

# Canadian Journal of Educational Communication

**Volume 15, Number 1  
Winter 1986**

## Editor

Robert M. Bernard

## Assistant Editor

Steven Shaw

## Editor, *Media News*

Cheryl Amundsen

## Editorial Assistants

Suzanne Daningburg

Ida Eva Zielinska

Carolyn Ross

Mary Geneva

## Column Editors

Al LeBlanc, Media Managers

Nancy L. Lane, Mediography

Richard Ellis, From the Media Periodicals

Len Proctor, Computer Microware

Suzanne Daningburg, Book Reviews

ISSN 0710-4340

The *Canadian Journal of Educational Communication* is published quarterly by the Association for Media and Technology in Education in Canada, 500 Victoria Road, North Guelph, Ontario N1E 6K2, Attention: Ron Eyre, Secretary/Treasurer. Notification of address change should be sent to the above. All articles are copyrighted by AMTEC and may be reproduced for non-profit use without permission provided credit is given to *CJEC*. *CJEC* is indexed in the *Canadian Education Index*.

Second class postage - pending

## EDITORIAL

*Robert M. Bernard*

## ARTICLES

Formative Evaluation of Instructional Materials: An Overview of Approaches  
*Cynthia B. Weston*

Interactive Television Brings University Classes to the Home and Workplace  
*Joyce Carver and Ruth C. MacKay* 19

Formative Evaluation of Educational Television  
*Jon Baggaley* 29

Formative Evaluation: The TVOntario Perspective  
*Patricia Parsons and Anne-Marie Lemire* 45

## COLUMNS

Media Managers 53

Microware Review 55

From the Media Periodicals 59

Mediography 62

Book Reviews 65

All correspondence should be addressed to:

Dr. Robert M. Bernard, Editor  
Education Department  
Concordia University  
1455 de Maisonneuve Blvd. W.  
Montreal, PQ, Canada H3G 1M8

# Formative Evaluation of Educational Television

**Jon Baggaley**

**Abstract:** Media and communication researchers have attempted to isolate the effects of television since its inception. They have dwelt at length on the negative contributions of television to society, and have raised numerous concerns about the potential of television to corrupt and manipulate. However, media research has contributed little to our understanding of how to use television as a reliable educational tool. As a result, there are few guidelines that we can give to the television producer who inquires about the techniques needed for "a good TV programme".

The following paper suggests some of the benefits of a "formative evaluation" approach to educational TV development. Examples are given of recent Canadian studies in this field, and of their theoretical and practical implications. The variables underlying audience responses to media materials are set in the context of attribution theory.

## DESIGNING "GOOD TELEVISION"

When asked about the types of insight they would like to gain from media research, television producers invariably have specific demands. They need to know about the impact of their programmes on a wide variety of audiences - adults and children, urban and rural. They need feedback on whether a programme fulfilled specific intentions, and whether or not particular production techniques were as useful as had been hoped. If a programme fails, its producer needs to know the reason it failed. He or she needs precise details about the types of performer to be used (or avoided) in future programming, and asks whether a programme will continue to be useful with repetition. Moreover, whenever they pose such questions, producers usually need the answers with speed!

The reasons for failure in a TV programme are often quite impossible to predict. On the one hand, the design team may have completely misjudged the capabilities and interests of its audience; on the other hand, a programme designed with the best theoretical intentions may be jeopardized by a momentary lapse in production skill. Camera angle, styles of editing - the ways that you say it as well as the things you say - are responsible for profound effects upon audience reactions.

**Jon Baggaley** is associate professor of educational technology at Concordia University. His research and teaching interests include the psychology of communication, audience research and educational TV and film research.

The author gratefully acknowledges the primary support of the Canadian Cancer Society to research reported in this paper.

This paper is based on addresses to the Advanced Seminar on Formative Evaluation, Human Sciences Research Council, Pretoria, November, 1984.

The most substantial body of research regarding TV production effects is found in the literature of educational technology (Baggaley & Duck, 1976; Coldevin, 1976). Coldevin reviews over 40 empirical studies dating since the early 1960's, concerning:

- (a) Presentation/technical variables (e.g., camera factors, setting, colour versus monochrome, still versus motion pictures, visual/audio reinforcement, and speed of presentation);
- (b) Content/subject matter organization (e.g., televised lecture, interview and discussion formats, audience reaction inserts, review strategies, direct explanation versus inserted questions); and
- (c) Performer characteristics (e.g., sex, age and appearance, dress, camera-eye contact).

Many of these studies indicate subtle effects of technique upon attitudes and learning that even an experienced TV practitioner would be unlikely to predict. However, as Coldevin notes, little consistency is to be found in the outcome of the various studies, since few of them paid adequate attention to the control of the experimental variables, and/or were followed up by further research.

Considering the immense array of production techniques offered by television, it is not surprising that research has been slow to shed light on their audience effects. In fact, we may wonder whether conventional research methods will ever be precise enough to predict the massive number of effects that TV techniques combine to produce. The traditional empirical approach to media production research has involved, as in the psychological laboratory situation, the testing of specific hypotheses. On this basis, deductions are made concerning the variables responsible for programme impact. If the factors underlying a programme's success or failure can be clearly defined, a *hypothetico-deductive* approach to its study can yield useful results. However, the normal TV production is too complex to support a single set of hypotheses, and a different research approach is clearly required.

In the last ten years a type of study has emerged which is aimed less at the isolation of production effects on hypothetical bases than upon their inspection in actual programme contexts (Baggaley, 1979-1980, 1984). This approach may be characterized as *inductive*, by which specific production guidelines are inferred from general observation. The approach derives in large measure from the fruitful relationship between producers and researchers at the Children's Television Workshop in New York during the 1970s (Dennis, 1979; Mielke & Chen, 1981). The design of CTW programmes such as "Sesame Street" and "3-2-1 Contact" was then, as now, the focus for intensive research and evaluation. Methods were devised for supplying producers with evidence regarding a programme's impact in time for modifications to be made as appropriate.

