximate chronological order of occurence.

A. Behavioral Goals

As the initial step toward establishing its educational goals, CTW, in the summer of 1968, conducted a series of five three-day seminars dealing with the following topics:

1. Social, Moral, and Affective Development
2. Language and Reading
3. Mathematical and Numerical Skills
4. Reasoning and Problem-Solving
5. Perception

The seminars, organized and directed by Dr. Gerald S. Lesser, Bigelow Professor of Education and Developmental Psychology at Harvard University, were attended by more than a hundred expert advisors, including psychologists, psychiatrists, teachers, sociologists, film-makers, television producers, writers of children's books, and creative advertising personnel. Each seminar group was asked to suggest educational goals for the prospective series and to discuss ways of realizing the goals on television. The output from the initial seminars was then systematically organized, refined, and operationalized by the CTW staff and Board of Advisors. This work resulted in specific goals stated in behavioral terms. These behavioral objectives served as a common reference for the program producers and the designers of the follow-up achievement tests.[10] Essential coordination of production and evaluation was thus assured.

B. Existing Competence of Target Audience

While the statement of goals specified the behavioral outcomes the program hoped to achieve, it was necessary to ascertain the existing range of competence in the chosen goal areas among the target audience. The Workshop research staff therefore undertook as its initial formative research effort a compilation of data provided in the literature, as well as some testing of its own, to determine the competence range. The resulting information helped guide the producers in allocating program time and budget among the goal categories and in selecting specific learning instances in each goal area.

C. Appeal of Existing Materials

To be successful, CTW had to capture its intended audience with

an educational show whose highly attractive competition was only a flick of the dial away. Unlike the classroom teacher, the Workshop had to earn the privilege of addressing its audience, and it had to continue to deserve their attention from moment to moment and from day to day. At stake was a variation in daily attendance which could run into millions. Measuring the preferences of the target audience for existing television and film materials was therefore crucial in the design of the new series.

D. Experimental Production

Seminar participants and CTW advisors had urged using a variety of production styles to achieve the curriculum goals adopted. Research had confirmed the appetite of the target audience for fast pace and variety. Accordingly, the CTW production staff invited a number of live-action and animation film production companies to submit ideas. The first season of "Sesame Street" would eventually include the work of thirty-two different film companies.

Prototype units of all film series produced by or for the Workshop were subjected to rigorous preliminary scrutiny and empirical field evaluation. Scripts and storyboards were revised by the Workshop producers on the basis of recommendations from the research staff; further revisions were made after review by educational consultants and advisors; and finished films were tested by the research department with sample audiences. Some material never survived the process. Four pilot episodes were produced for a live-action film adventure series entitled "The Man From Alphabet," but when the films were shown to children they failed to measure up, either in appeal or educational effect, and the series was dropped. Sample videotaped material went through the same process of

evaluation, revision, and occasional elimination.

By July of 1969 a format for the program had been devised, a title had been selected, a cast had been tentatively assigned, and a week of full-length trial programs had been taped.

Completed prototype production elements were tested by the research staff in two ways: (1) the appeal of the CTW material was measured against the appeal of previously tested films and television shows, and (2) the CTW material was tested for its educational impact under a number of conditions. For instance, field studies were conducted to determine the effect of various schedules of repetition and spacing, of providing the child with preliminary or follow-up explanation, of presenting different approaches to a given goal separately or in combination, and of the relative effectiveness of adult vs. child voice-over narration. Extensive observation of viewing children provided information regarding the child's understanding of various conventions of film and television technique. Upon conclusion of each research study, the results were reported to the producers for their use in modifying the show components tested and for guidance in the production of subsequent elements. It should be noted that this progress testing also served a formative research function for the ETS staff by field-testing the summative evaluation instruments and administration procedures.

E. The Progress Testing

The evolution of "Sesame Street" did not end with the first national broadcast on November 10, 1969. Formative research continued throughout the six-month broadcast season. During this time, it became possible to begin examining the cumulative impact of the series. Accordingly,

the research staff instituted a program of progress testing of the show's effectiveness, using the summative evaluation instruments designed by the Educational Testing Service (ETS) of Princeton, New Jersey. A sample of day-care children, predominantly four- and five-year-olds, was pretested prior to the first national telecast. One-third were tested again after three weeks of viewing the show; the same one-third and an additional one-third were tested after six weeks of viewing; and the entire group was tested after three months of viewing. Comparisons between experimental (viewing) and control (non-viewing) groups at each stage of the testing gave indications of strengths and weaknesses both in the execution of the curriculum and in the production design. Other independent formative studies of program appeal and of the responses of viewing children also influenced production decisions during this period.

