

Formative evaluation of educational children's programs: A comparison of Australian and American models

Marianne Treuen
Queensland University of Technology

This paper examines the trial and pilot stage of formative evaluation, concentrating on the two programs "Sesame Street " and "Finders Seekers ". For each trial it investigates the setting, subjects, methods, aims and purpose of the trial as well as which sections of the programs were selected and how responses by children and adults were reported and recorded.

While Australian producers have limited budgets, the following observations are relevant to Australian trials. Settings for trials should be varied. Trials should reflect the differences among subjects. While testing segments may be useful, pilots should be ready to use in the trial process. Observation as a method needs careful analysis. Reporting needs to take into account the literacy skills of children and the editing done by evaluators.

Introduction

This paper discusses a method used to evaluate success of educational children's programs. The method concentrates on formative evaluation or "research into audience responses before the program is finalised" (Bates, 1984: 198), more particularly the trial and pilot stage of formative evaluation of two programs.

"Sesame Street" is one of the best known and most widely researched programs and needs no introduction (Lystad, 1989:21). "Finders Seekers" is a junior science program made by the Australian Broadcasting Corporation (ABC) in Brisbane (Thrupp, 1989). Formative evaluation for this program was the subject of a study I carried out with the help of its producer, D. Parmiter, and the Queensland Liaison Officer, A. James.

In this paper an Australian trial will be described with the aim of delineating a model of the trial phases of new programs. The original

working model of a trial was developed by the Children's Television Workshop (CTW) in the late 1960s for "Sesame Street". This paper will investigate the trials of "Sesame Street" and an ABC educational television series, "Finders Seekers".

Interactions between schools and television have varied (Lusted, 1985 and Schwager, 1991) and in Australia the survival of ABC educational programs has been questioned (Dawson, 1988 and Ramsay, 1988:44). In spite of the changes, while producing fewer programs, the ABC is still broadcasting educational programs in its daytime slot.

In view of recent concerns regarding the role of children's television and the scarcity of resources, the process of developing and trialing a new program assumes importance. The particular methods used during this stage of production need to be examined.

Different aims and learning theories do have implications for trials. The aims of "Finders Seekers" and "Sesame Street" are not identical. Learning theories related to educational programs are still being researched (Clark, 1990, Choat, 1988, and Choat et al, 1986 and 1987, Kozma, 1986, Rendall and Nixon, 1980). If the design of an educational program has been based on a particular learning theory, the evaluation of the program would need to account for this.

Defining educational programs

There are a number of possible definitions of educational programs (Cambre, 1981:22, Gilmour, 1979). For the purpose of this paper I wish to define educational programs very broadly as programs useful for education which may or may not be made specifically for schools.

Formative evaluation

Formative evaluation aims at feedback and modification during the development of the program, as distinct from summative evaluation carried out when the program is complete (Cook et al, 1975:46). A trial is just one of a number of facets of formative evaluation. Usually trials are considered ephemeral, short term feedback for the producer so the Children's Television Workshop's published work on trials is important.

Evaluation undertaken by government bodies concerning the educational implications of television, is usually summative in that the study is usually conducted at the conclusion of the program. Some examples of evaluative studies are the Senate Standing Committee on Education and the Arts Reports (1978 and 1981), the National Survey of Primary School Teacher's use of ABC Programs (Black and Palmer, 1987 and 1989) and the Evaluation of ABC Primary Schools Broadcasts (Queensland Department

of Education, 1980). For the Children's Television Workshop a national study was undertaken by the Educational Testing Service (Ball and Bogatz, 1970).

Although trials differ and some factors in a trial are hard to control, such as the effect of a committed team and the "halo effect" of a new program, the methodology needs to be designed to take account of the specific features of a particular program in order to draw some conclusions from the trial.

Data collected for formative evaluation usually includes data on "program appeal and audience attention, recall, comprehension and cognitive gains" (Cambre, 1981:21). Eidlitz has also considered affective or non cognitive knowledge (Eidlitz, 1979). Most of the data which arises from the trial stage of formative research could be classified as relevant to interest holding, comprehension or learning.

The children's television workshop trial

The Children's Television Workshop model of formative research had a number of stages. The first was to test audience interests. The audience tested was usually children but in variations of the model it could be adults. This was done by breaking the program up into segments in a laboratory setting which allowed observers to rate the distraction possible in each 7.5 seconds of a child's viewing (Cook et al, 1975:48). The segments were also used for simple "before and after" tests designed to test the recall and learning and test how "films could be improved as teaching instruments"(Cook et al, 1975:48).

In the second stage a pilot whole program is shown. Pilot shows aim to test the ability of the program to entertain as well as testing comprehension and cognitive and affective learning through pre-testing and post-testing.