The emphasis here was upon effective formative evaluation (Scriven, 1967) which aims to monitor and to recommend modifications to the impact of a production during its formative process. The approach contrasts with more traditional "summative" styles of evaluation, which are conducted when production is completed and modifications are no longer possible. The objectives of formative evaluation are usually more pragmatic than those of summative studies, although they are also more limited than the objectives associated with full-blown research studies. In Scriven's words, research studies usually aim to prove something, while formative evaluation is content to improve something. Evaluation studies commonly use the same techniques as research studies, and they are a valid form of research activity; but their aims and accomplishments are usually more specific.

In addition to providing rapid feedback to the designer of a product, an effective formative evaluation must also give precision of feedback. Traditional evaluation methods, involving pretest and posttest procedures (Borich, 1981; Dick & Carey, 1978 ) are often

incapable of the precision required by TV producers. In order to measure the precise effects of TV production technique, methods for monitoring a programme's moment-by-moment impact upon its audience are needed. Attempts to develop technologies for time-based audience reaction measurement date back to the 1930s (Cambre, 1981); and with the advent of the microcomputer in the 1980s, such facilities have gained immensely in speed, precision, portability and general practicality.

One such system - the Program Evaluation Analysis Computer (PEAC) - records fluctuations in programme impact as fast as every quarter-second. It employs a set of push-button hand-units to record the responses of individual audience members on measures such as interest value and credibility (Nickerson, 1979, 1981). Developed as a collaborative venture of the Children's TV Workshop and research staff at the Ontario Educational Communications Authority, the PEAC system is currently finding a particular niche within the media advertising industry (The Program Evaluation Analysis Computer (PEAC) is a product of PEAC Media Inc., Toronto.). To the producer of a TV commercial, second-by-second evidence of appeal and persuasiveness can obviously be invaluable. It can be equally useful, for that matter, to the designer of a political campaign, in determining the detailed impact of campaign strategies or candidates. The scope for abuses of the new methodology is thus only too clear; and it is to be hoped that educational broadcasters will come to use it with the same enthusiasm as their advertising and propagandist colleagues.

At Concordia University in Montreal, and at the Human Sciences Research Council, Pretoria, programmes of formative research and evaluation using the PEAC system are currently underway. The present paper, based on the Montreal experience, indicates various educational applications of the new formative evaluation methods. Separate applications are discussed in areas of (a) needs assessment, (b) product development, (c) product utilization, and (d) process development. Studies conducted in each of these areas may have immediate application within specific media projects, and more long-term implications for media research. Ultimately, the potential of the new methods extends beyond the media field altogether, to the study of human communication processes previously incapable of measurement. With these broader implications of formative evaluation methodology in mind, a theoretical perspective is now suggested for the analysis of communication effects.

## TELEVISION AND ATTRIBUTIONS

Television is neither good nor bad. The same techniques may be employed to teach via television as to manipulate and persuade. As with any medium, the morality of television depends upon the intent of its users (Jamieson, 1985). The value of television's effects, moreover, can differ widely from one individual to the next; and with the same individual they can vary across time. We are reminded of these factors by attribution theory, which examines the desire to attribute causes to the events and phenomena encountered, and the bases on which a person's attributions may be predicted (Heider, 1958; Kelley, 1967).

To understand the effects of television, attribution theory warns, we must examine:

- (a) Individual differences, due to psychological and social factors, between the responses of audience members;
- (b) Differences over time in the responses of individual viewers;
- (c) Technical effects: that is, differences in audience response due to the manner in which televised material is mediated; and
- (d) Parallel influences upon audience responses by the wide range of alternative factors thought capable of such effects (Baggaley, 1980, pp. 110-11, 162-5).

To mount a comprehensive study of television's impact, we must ideally be capable of controlling or at least recording each of these complex sources of variance simultaneously.

While comparison between individuals - (a) above - is a relatively simple matter, the measurement of all parallel, alternative influences, (d), may be totally impossible. Using traditional methods of programme evaluation, the influence of outside factors upon audience reactions can at least be kept to a minimum. A baseline reading of audience attitudes or ability may be obtained; a TV "treatment" is presented; and post-test measures are then administered in order to gauge aspects of the programme's immediate impact.

The major problem occurs in attempting to measure long-term effects of the programme (b) and (d), for the extraneous influences that may operate in the long term are too numerous to conceive. The necessary controls for long-term effects are often cumbersome and expensive (Glass, Willson & Gottman, 1972; Ostrom, 1978), and they may preclude the pragmatic approach to media research required in most broadcasting contexts. Within the context of a single programme or programme series, however, insights into audience effects can now be obtained with a new sophistication. The modern response analysis systems allow for precise measurement of audience responses during the time-span of the programme itself - (b) above. The second-by-second precision of this measurement also allows the analyst to identify effects due to particular variations in production technique (c).

The new opportunities provided by time-based response analysis are indicated by the following case study.