F. Summative Evaluation

The summative research and evaluation carried out by Educational Testing Service on each of the program's first two seasons[11] followed a plan developed in consultation with CTW staff and advisors. Participation of ETS representatives in all main phases of pre-broadcast planning helped to ensure coordination between program development and follow-up testing.

For the first-season study, ETS developed and administered a special battery of eleven tests covering the major CTW goal areas to a sample of children from Boston, Philadelphia, Durham, and Phoenix. The groups included three-, four-, and five-year-olds in urban and rural settings, from middle- and lower-income families, in both home and day-care situations. A special side study related to children from Spanish-

speaking homes. Other measures assessed home conditions, parental expectations for the children, and the like. Where the results of the first season's summative research fed into production decisions for the second season, they took on a formative function. For example, the summative data indicated that children's prior knowledge and skills had been underestimated in some goal areas and overestimated in others. This was taken into account in programming the second season of "Sesame Street."

G. Writer's Notebook

As the producers and writers began to develop scripts, animations, and films addressed to particular behaviorally-stated goals, it became apparent that the goal statement was not a wholly adequate reference. After having been given several successive assignments in the same goal area, they began to express the need for extended and enriched definitions which would provide creative stimulation. Gradually, through trial and error, a format for a Writer's Notebook was developed which the producers and writers themselves found useful.

The Notebook emphasized four criteria:

1. To focus on the psychological processes involved in a particular form of behavior.

2. To exploit and extend the child's own experiential referents for such behavior.

3. To prompt the creation of various similar approaches by the producers and writers themselves, by presenting them with highly divergent examples.

4. To provide suggestions free of any reference to particular characters or contexts from the television program, so that the ways in which the suggestions could be implemented would be left as open and flexible

as possible.

These features of the Writer's Notebook may be highlighted through an example. In the broad area of Symbolic Representation, the word-matching objective is stated as follows: "Given a printed word, the child can select an identical printed word from a set of printed words." To implement this objective, the Notebook would encourage the producers to use words with different numbers of letters; to vary the location within the word of the letter or letters which fail to "match"; to present various matching strategies, such as comparing two given words letter by letter, moving initially separated words into physical superimposition, or spelling out each of two given words and comparing to see if one has made the same sounds both times. Another recommended approach was to make use of the "sorting" format already familiar to viewers, wherein three identical things (in this case, words) and one odd thing are presented simultaneously, along with a standard song which invites the viewer to find the one which is different. Still another was to construct a letter-by-letter match for a given word by choosing from a large pool of letters. To encourage still further divergent approaches, another recommendation was to present pairs of words which matched in one sense but not in another, e.g., a pair in which the same word is presented in different type faces, or in which one member of the pair is the upper- and the other the lower-case version.

Similar suggestions were developed for other goal areas, as requested by the producers and writers. Again, suggestions were solicited from advisors and consultants. In addition, the Notebook provided a place and a format for collecting the ideas of the in-house research staff

and a channel for helping to ensure that these ideas would be seen and used.

II. A Model for Research on Presentational Learning

Learning by way of televised presentations does occur, and the objective of the formative research at CTW is to discover principles of program design by which this type of learning may be improved. In the specific case of "Sesame Street" and "The Electric Company," the research seeks principles of presentational learning appropriate to their particular educational goals, audiences, and production techniques. However, there are more basic objectives also, one of which is to create generalizable formative research methods and practices, and another to discover generalizable principles of presentational learning.

In pursuit of these objectives, a model for research on presentational learning is being developed. A summary outline of the model is given in Figure 1. One point highlighted by the model is the need to

---

Insert Figure 1 About Here

---

identify and rigorously define features of program design (independent variables) which are reliable predictors of learning and learning-related outcomes among viewers (dependent variables). The statements which relate program-design features to empirical data are principles of presentational learning, which are potentially generalizable to other televised presentations and to other media. Moreover, they are seen as principles which will have to be taken into account within any comprehensive theory of presentational learning.