The generally accepted model for formative evaluation is that of the Children's Television Workshop originally used for "Sesame Street". A short description follows but the model will be described later in more detail. In this model during the preliminary production phase researchers assess the comprehensibility and appeal of segments to viewer populations and - in some cases - test their acceptability to teachers. The in-house researchers measure the degree of concentration the children exhibit toward the filmed segments, rejecting or modifying those segments that fail to hold attention. Tests of comprehension are also given and segments that do not cause short-term comprehension are also rejected or modified. The pilot shows are tested to estimate audience appeal and comprehensibility in less fragmented and somewhat more mundane contexts than the research laboratory (Cook and Curtin, 1985:94).

People using this model in other situations may raise some of the following questions. What is the budget available? Can dramatic sequences be tested (and reshot if necessary) as easily as segments? Can motivation be measured as easily as liking and comprehension? (Cook and Curtin, 1985:109). A short comparison of "Sesame Street" and "Finders Seekers" illustrates the importance of these questions.

In Australia questions concerning budget and methodology arise as the "Sesame Street" trial had a large budget allowing for a team of formative (and other) researchers which may not be available to the ABC or other institutions. "Sesame Street" is made in segments which fit loosely into a story line, so that a segment which fails in a trial can easily be reshot without interrupting the story line. A program with a dramatic format, including a chronological story line (such as "Finders Seekers") will need more of the program reshot if a section fails. "Sesame Street" aimed to have children like, understand and learn from the program, so the trial aimed to test whether this was happening. "Finders Seekers" aimed to motivate the children to try out experiments themselves. It would be much harder to carry out a trial which aimed to investigate what children did in their homes or other places because they were "motivated" by a program.

Although the Children's Television Workshop Model has drawn criticism (Postman, 1985, Noble, 1975 and Noble et al, 1979 and 1985), at the same time it has achieved recognition for its importance as a formative evaluation tool (Cook and Curtin, 1985:96). Program makers after 1975 saw the pilot show testing of a trial program as crucial and according to Cook and Curtin, (1985:107) they "borrowed this idea and made it the sole focus of evaluation". However, Cook and Curtin point out that "testing pilots is only a small part of the total Children's Television Workshop model in any of its versions". The ABC has used the trial as an important part of its formative evaluation.

Trial comparisons

The Children's Television Workshop model was used in the trials of "Sesame Street" and the following information utilises my brief summary from the work of Cook et al(1975), Cook and Curtin (1985) and Chen et al (1988).

Information on the "Finders Seekers" trial was obtained from interview with ABC personnel (Parmiter, 1989 and James, 1989), from observation of a trial at a Brisbane State School and from the records of trials kept by the ABC Toowong, Brisbane. The majority of the trials of "Finders Seekers" took place in Queensland, although some trials were held in other states.

In making this comparison an important consideration is that in formative evaluation "the speed of feedback" is very important(Cook et al, 1975:47).

The consequence of speed in the ABC trials was pointed out by the former Head of ABC Audience Research, Palmer, in interview, when she claimed "ABC trials do not pretend to be rigorous research" and that the trial forms were "made for busy people"(Palmer, 1989).

Using the Children's Television workshop model as applied to "Sesame Street", ten points are used as a basis of comparison with the procedure used to trial "Finders Seekers".

1. Setting

"Sesame Street" was trialed in two settings, a controlled research laboratory and in the situation it was likely to be viewed, in the home. The "Finders Seekers" trials were undertaken under normal school viewing conditions. Recorded viewing conditions included a variety of class sizes from 16 to 60, morning and afternoon viewing and, occasionally, a little information was recorded about the condition of the television set. On the trial forms aspects of the viewing conditions were often left blank. There were no laboratory controlled conditions.

2. Subjects

The "Sesame Street" subjects were a group of four year olds who were sampled to represent the target audience of "3-5 year old disadvantaged children (with emphasis on 4 year olds) living in ghetto- type, inner-city areas in various parts of the country" (Ball and Bogatz, 1970: 13). Similar Children's Television Workshop evaluations used a random sample of approximately a thousand children from preschools in selected disadvantaged districts in five cities(Ball and Bogatz, 1970:23).

For "Finders Seekers" no attempt was made to use a random sample. Liaison Officers simply chose cooperative subjects from a selection of schools, districts and states which approximated the target audience of students in years four and five. Interstate ABC Liaison Officers tried to obtain "a cross section of schools" in the selection (James, 1989) In Queensland, attempts were made to include rural communities, but the rural areas chosen were close to the capital city because of financial restrictions. Thus certain types of rural community were not consulted. Ethnic minority groups were included to ensure the trial group approximated the target audience.

The large number of students' responses were not reported in full but were edited for the report. Interstate Liaison Officers collated the results before they got to the producer in Queensland, so he saw an edited version.

Teachers were considered subjects but teacher responses were seriously limited as only one teacher response sheet was available, although probably fifteen teachers would have viewed the program. Discussion

between teachers and the Producer and Liaison Officer were written well after the trial and depended on the Liaison Officers' and Producer's memory. Series Advisory Teachers' comments were compiled.

3. Parts of the program selected for testing

"Sesame Street" programs were trialed in two stages, firstly a test of segments of the program, then pilot testing of episodes.