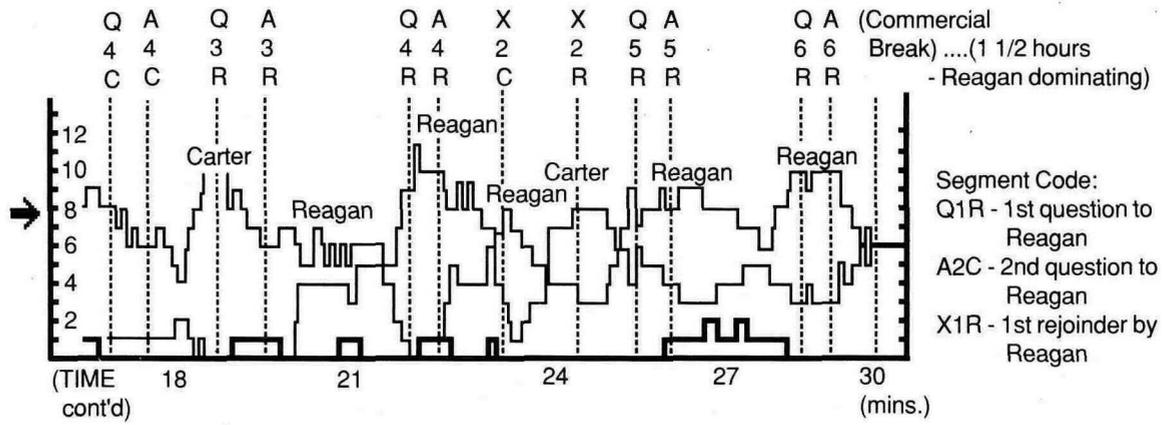
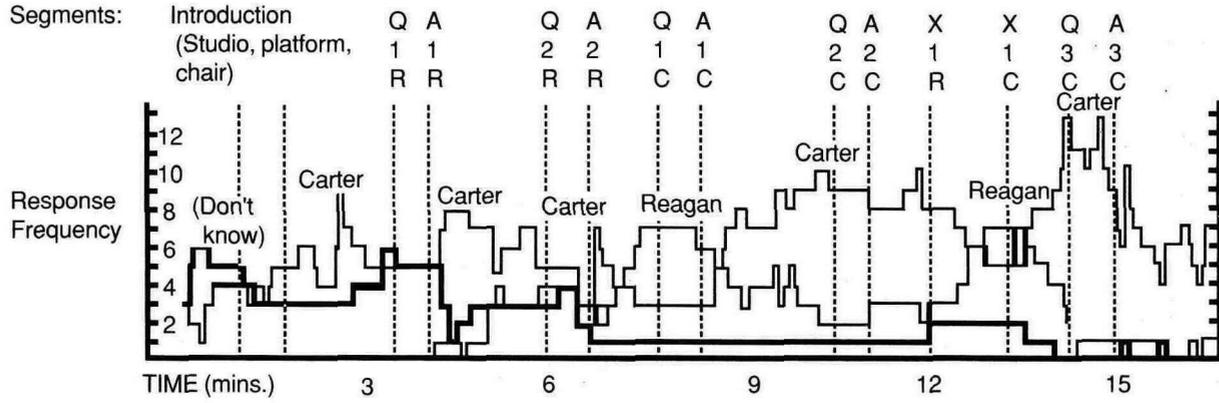
#### CASE STUDY #1: THE U.S. PRESIDENTIAL DEBATES

In November 1980, American President Jimmy Carter and Presidential challenger Ronald Reagan engaged in a 90-minute debate televised live throughout North America. In St. John's, Newfoundland, a panel of two dozen adult viewers gave their reactions to the debate, second by second, via the hand-units of the Program Evaluation Analysis Computer system owned by Memorial University of Newfoundland (Baggaley, 1985a). The viewers each watched the debate in their own homes, to which the response units were delivered in advance. The study was commissioned by the Canadian Broadcasting Corporation as material for a radio report on the debate the following morning.

The units were programmed to sample responses to the following question: "Who, from one moment to the next, is winning the most votes?" Three responses were available to the question, on buttons labelled CARTER, REAGAN, and DON'T KNOW. Audience reactions to the debate were sampled, on this basis, at 4-second intervals. In Figure 1 (See next page), moment-by-moment fluctuations in the perceived success of the two candidates are plotted for the first 30 minutes of the debate.

As Figure 1 indicates, the candidate perceived as winning the most votes at the beginning of the debate was Jimmy Carter. As the incumbent President, he was perceived as winning even before he first spoke. In the 19th minute, however, audience responses began to favour Ronald Reagan. On synchronizing the animated computer display of these results with a videotape of the debate, this shift in viewer response was found to coincide with verbal points made by Reagan concerning the Carter administration's economic record. A comparison of individual responses to the debate indicated that the shift to Reagan was due primarily to male viewers in the sample. Thirty seconds later, the same shift was observed in the responses of the female viewers; this effect coincided with a repetition by Reagan of the same verbal argument. The internal validity of these conclusions is demonstrated by statistical procedures beyond the scope of the present paper.

The Carter - Reagan debate (November, 1980): First half-hour



(Audience question: "Who, from one moment to the next, is winning the most votes?")

Segment Code:  
 Q1R - 1st question to Reagan  
 A2C - 2nd question to Reagan  
 X1R - 1st rejoinder by Reagan

The external validity of such results is, of course, debatable. A panel of viewers randomly selected from the population of Newfoundland, Canada, can hardly be regarded as representative of the voting population of the United States. The panel's reaction could in no way be assumed to predict the outcome of the Presidential election one week later. On the other hand, specific conclusions of the Newfoundland study were identical to those later put forward by Wingerson (1982). A tendency was noted for the Newfoundland panel to give a response against Carter at moments when the camera presented him from a particular side-shot. It was concluded that Carter appeared more tense from this camera angle than he did from others; and the conclusion was reported on CBC-Radio the morning after the debate. Eighteen months later, Wingerson's discussion of the debate was published, crediting the visual evidence of Carter's tension with being a major reason for his defeat.

The Newfoundland study indicates the greater speed as well as precision of a moment-by-moment analysis in comparison with conventional measurement techniques. If it had been replicated simultaneously with a contrasting range of American voters, there is no reason to doubt the predictive value of its conclusions. Four years later, the 1984 TV debates between Presidential contenders Reagan and Mondale have been subjected by American polling organizations to exit-surveys and other analytical procedures far more suspect than the methodology reported here. A further set of electronic analyses of the 1984 debates has been conducted by Concordia University researchers in Los Angeles and New York State, lending support to this view.