The model is tentative, intended more as a point of departure for

MAJOR PROGRAM ATTRIBUTES (Categories of Independent Variables)	VIEWER OUTCOMES (Dependent Variables)	PRINCIPLES OF PROGRAM DESIGN (Statements Linking Specific Independent and Dependent Variables)
APPEAL	Visual orientation, attention, attitude, channel selection, etc.	Within each major attribute category, any number of specific features may operate to affect viewer outcomes. Statements linking specific and well-defined program features to learning outcomes belong in this column.
COMPREHENSIBILITY	Comprehension.	
ACTIVITY ELICITING POTENTIAL A. Potential to elicit Motor and Psychomotor Activity	Verbalization, gross physical acts, imitation, direction-following, etc.	
B. Potential to elicit Emotive Activity	Arousal, attitude, etc.	
C. Potential to elicit Intellectual Activity	Synthesizing or integrating, forming concepts or principles, generalizing, comparing, evaluating, predicting, etc.	
INTERNAL COMPATIBILITY OF ELEMENTS	Attention to signal vs. noise; integration of elements (visual-visual, auditory-visual, auditory-auditory); etc.	

Fig. 1. -- A model for research on presentational learning. Major attribute categories are very general dimensions of the televised presentation. Viewer outcomes represent the effects of the presentation on viewers. Within each attribute category, any number of specific program features (variables) may operate to affect viewer outcomes. Principles of program design are hypothesized or well validated relationships between scientific program features and specific viewer outcomes.

further exploration than as a fixed scheme. However, as it stands, it plays a central role in guiding CTW's formative research studies, and in organizing their results. It features three main categories of program attributes, including Appeal, Activity Eliciting Potential, and Internal Compatibility of Elements. Activity Eliciting Potential presently has three main subcategories: (a) motor and psychomotor activities, including imitation; (b) emotive activity; and (c) intellectual activity. The category of intellectual activity encompasses, in turn, a number of program attributes, including those responsible for such viewer outcomes as integrating (e.g., integrating the audio and visual portions of the presentation, or the successive elements of an unfolding plot); generalizing (relating the presentation to past experience or to future possibilities); anticipating (predicting possible upcoming events in the presentational sequence); forming new concepts and principles; imputing the motives and intentions of characters, or the program's producers; evaluating, as in assessing the credibility of the plot, message, or character premises, judging the quality of the performance or the technical quality of the production; and the like. A more extensive discussion of the main attribute categories contained in the model, and of some of the research methods employed within each, is presented later. But first, there needs to be a clarification of the intended role of the model in the formative research process.

The development of this model is not so much an outgrowth of the formative research process as an integral part of it. It was developed in its present form with a number of specific objectives in mind. First, in form, it is highly simplified. It is intended that by virtue of its simplicity, it will serve as a convenient checklist for both producers and researchers, helping to suggest program attributes which need to

be taken into account in creating new segments or in designing new formative field studies. It is also intended to serve a number of useful organizing functions. First, the various field research methods are organized according to the categories of attributes to which their results relate. This is important, in that the various methods within a given attribute category tend to elucidate complementary sets of program-design features and viewer outcomes, and their organization within broad attribute categories tends to highlight this complementarity.

Still another intended function of the model is to bring together within a small and therefore convenient number of attribute categories, the great number of hypothesized program-design principles growing out of the formative research.

Finally, the model is open-ended, in that it necessarily falls short of presenting an exhaustive list of program attributes and audience outcomes. This open-endedness gives rise to another of the model's potentially valuable qualities, which is that it invites other researchers and producers to identify other potentially significant program attributes, audience outcomes, research methods, and principles of program design. To cite an example of this effect, CTW researcher Dr. Joyce Weil, while participating in the research planning for a new adult television series on health, extended the categories of the model to encompass two variables considered important in developing such a series, namely: (1) credibility, on the independent variable side, which relates to belief; and (2) memorability, which relates to long-term recall and to factual distortion.

Principles of presentational learning such as would appropriately appear in the third column of Figure 1 were characterized earlier as statements which link features of the presentation to empirical outcomes.

One important consideration in establishing these principles is to define program variables with enough precision to yield high interrater reliability, i.e., so that two or more raters ranking the same program segments according to a given definition produce the same or nearly the same rankings. Second, the principles of presentational learning based upon these definitions need to possess predictive validity, such that rankings assigned to segments in terms of a given program variable yield better-than-chance predictions of their effectiveness.