The information derived from such studies has obvious value for TV production staff. Since it was analyzed and reported after the Presidential debate's completion, the current study is essentially summative; but it has distinctly formative implications for the designers of future debates. For the television producer, the study indicates camera techniques that may bias audience reactions, and which should therefore be avoided. For the politicians and their advisors, such information can suggest strategic manoeuvres both verbal and visual. The findings can thus be used in the planning of new TV productions, as well as the post-mortem stage of the existing production.

The same methodology may equally be applied in educational broadcasting to investigate the influence of presentation techniques before, during, and after the production process. Examples of these applications are given in the next study.

## CAST STUDY #2: THE IMPACT OF SMOKING PREVENTION FILMS

In 1981, the Canadian Cancer Society (CCS) commissioned a summative evaluation of four films commonly used to inform the public about the dangers of smoking. The Society was particularly concerned to determine the films' impact upon under-educated (or functionally illiterate) viewers, representing a quarter of the Canadian population and particularly susceptible to lung and other cancers. Accordingly, the films were shown to a contrasting sample of viewers in urban and rural communities of Newfoundland and Labrador. The viewers' reactions were obtained by a range of conventional pre-test and post-test procedures, and via the time-based procedures of the Programme Evaluation Analysis Computer (a 4-point scale of approval across time was used, ranging from GOOD to POOR). The effects of prior attitudes to smoking and cancer upon perceptions of the films were assessed, also the moment-by-moment impact of the films upon prior attitudes. Finally, production techniques were recommended for future productions aimed at audiences varying in age, education, and sex.

Reported to the CCS by Baggaley (1982a), particular conclusions of the study were as follows:

- (a) Male and rural members of the sample have significantly less active concern for cancer prevention than female and urban groups.
- (b) Adult illiterate persons (reading grade 8 and lower) gain little or no benefit from the films, nor from the print materials accompanying them.
- (c) Adult illiterate persons have the highest incidences of smoking, cancer in the family and general pessimism about cancer issues.
- (d) Schoolboys (15-17 years old) exhibit unusually high incidences of smoking/pessimism also.
- (e) Television is the medium with the strongest potential for reaching these audiences, although radio may also be useful for communicating with housewives.
- (f) Smoking prevention films in current use serve to reinforce non-smokers in their distaste for smoking, though are generally received by smokers with defensiveness and hostility; the films under test were thus considered more useful for education about the prevention rather than the cessation of smoking, although in both connections their presentation should be accompanied by careful group discussion
- (g) Smokers are willing to consider practical guidelines for smoking cessation, being inclined to respond most positively to the films during segments when practical "quitting tips" are given; however, they respond negatively to any suggestion by the films (usually visual) that quitting will make them appear socially eccentric, and they require evidence that the effort to quit will be worthwhile.
- (h) A successful formula for films encouraging smoking cessation would be to increase viewers' pessimism regarding the dangers of smoking, while reducing pessimism regarding cancer prevention and cure.

As in the earlier case study, these conclusions have formative as well as summative uses. Their summative implications concern the distribution and utilization of the films. The conclusions regarding audience knowledge and opinion, and about the effects of specific production techniques, can be of value at the needs assessment and production planning stages of future health campaigns.

The benefits of a formative evaluation approach are thus seen from the earliest to the final stages of programme development. These possibilities have been described by previous writers in terms similar to those used here. For example, Sanders & Cunningham (1973) have identified four types of formative evaluation, as follows:

- (a) Predevelopmental Activities - audience needs assessment and other evaluation procedures occurring prior to actual product development;
- (b) Evaluation of Objectives - assessment of the formal goals and objectives defined by the product developer;
- (c) Formative Interim Evaluation Activities - assessment of the product at its early stages of development;  
and
- (d) Formative Product Evaluation Activities - assessment of the product at its final draft stage.

Defined in this way, the very distinction between formative and summative evaluation conceived by Scriven (1967) begins to blur. It becomes apparent that evaluation activities can, or at very least should, aim to generate some formative recommendations before and after production as well as during it. This contention would certainly be supported by media producers, for whom purely summative evaluations have little or no applied value. Since summative evaluation customarily serves to expose the weaknesses of a product when it is

too late for improvements to be made, producers can also regard it as threatening. The advent of today's microcomputer-based techniques, with a speed and precision quite unforeseen by writers on this topic ten years ago, makes it even more possible for all evaluation activities to have a formative outcome.

In one sense, it would be desirable if the distinction between formative and summative types of evaluation were now disbanded. The distinction has served a useful function in pointing out the need for formative conclusions at a time - the late 60s and 70s - when evaluation studies were almost exclusively summative. However, evaluators of educational TV products are now capable of fulfilling a formative role whether the product is technically completed or not. They should aim to fulfil this role automatically, in order to be maximally useful to production personnel and to avoid having a threatening effect upon them. Without a relationship of mutual support and trust, the validity of an evaluation study is jeopardized in any case, for effective collaboration between a media producer and researcher becomes all but impossible.

The final case study describes a collaboration between film production personnel and an independent team of formative evaluators in the planning and development of a film on skin cancer.

### CASE STUDY #3: FORMATIVE DEVELOPMENT OF A SKIN CANCER FILM

In 1983, the Canadian Cancer Society (CCS) commissioned a formative evaluation study leading to the production of a new film on skin cancer prevention (Baggaley, 1985b). The study was prompted by the findings of Case Study #2 (above) - that male and rural people show little concern for conventional cancer education films compared with female and urban groups. This observation provides cause for some concern, for rural males are known to have an exceptionally high incidence of skin cancer, largely in view of their greater exposure to the sun.

The only existing film in use by the CCS was already fifteen years old, and the facts it contained were now out-of-date. A new film was required whose primary audience would be the rural male; a secondary target audience would be those individuals - urban as well as rural - who risk over-exposure to the sun in the course of leisure activities (e.g., sunbathing, gardening, skiing). The need to design a film for such diverse target groups presented a considerable challenge, and a formative research and evaluation plan was devised running parallel to but independent of the film's production schedule.

The research was in four stages, broadly equivalent to the four types of formative evaluation recommended above (Sanders & Cunningham, 1973).

#### *Needs assessment*

Prior to the planning and scripting of the film, a summative evaluation was conducted of the existing film ("Sense in the Sun", produced by the American Cancer Society in 1968). Two-hundred and fifty people from the provinces of Newfoundland and Quebec gave their reactions to the film on the PEAC system's electronic hand-units and on pretests and posttests, as in the second case-study. The sample was divided according to geographical location (urban versus rural), age, sex, and education. Urban audiences found the film old-fashioned and insufficiently informative, while rural people enjoyed its location in a coastal fishing environment similar to their own. Rural males, however, tended to react suspiciously to the film's message about the dangers of exposure to the sun, and to disagree that people who work out-of-doors should protect themselves from the sun. The second-by-second reactions (on a 4-point scale of approval from GOOD to POOR) indicated the need to

present material to the urban and rural groups at different rates. Once again the particular defensiveness of male and rural viewers to the cancer topic was observed.

Based on these and other findings, an outline - or storyboard - was prepared for the new film. In order to be acceptable to lesser educated viewers as well as those with high-school education or above, the CCS decided not to include detailed information in the film, but to aim for a product which would motivate people to seek further information from other sources. Findings of the research concerning production technique, however, were applied in detail. It was decided, for example, that the new film should cater to the urban and rural audiences in a series of alternating segments. The types of information required by the urban audience would be conveyed by segments using speedier cutting rates and an urbane story treatment. The segments for rural viewers would be motivational rather than fact-packed, filmed in a rural context and presented at a gentler rate. The effort to maintain the interest of each type of audience during segments intended for the other, would be a challenge requiring careful monitoring.

The needs assessment stage of the study, featuring the evaluation of an existing film, provides a good example of a summative evaluation with formative implications. For want of a shorter term, such an evaluation may be described as quasi-formative.

#### *Evaluation of the Film Concept*

The objectives and storyboard of the new film were now evaluated in discussions with a further 107 people from the same urban and rural areas. Subjects favored a storyline approach using a logical rather than emotional approach. They stressed the need to feature real people in the film rather than actors, and they suggested a wide range of characters with whom they could identify. They requested explicit details of the symptoms and effects of skin cancer, and of the types of person most susceptible to it. They remained unconvinced that a Canadian audience would be particularly anxious about the dangers of sun exposure, and a need was indicated to stress facts concerning the sun's effects through cloud. The design of pamphlets and posters to accompany the film was also discussed.

The storyboard evaluation ascertained that the findings of the earlier needs assessment had been correctly interpreted by the production and scripting team, and generated new ideas for specific script additions and changes.

#### *Formative Evaluation of the Film Roughcut*

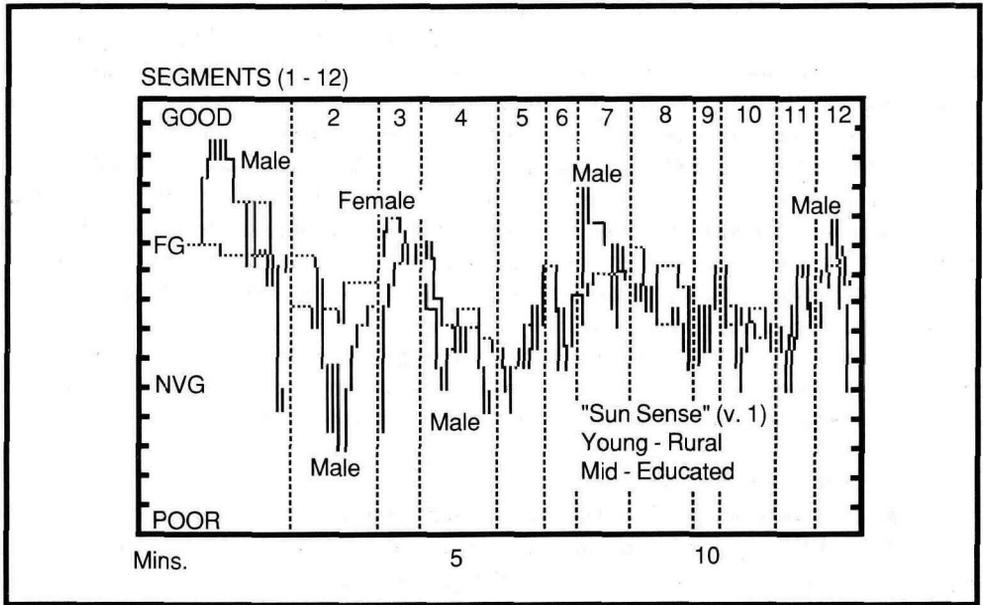
In June 1984, a 13-minute roughcut of the new film was tested upon a further 128 urban and rural viewers from Newfoundland and Quebec. Responses to the film were collected via pre- and post-tests, and on a second-by-second basis as in the needs assessment. It was found that the film was creating an appropriate set of general attitudes about skin cancer, though was not as yet convincing viewers of the need to take preventive measures. Specific moments in the film were strongly disapproved by urban and rural viewers alike. The second-by-second responses of male and female viewers to the film (18-25 years old, rural, high-school educated) are presented in Figure 2 (See next page).