CTW is exploring a number of procedures for identifying program-design variables and linking them empirically to their effects on viewers. A procedure being developed by Rust has been applied to data on appeal.[12] In this procedure, empirically derived scores on appeal are determined for each 7.5-second interval throughout a program. The scores for the various intervals are converted to unit normal (z-score) form. Lists are then made, so that highest and lowest scoring segments are displayed separately. These are then scanned by researchers and producers in an attempt to define program-design variables which appear to differentiate the high from the low. The definitions of variables so derived are then tested for interrater reliability, and improved, if necessary. Finally, the definitions are applied, a priori, to new segments, and evaluated for their power in predicting the measured appeal of those segments. The procedure may be applied with other program attributes as well. Another procedure is currently being explored for CTW by Dr. Gavriel Salomon of Hebrew University in Jerusalem. In Salomon's procedure, each segment is scored according to a number of different program-design variables which may contribute to measured outcomes along a given dimension on the criterion side. As with Rust's procedure, it is important that

the definitions of these program-design variables possess high interrater reliability. Through the application of multiple and partial regression analyses, Salomon then can evaluate both the individual and interactive contributions of the predictor variables (the program-design variables) toward producing the measured outcomes. A third method, currently being explored by the author and CTW researcher Mr. Lewis Bernstein, is an attempt to apply to the same general problem Guttman's facet analysis procedures.[13] This approach differs from the others mentioned above primarily in offering a range of alternative models, both parametric and non-parametric, for relating program features to outcomes.

A. Formative Research on Appeal

Appeal research bears on a wide range of program decisions. It reveals the effects of various forms and applications of music, and of music as compared with other types of elements. It indicates the most and least popular forms of live-action films, animations, puppets, and live performers. It indicates the attention-holding power of various types of individual or interpersonal activities, such as showing one person guiding another through a difficult task in a supportive versus demeaning manner; presenting conflict resolution through the arbitrary use of power versus cooperation; revealing the simultaneous perspectives of different characters; and portraying the struggle of an individual toward an achievement goal or toward improvement upon his own past performance, to mention a few.[14]

Appeal research also helps to indicate for various conditions the amount of time over which attention can be maintained; the optimum amount of variety and the optimum pacing of events; the relative holding power of program elements which are and are not functionally relevant

to the action; the ability of a segment to bear up under exact repetition; the most and least salient (memorable) characters; and the effectiveness of special techniques such as pixilation, fast and slow motion, and unusual camera angles. In addition, research on appeal can show growth or decline in the popularity of specific program elements over time; the most and least effective uses of dialogue, monologue, and the voice-over technique; the relative effectiveness of ordinary or caricaturized voices; and the effect of sparse and pointed versus sustained verbalizations. It also can reveal the effects of incongruity, surprisingness, or fantasy, as compared with straightforwardness, predictability, and realism; the effect of different motives or intentions on the part of the characters; of episodic versus linear styles of continuity; and also of familiar versus unfamiliar conventions and symbols dealing with time, sequence, interpersonal relationships, and the like. Finally, this type of research can be used to investigate characteristic individual or group preferences vis-a-vis such program design features.

Since both "Sesame Street" and "The Electric Company" were designed according to a magazine format, with successive brief segments addressed to very explicit educational objectives, it was important to maintain high program appeal on a moment-to-moment basis. Accordingly, a method, referred to as the distractor method, was introduced which yields appeal data throughout the course of a program. The distractor method consists of placing one child at a time in a simulated home-viewing circumstance. A black-and-white videotaped recording of a television program is presented simultaneously with a color slide show flashed at 7.5-second intervals on a rear-projection screen equipped with an eighty-slide carrousel. The rear-projection screen, which is approximately the same size and

height from the floor as the television screen, is placed at about a forty-five degree angle from the child's line of vision to the television screen. The child himself is seated in a chair three to four feet away from and facing the television, but is free to move about within the confines of the room at any time. A continuous record indicates when the child's eyes are directed toward or away from the set. For each viewer, the eighty-slide carrousel is started at a different slide, so that the stimulus competing with a given 7.5-second interval of the televised presentation is different for each viewer. Composite graphs of the results are studied by the researchers and producers in various ways to attempt to identify the elements of program content responsible for high and low appeal.

A frequently used complementary form of appeal testing consists of taking observations upon successive sets of viewing groups, where each group typically contains from three to five viewers. Usually, four to six such groups are observed in testing a program. A detailed record is kept according to predefined categories of visual, verbal, and motor behaviors. The visual behavior of children in viewing groups provides a cross-check on the distractor results. The record of verbal and motor responses, in addition to reflecting upon program appeal, helps to identify program approaches most and least effective in eliciting active participation. The fruitfulness of this particular approach is very much a function of the training and the creative interpretive skills of the researcher.

Audience surveys can provide much additional material on program appeal, as can structured interviews, in which the salient and lasting as opposed to the immediate appeal of various program features can be determined. The salient appeal of an element is particularly important

where it is necessary to attract a voluntary audience day after day.