The results of the formative evaluation were reported to the film's producer and sponsor in July 1984. The producer responded by shortening various segments, by changing the voiceover to strengthen character identification, adding montage and graphic sequences to emphasize some of the educational points, repeating certain sequences in order to create relief, and adding music to heighten mood and structure. He also prepared draft versions of a printed leaflet reinforcing the film's main points.

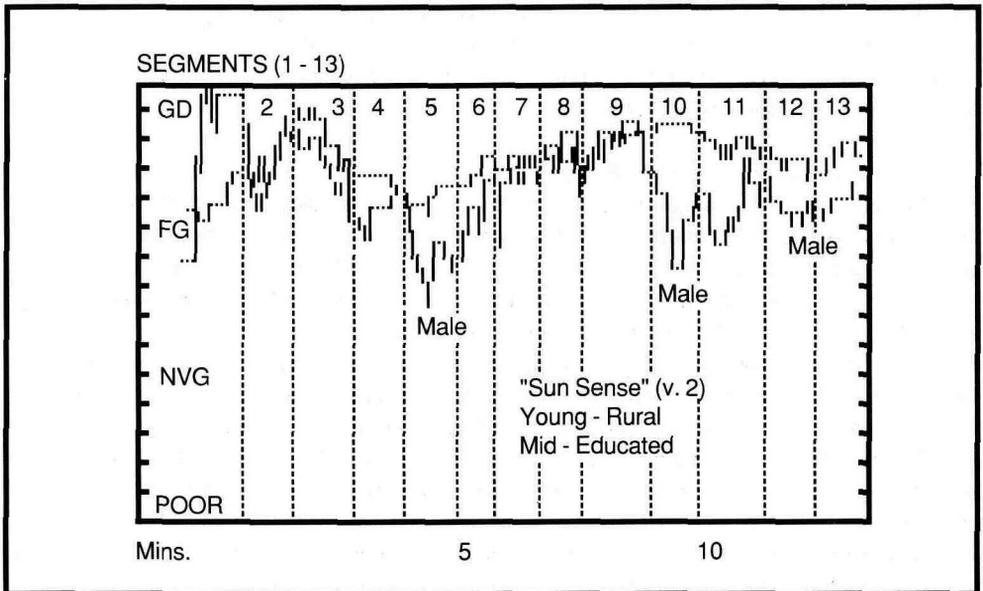
#### *Summative Evaluation of the Modified Film*

The final evaluation of the film was conducted in Newfoundland and Quebec during

FIGURE 2.  
Audience Reactions to New Film (pilot version).



FIGURES.  
Audience Reactions to Modified Version.



Note: Both Figures feature the responses of a representative subgroup of viewers: rural, 18 - 25 years old, with highschool education; n (pilot version) = 18; n (modified version) = 19.

October 1984 and January 1985, upon a further 194 viewers. During the same sessions the supporting leaflet was evaluated, and was found appropriate for audiences with a reading level of Grade 5 and above. The need was indicated for modifications to the leaflet's layout.

Overall reactions to the film were now positive. It was regarded as an effective means of persuading people to take preventive measures against the sun. Even the rural male viewers approved of the film, which seemed to have overcome their earlier reluctance to regard exposure to the sun as potentially dangerous. Second-by-second responses to the film by male and female viewers (18-25 years old, rural, high-school educated) are presented in Figure 3 (See previous page).

A comparison between Figures 2 and 3 reveals the effects of the modification made to the film between its roughcut and final stages. Reactions to the four segments eliciting low approval at the roughcut state (Figure 2) are now significantly improved (Figure 3). The relatively negative response to the first of these segments - set in a doctor's surgery - is eliminated altogether.

At each stage of a formative evaluation - including the final one - it is always possible that further modifications may be made in the attempt to attain perfection. The three segments in the skin cancer film which continue, at this summative stage, to elicit relatively negative responses, might certainly be changed on this basis or cut out altogether. It is easy to detect a pattern in the audience responses to these segments, for all three feature the same character (a young lady artist) and setting (an artist's studio). The lower rates of approval for the segments seem due to the lack of specific information within them, and to an attempt by the script to develop human interest.

If further evaluations are not contemplated, however, it is important that further modifications at this stage are made with caution. It is quite possible, for example, that the removal of these segments could upset the balance and pacing of the film, so carefully calculated on the basis of the earlier research. Post-test reactions to the artist character are very positive; and in view of the approving audience reactions to this version of the film in general, the relative weakness of the three studio segments may be regarded as of little consequence.

During this final evaluation of the skin cancer film, however, one audience subgroup remained more negative than the others. Highly educated urban viewers continued to disapprove of certain segments in the film, and of its general pacing and lack of detail. The evaluators recommended that in the future use of this film with urban groups, accompanying print materials or discussion should be designed for different target groups, providing an inexpensive way of maximizing the film's utility.

It should be noted, of course, that the urban, educated population was not the film's primary target audience, and that the crucial evidence for the educational value of the film was provided by the rural viewers. Many of the production features criticized by urban viewers - including repetition of segments and design of characters - were based quite deliberately on the earlier feedback from the rural viewers. It is possible that urban viewers, however highly educated, are less expert in the matter of what makes a good TV programme for a rural audience than are the rural people themselves. Yet most instructional films are made in an urban, middle-class environment, based on urban, middle-class values. It is clearly important to provide viewers from other environments with the opportunity for input into the programmes which concern them. The role of formative evaluation is to create this opportunity.

FORMATIVE EVALUATION: THEORY AND PRACTICE

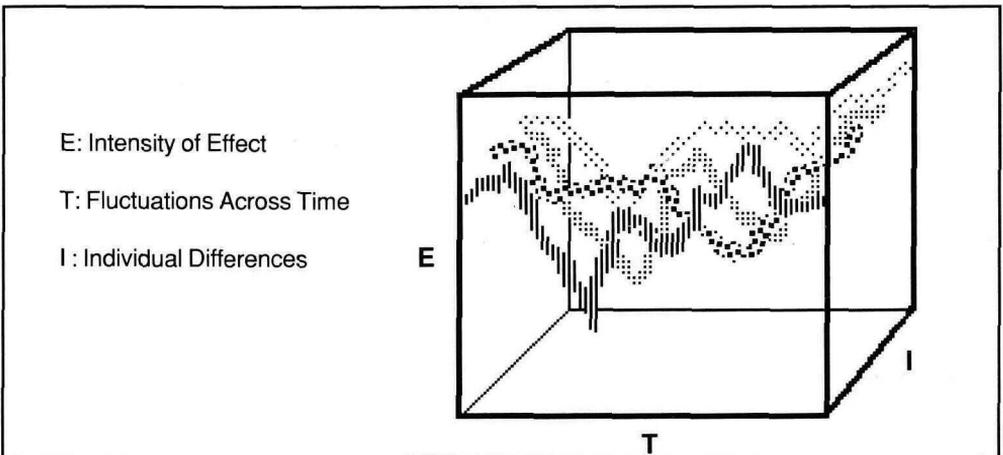
As an increasing range of media productions is evaluated by these methods, our knowledge of the communication processes underlying educational TV production will increase dramatically. From the study of the smoking prevention films in the second case study, it was possible to generate guidelines for the production of the skin cancer film in the third case study. From the evaluations conducted in that context, information has been derived which can be applied in film-making on other aspects of health education. On this basis, we should ultimately be in a position to define the techniques required for effective television on any topic, and for all types of audience. For we will have established - inductively - the impact of all production methods common to the conventions of the day. We will always, of course, be limited in our ability to predict unconventional production techniques. The creative freedom of the media producer will therefore remain intact; indeed, his or her artistry may even be stimulated by the insights gained into audience psychology.

The ability of research to penetrate the effects of media communication is seen to have been increased substantially by the development of new techniques for electronic response measurement. It is now possible to record, and to make speedy comparisons of the effects of a TV programme upon different individuals; simultaneously one may inspect second-by-second variations in these effects due to presentation technique. Three independent dimensions of measurement may therefore be controlled, relating to:

- (a) the intensity of the effect under scrutiny (e.g., cognitive, affective, or behavioural);
- (b) fluctuations in the effect across time; and
- (c) individual differences in the effect (due to demographic, geographic, or psychographic factors).

When data are available on all three of these dimensions at once, we have the means to predict and explain the effects of communication according to attribution theory (see earlier section). The interaction of the three measurement dimensions accounting for media effects is summarized in Figure 4.

FIGURE 4.  
*Interaction of Dimensions Accounting for Audience Attributions to Television.*