B. Formative Research on Comprehensibility

Once children's attention is assured it becomes critical to assess the information extracted from the program by the children. What did they grasp of the intended instructional points? Can they interpret the motives or intentions of the characters? Do they remember the relevant or the trivial?

Comprehensibility testing, while useful in evaluating a viewer's understanding of the dramatic action, is undertaken primarily for the purpose of pointing up program design features important in producing instructional effects. As such, it focuses upon the qualities of the televised message as these interact with learner characteristics. It is instructive to the producers to have an empirical check on their own assumptions about the comprehensibility of program design features they are employing, and even limited amounts of field research can help them to maintain a generalized sensitivity to this important attribute.

Although CTW's research in the area of comprehensibility has just begun, the ultimate objective is to identify and set down specific program-design principles. These principles, though tentative, include the use of production approaches which can help to clarify the relationship between an event occurring on the screen and the theme, the plot line, or the logical progression of the dramatic component, or between the instances and non-instances of a concept, the referents and non-referents of a term, or the most and least effective of a set of proposed solutions to a problem.

The unique conventions and capabilities of the television medium

these conventions are presented will determine their comprehensibility to the viewer, and thus their effectiveness in communicating the meanings intended. These include the use of the flashback technique, of special lighting effects or special combinations of music and lighting, the use of various camera perspectives, of fast or slow motion, of pixilation, and of the matched dissolve between objects. They also include the close juxtaposition of events in order to establish a metaphoric or analogic relationship between them, and the use of conventions having to do with fantasy, such as presenting puppets and cartoon characters who move and talk like humans. Still others include the creation of "magical" effects, such as making an object instantly appear or disappear from a scene, or grow smaller or larger, and the use of exaggerated motions and consequences, as with slapstick and "banana peel" humor, to mention a few. Other conventions which can be used in more or less comprehensible ways are the speech balloon, the rules of games presented for instruction or entertainment, and rules involved in reading, spelling, mathematical operations, the interpretation of maps, and the like.

Still other facets of comprehensibility relate to timing, sequencing, and the use of redundancy, as in repeating an event exactly or with an illuminating variation, in restating a point from alternative perspectives, and in making use of introductions or reviews. The list could go on indefinitely, a fact which helps suggest the significance of this attribute in educational television research.

Again, as with appeal testing, this area employs not one but a family of complementary research methods. One very useful approach is

so as to "freeze the frame," and then to ask viewers about events le
up to or likely to follow from the pictured situation.

If the research concern has to do with a character premise or with a character's motivation, the viewer might be asked, "What kind person is he?" or "Why did he do (say) that?" or "What do you think will do next?" "Why do you think that?" and so on. In one segment designed for "The Electric Company," The Short Circus, a musical roc group made up of children, was shown singing a song which contained letter combination "ow" several dozens of times. As the "ow" song sung, the printed "ow" was shown a number of times simultaneously. The intention was to provide repetitive practice in associating the spoken and printed forms of this particular letter combination. By using the method of freezing a single frame, it was possible to eva the extent to which members of the target audience actually perceiv the speech-to-print correspondence. In this case, the letter combi tion was frozen on the screen at a point late in the song, and as t experimenter pointed to the printed letter, the subjects were asked question of the form: "Why is that there?" "What does it mean?"

In a related method, a program or segment is played once or tw through. It is then presented once again, but this time without th sound (or, in a variation upon the method, with the sound but witho the picture), and the viewer is asked either to give a running acc of what is happening or to respond to specific questions.

A strength of comprehensibility testing relative to traditional forms of summative evaluation is the opportunity it provides for discriminating between the most and least effective of the many individual segments devoted to a particular achievement objective. A potential but largely surmountable limitation is the tendency for these methods to produce biased results. Because comprehensibility testing is performed as the program is being viewed, and because the viewer knows he will be questioned, there is typically an over-estimation of a segment's effectiveness. In practice, this bias can be subjectively discounted, at best, and must further be weighed against the possibility that segments which produce no measurable learning when presented in isolation may be effective in combination or when presented along with an appropriate introduction or review. However, these limitations do not detract seriously from the usefulness of such methods. The bias can in fact be turned to an asset, as when it can be shown that a segment of questionable value fails to make its point even when evaluated by means of a liberally-biased method.

C. Formative Research on Activity Eliciting Potential

A frequently expressed point of view about the potential of television for instructional purposes is that due to passivity of the viewer, the medium is virtually powerless to produce learning. However, since it is patently obvious that television does teach, it seems desirable to explore, conceptually, how this capability comes about, and operationally, in what proper and constructive ways it may be exploited. The position taken here is that in spite of the apparent passivity of television viewing, the medium's activity eliciting potential is perhaps the chief basis for whatever instructional value it possesses.

One significant form of activity television can elicit is intellectual activity. Others include verbal behavior and gross physical acts, ranging from television-modified performance on tests of attitudes and achievements to the imitation of televised models. It is important to note that the concern of the medium can be either to exploit these effects as instruments of instruction or to foster them as instructional objectives.

Some examples of intellectual activities include integrating separately-presented items of information, anticipating upcoming events, forming new concepts, imputing the motives and intentions of characters, following progressively developed dramatic and instructional presentations, and guessing answers to questions. Also, the viewer may take an active role in evaluating relationships between premises and conclusions, between information given and interpretations made of it, and between behavioral ideals and the actual behaviors carried out by the performers. The viewer may also relate televised information to his own prior experiences and to his future plans.

Tentative indices from formative research on the activity eliciting capabilities of the medium suggest that many of its assumed limitations may be at least partially surmountable. For example, it is often assumed that learning through trial and error or through trial and reinforcement cannot occur through one-way televised presentations, on the basis that there is no opportunity for either reinforcement or information feedback to be tied to an action of the learner. This is not a trivial issue, from a practical standpoint, since vast amounts of money may yet be spent studying the use of two-way communication systems, which do make it possible to do so. It turns out that conceptually it is possible to effect trial-and-error learning through one-way television, simply by

the use of "if" statements. That is, the viewer may be offered a choice among provided alternatives, given time to make his choice (his point of most active involvement), and given reinforcement, or an accuracy check, of the form: "If you chose thus and so, you were correct (incorrect)." Empirical studies may or may not support the viability of such an approach, but it certainly deserves further investigation.

The notion that certain activities containing a motoric component can be learned only through direct experience is also questionable. For example, direct experience in the construction of alphabetical characters may have its most significant effect upon learning by controlling the scan of the eye over the configuration of the letter, by providing extended or repeated exposure to the letter, or by providing an occasion for the most common errors to be made and corrected. But all of these are features one-way television can either duplicate or simulate. We need to know more about the possibilities for television simulation of learning conditions in which direct, hands-on experience traditionally has been considered essential. We also need to know more about the entry skills required with such simulations in order for learning to occur, and about possibilities for employing simulation in order to facilitate subsequent performance in hands-on learning contexts.

All this is not an argument in favor of unduly widespread substitution of television for physical activity among children, by the way, nor is it intended to deny the great importance of extensive direct experience in learning, especially in early learning. It is intended, rather, to urge open and positive consideration of some of the possible but not yet systematically explored capabilities of the television medium.

D. Formative Research on Internal Compatibility

Internal compatibility is a program attribute which has to do with the relationship of different elements appearing within the same segment. The basic strategy underlying both "Sesame Street" and "The Electric Company" is to attempt to effect instruction through the use of television's most popular entertainment forms. To this end, it is essential that the entertainment and educational elements work well together. Without the entertainment, attention strays, and without the education, the whole point of the presentation is lost. In segments where these elements are mutually compatible, the educational point is an inherent part of the dramatic action and often is actually enhanced in its salience as a consequence. In others, the entertaining elements override and thereby actually compete with the educational message. Other cases in which the relationship of elements becomes a concern have to do with auditory-visual, auditory-auditory, and visual-visual compatibilities.

The objective of formative research in this area is to shed light upon the program design features which make for a high or low degree of compatibility. In one method used for assessing compatibility, a panel of judges is asked to rate each segment of a program according to a pre-determined set of categories defining the extent to which a segment's entertainment either facilitates or competes with the instructional content. Working from each segment's compatibility score, which is a composite of the ratings given by the various judges, it is possible to identify sets of high-rated and low-rated segments, and to present the producers with an interpreted list of each type. The interpretations identify program-design features to be emulated, revised, or avoided.

Another method involves eye-movement research, which is especially useful in the case of "The Electric Company," because of the extensive

presentation of print on the screen and the desire to find ways of motivating the child to read it. In most segments, the print appears on the screen along with competing stimuli. By using the well-known technique in which a beam of light is reflected from the cornea of the eye of the television-viewing subject and recorded on a photographic device for later interpretation, it is possible to identify the conditions under which the print is and is not read. Once again, the results indicate program-design features worth emulating and approaches which need to be revised or avoided.

Among the important program features focused upon by this method are the location of the print on the screen; the effectiveness of various ways of animating print; the effect of the exact repetition of segments upon the elements attended to; and the usefulness of special motivational devices, such as telling all but the punch-line of a joke, and then presenting that in print.