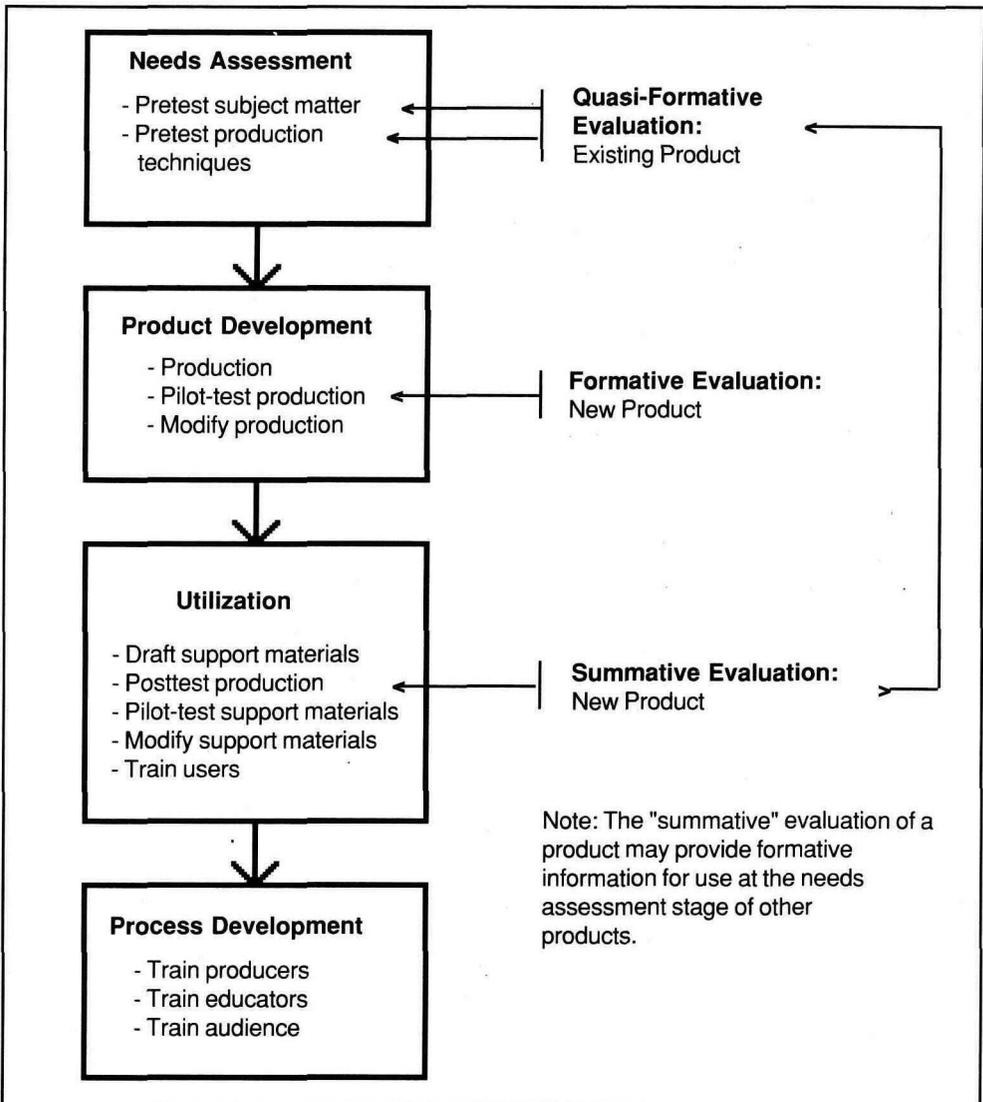


The precision of the new forms of data, and the speed with which they can be analyzed, have already been found to have immediate benefits for the media audience itself (Baggaley, 1982b). By providing a group with fast feedback concerning its reactions to media presentations, it can be sensitized a) to ways in which the media can manipulate public opinions, and b) to media techniques that the group may use for its own purposes. Baggaley & Smith (1982), for instance, have reported the use of formative research methods in a rural development context - a project which has since led to a fuller understanding by fishermen of social and economic problems facing them, and to the correction of some of these problems via the mass media. Implications of formative evaluation methods for "process" as well as "product" research are thus indicated, and for the use of media in a wide range of social situations.

With careful advance planning, a formative evaluation approach to the development and use of media materials can proceed as in Figure 5. A single project may use formative

#### FIGURES.

*Stages in the Development and Formative Evaluation of Media Materials.*



methods for one or more of four general reasons: needs assesment, product development and utilization, and process development. Of course, the true test of a methodology is that it should be able to identify phenomena that would not have been so readily apparent otherwise. There is little doubt that formative evaluation methods will be rapidly seized upon by the political and commercial users of media; the benefits of formative evaluation in these contexts may be inferred from the case study of the Presidential Debates. It is hoped that the new methods will also be used in educational television, leading to an increased awareness of the public's needs, and to an understanding of the differences and similarities between cultural groups. On this basis we may learn to use the medium in society with a greater accountability and effect.

#### REFERENCES

- Baggaley, J. P. (Ed.). (1979-1980). *Experimental research in televised instruction, Vols. . 2-3*. St. John's: Memorial University.
- Baggaley, J. P. (1980). *Psychology of the TV image*. Farnborough: Gower.
- Baggaley, J. P. (1982a). *Cancer prevention and literacy*. (Report to the Canadian Cancer Society). St. John's: Memorial University.
- Baggaley, J. P. (1982b). Electronic analysis of communication. *Media in Education & Development, 15*, 70-3.
- Baggaley, J. P. (Ed.). (1984). *Experimental research in televised instruction, Vol. 4-5*. Montreal: Concordia University.
- Baggaley, J. P. (1985a). The analysis of media impact. In B. Alloway & G. Mills (Eds.) *Aspects of Educational Technology, Vol. 18*. London: Kogan Page.
- Baggaley, J. P. (1985b). *Formative evaluation and development of a film on skin cancer*. (Report to the Canadian Cancer Society). St. John's: Memorial University.
- Baggaley, J. P., & Duck, S. W. (1976). *Dynamics of television*. Farnborough: Saxon Hous.
- Baggaley, J. P., & Smith, K. (1982). Formative research in rural education. *Media in Education and Development, 15*, 173-6.
- Borich, G. D. (Ed.). (1981). *Evaluating educational programs and products*. Englewood Cliffs, NJ: Educational Technology Publications.
- Cambre, M. (1981). Historical overview of formative evaluation of instructional media products. *Educational Communication & Technology Journal, 29*, 3-25.
- Coldevin, G. O. (1976) Comparative effectiveness of ETV presentation variables. *Journal of Educational Television 2*, 87-93.
- Dennis, R. (1979). Research on "Sesame Street": Designing the educational context. In J. Baggaley (Ed.), *Experimental research in TV instruction, Vol. 2*, St. John's: Memorial University of Newfoundland.
- Dick, W., & Gary, L. (1978). *The systematic design of istruction*. Glenview, IL: Scott, Foresman & Company.
- Glass, G., Willson, V., & Gottman, J. (1972). *Design and analysis of time series experiments*. Boulder: University of Colorado.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Jamieson, G. H. (1985). *Communication and persuasion*. Dover, N.H.: Croom Helm.
- Kelley, H. H. (1967). Attribution theory. In D. Levine (Ed.), *Nebraska symposium on motivation*. Lincoln: NE, University Press.
- Mielke, K., & Chen, M. (1981). Making contact: Formative research in touch with children. In J. Baggaley (Ed.), *Experimental research in TV instruction, Vol. 4*. Montreal: Concordia University.

- Nickerson, R. (1979). Formative evaluation of instructional TV programming using the Program Evaluation Analysis Computer. In J. Baggaley (Ed.), *Experimental research in TV instruction, Vol. 2.* , St. John's: Memorial University.
- Nickerson, R. (1981). Second generation electronic testing technology applied to TV evaluation. In J. Baggaley (Ed.), *Experimental research in TV instruction, Vol. 4.* Montreal: Concordia University.
- Ostrom, C. W. (1978). *Time series analysis.* Beverly Hills: Sage.
- Sanders, J., & Cunningham, D. (1973). A structure for formative evaluation in product development. *Review of Educational Research* , 43, 217-36.
- Scriven, M. (1967). The methodology of evaluation. In B. Worthen, & J. R. Sanders (Eds.), *Educational evaluation: theory and practice.* Worthington, OH: Charles A. Jones.
- Wingerson, L. (1982). Nice guys finish first. *Discover*, March, 66-7.