Methods for measuring eye-movement obviously have implications also for assessing other program attributes, such as appeal and activity eliciting potential.

E. Extending the Methods and Findings of Formative Research

The search for ways to improve the contributions of formative research is a continuing process at CTW. Suggestions for new research methods and critical appraisals of current methods are encouraged not only from the CTW research staff, but also from a wide variety of outside consultants (e.g., Rust,[15] Mielke and Bryant,[16,17] O'Bryan[18]). Although not a primary mission at CTW, there is also high interest in relating

the ever-expanding body of formative research findings to more general theories of learning. An excellent contribution has been made by Dr. Gerald S. Lesser.[19]

III. Organizational and Interpersonal Factors

As technologically sophisticated forms of instruction come into increasing prominence, it will be necessary to make increased use of production teams whose members possess a diversity of highly specialized talents. In anticipation of this trend, we need to know more about related organizational and interpersonal conditions. These conditions deserve attention in any attempt to establish a working partnership between television research and production groups, and they play a strikingly more prominent role in the formative research context than in the context of more traditional approaches to educational research.

An important factor in CTW's case has been the opportunity afforded by an eighteen-month pre-broadcast period for the members of the two groups to learn about each others' areas of specialization. Another has been the attitude that every new formative research approach is an experiment, to be continued or discontinued depending on its merits as evaluated by the producers themselves.

The fact that CTW's researchers and producers possess not the same, but complementary skills is also significant, largely because it makes for clear and distinct functions on the part of each group. Still another factor is that the producers, before joining the project, made the commitment to try to work with formative research. This prior commitment helped to support the cooperative spirit through the early, more tentative period of the effort. Also, research never takes on the role of adversary, to

be used against the producers in winning a point or pressing for a particular decision. The producers hold the final power of decision and are free to ignore research suggestions if production constraints require it.

In all, the factors consciously dealt with in the interests of researcher-producer cooperation have ranged from the careful division of labor and responsibility to housing the two staffs in adjacent offices, and from patience and diplomacy to occasional retreat.

IV. The Distinctive Role and Functions of Formative Research

Formative research is distinguished primarily in terms of its role as an integral part of the creative production process. It is important to maintain a clear distinction between this type of research, on the one hand, and that undertaken in order to test the validity of a theory or the measureable impact of an educational product or practice (i.e., summative research), on the other. Research undertaken in the context of scientific validation is concerned with effects which have been hypothesized, a priori, within the framework of a broader deductive system; with the use of empirical and statistical procedures well enough defined so as to be strictly replicable (at least in principle); and with the highest possible degree of generalizability across situations. In contrast, while research carried out within the formative context can possess all these same characteristics, it need not and typically does not. The main criterion for formative research recommendations is that they appear likely to contribute to the effectiveness of the product or procedure being developed. It is neither expected nor required that they be validated by the research out of which they grew. Establishing their validity

is the function of summative research.

As this view implies, to achieve the objectives of formative research, it is often necessary to depart from traditional research practices and perspectives. This is not to say that experimental rigor has no place in the formative context. However, for example, even where strict experimental and control conditions have been maintained, there is seldom anything to be gained by using tests of statistical significance. The creative producers often prefer to work directly with information about means, dispersions, and sample size. Also, whereas matching of experimental and control groups on the basis of pretest scores is discouraged where inferential statistics are to be used because of the conservative effect on the usual tests for the significance of the results, such matching can be very useful, for efficiency, and to maximize the reliability of information based on small samples.

In the area of sample selection, it also can be useful to depart from the traditional practice of including all age and socioeconomic groups for which the educational materials are intended. Time and effort may often be saved by selecting a sample of average performers, of performers from the high and low extremes, or, where the intent is mainly to upgrade the lowest performers, a sample of only those. In general, where biased methods of sampling and biased methods of testing are more efficient than unbiased methods, and where the objective is not to make accurate population estimates, it is often useful to exploit the very biases which quite properly would be avoided in other research situations.

In practice, it tends to be difficult for researchers trained and experienced in traditional approaches to adopt an appropriate formative

research point of view. In the formative situation, their first responsibility is to improve a specific product or practice, and not to contribute to a general body of knowledge (though the two objectives certainly are not incompatible). Studies must first address the information needs of the product designers and not primarily the individualistic or special theoretical interests of the researchers. Covering a wide range of empirical questions may deserve priority over rigorous reporting or establishing careful experimental conditions, where it is economically impossible to achieve both, and where the usefulness of the results is not unduly compromised as a consequence. Quantitative indices such as percentages, and highly detailed item-level data, if they communicate most effectively with the creative producers, are to be preferred over those which conform to standard practice for research reports. Broad, speculative interpretations of empirical results are typically more useful than interpretations limited to the more strict implications of a study. And, as indicated earlier, biased methods of sample selection and testing often can be employed to good advantage. However, in following these departures from standard research practice, there is a risk of producing misleading results. Accordingly, it is essential that resulting production recommendations be appropriately qualified.

Formative research, in the view taken here, is properly eclectic and pragmatic. In these respects, it is highly compatible with the current trend toward the very explicit definition of instructional objectives, followed by the development through systematic trial and

sciences in education or the usefulness of existing theory and knowledge. Rather, it holds that a useful step between basic research and educational practice is additional research of a formative sort, far more directly concerned with specific combinations of educational objectives, instructional media, learners, and learning situations. This is not to say that formative research is exclusively concerned with putting theory into practice. An equally valid function is that of starting with practice and transforming it into improved practice. Still another is that of providing hypotheses for further research and theoretical development. This is, incidentally, what is coming to be the dominant conception of the technology of education, not a commitment to teaching in the older audio-visual tradition, but a commitment to the achievement of a planned educational effect.

One long-standing point of view in education holds that theories and results growing out of the "mother" disciplines of psychology, sociology, anthropology, and the like, will filter into effective educational practice if enough educators have been trained in these basic disciplines. While this approach has been useful to a degree, it has not produced broadly satisfactory results. Meanwhile, creators of new educational products and practices have proceeded largely without the benefits of measurement and research. This is partly because skill and training in these areas have been linked to the process of theory construction and validation, and partly because of an inappropriately rigid adherence to traditional research practice within the product developmental context. Formative research procedure promises to help in creating a mutually constructive relationship between these two overly isolated realms -- the science and the technology of learning.

Footnotes

[1]Joan G. Cooney, "Report: The Potential Uses of Television in Pre-School Education" (The Carnegie Corporation, 1967).

[2]Michael Scriven, "The Methodology of Education," in Perspectives of Curriculum Evaluation, ed. R. W. Tyler, R. M. Gagne, and M. Scriven (AERA Monograph Series on Curriculum Evaluation; Chicago: Rand McNally, 1967).

[3]Lee J. Cronbach, "Course Improvement Through Education," Teachers College Record, 64 (1963), 672-83.

[4]J. Thomas Hastings, "Curriculum Evaluation: The Why of the Outcomes," Journal of Educational Measurement, Vol. III, No. 1 (Spring, 1966), 27-32.

[5]Scriven, op. cit.

[6]George L. Gropper and Arthur A. Lumsky, Studies of Televised Instruction (Metropolitan Pittsburgh Educational TV Stations WQED-WQEX and American Institute for Research, 1961).

[7]Walter Dick, "A Methodology for the Formative Evaluation of Instructional Materials," Journal of Educational Measurement, Vol. 5, No. 2 (Summer, 1968).

[8]Roger O. Scott and Morris F. Martin, "The 1969-70 Classroom Tryout of the SWRL Instructional Concepts Program," SWRL Technical Memorandum No. TM-3-70-4, June 9, 1970.

[9]Roger Scott, Jennifer Castrup, and Emily Ain, "The SWRL Kindergarten Art Program," Southwest Regional Laboratory Technical Memorandum, May 18, 1970.

[10]Samuel Ball, Gerry Ann Bogatz, "The First Year of 'Sesame Street': An Evaluation," (Princeton: Educational Testing Service, 1970).

[11]Ibid., and Gerry Ann Bogatz and Samuel Ball, "The Second Year of 'Sesame Street': A Continuing Evaluation," (Princeton: Educational Testing Service, 1971).

[12]Langbourne A. Rust, "Attributes of 'The Electric Company' That Influence Children's Attention to the Television Screen," (in press).

[13]For perhaps the best presently available discussion of facet analysis, see Dov Elizur, Adapting to Innovation (Jerusalem, Jerusalem Academic Press, 1970).

[14]E.g., see Rust, op. cit.

Footnotes, ii

[16]Keith Mielke and Jennings Bryant, Jr., "Formative Research in Comprehension of CTW Programs," (in press).

[17]Keith Mielke and Jennings Bryant, Jr., "Formative Research in Attention and Appeal," (in press).

[18]Kenneth O'Bryan, "Report on Children's Television Viewing Strategies," (in press).

[19]Gerald S. Lesser, Lessons from Sesame Street (in press)