The trialing process for "Finders Seekers" was carried out as a pilot test using two complete programs. The Liaison Officer and Producer observed and wrote comments on students' reactions to specific segments. While students could comment on specific segments there was no mechanism for responses of teachers to specific segments of the program.

4. Purpose of the trial test

Interest holding capacity, comprehension and learning were tested for "Sesame Street".

For "Finders Seekers" only interest levels of students were tested. Interest levels are shown by observation of concentration levels, and in anecdotes. Children's comments on what part they liked and disliked were used to reveal interest levels.

Learning was not tested. While the producer had a very specific image of science and scientists that he wanted to present, this image was not tested at all in this trial. The messages of the program are the attitudes outlined in ABC Daytime (1989:200) which include the ideas that learning in Science is fun, that occasional failure is to be expected and that children can carry out experiments. Some teachers understood and liked the ethnic variety and handy(wo)man touch, but students didn't seem to comment on these.

There was no comprehension test, and any information on comprehension depended on accepting it as a factor in holding interest.

5. Methods

For "Sesame Street" observation and pre-test/post-test were used. Observation was considered suitable for children with few reading skills.

The "Finders Seekers" trial used observation and the written trial forms to record interest. Observation of children by the Liaison Officer and the Producer was the main method. With only two observers this method has limited accuracy. The two observers tried to record children's concentration ratings for each section of the program, and write anecdotal information (for example, "total still, mouth open" from a trial in August, 1988) for up to sixty children at a time.

Written trial forms from the students included comments concerning what they found interesting. One set of responses I saw included children's drawings.

Teachers' replied in written form to a set of questions, and gave oral feedback when time allowed.

6. Reporting students' responses

For "Sesame Street" these were obtained and recorded by observers and were not written by the children.

Self report forms completed by teachers and students watching "Finders Seekers", were then collated by the Liaison Officers. In these self-report forms students were asked to write what they liked about the program and what they learned. Most responded in sentences but there were drawings. Concentration was observed and recorded on forms by observers. Results were collated onto the final OFFICIAL TRIAL REPORT.

7. Recording

"Sesame Street" aims and results were thoroughly recorded in logs for the purpose.

Written and oral findings were recorded on "Finders Seekers" trial forms and typed as an OFFICIAL TRIAL REPORT, collated by the Liaison Officer. This involved a great deal of editing from the written raw data. Generally the oral raw data took the form of classroom discussion and was not recorded at the time of collection but written up later by the Liaison Officer.

8. Statement of aims

The aims of "Sesame Street's" formative evaluation team were stated in terms of interest holding and of specific learning goals. They were known to the team and written down.

Aims to test specific interest or problem areas of "Finders Seekers" programs were never recorded. This "special agenda" was not stated in any written form, on trial documents or the official report. In interview the Liaison Officer spoke at some length on the dangers of pre-empting "if we tell teachers we are unsure of an actor's speech, then it can have the same effect as saying he is unsuitable"(James, 1989). Yet, this agenda could have had some effect on the process of editing comments.

In the trial of the episode "How Does Your Garden Grow?" the Producer's "agenda" possibly influenced the Liaison Officer's selection of student comments. When interviewed the Producer seemed to be concerned about a certain actor's speech and sense of humour (Parmiter, 1988). The Producer's aim was never revealed to the teachers or the children and was not recorded, yet the form collated by the Liaison Officer contained many comments on this actor.

9. Adult Responses to the trial program

In the original CTW model adults reactions were not part of the trial. [Some variations used pre-test post-test (Eidlitz, 1979) or post-test and Likert scale to test attitudes (Burton, 1979)].

There was one response set of questions for teachers to answer after viewing the "Finders Seekers" pilot. It was only filled in by a small number of teachers, about one in fifteen teachers. Advisory teachers also viewed the pilots. Advisory teachers did not give written responses and their oral comments were summarised by the Liaison Officer.

10. Style of the Series

"Sesame Street" was presented as segments, linked by a loose story line.

The style of the "Finders Seekers" series was dramatic. In the story two children visit their "Aunty" in her home, which is the setting for experiments.

Summary of findings**1. Setting**

The Children's Television Workshop used a laboratory for segment tests while the ABC did not include laboratory observation of children. A number of aspects of the ABC trial were not controlled, for example the condition of the television set and the quality of the picture. Goldman, Hart and others have shown possible effects of a degraded picture on the viewers reaction to a program (Goldman et al, 1982 and Hart 1986). Both series used the place where they would be viewed as the real testing ground for pilot shows. The "Sesame Street" program was trialed in children's homes, the "Finders Seekers" program was trialed in classrooms.

2. Subjects

The "Sesame Street" program had a targeted disadvantaged preschool group for its audience and so the trial was aimed at a similarly disadvantaged sample of what was considered a very homogeneous audience. The "Finders Seekers" program aimed to reach students like its target population (all year four and five students in Australia), so its trial aimed to reach a large cross section, across geographic, cultural, ethnic and social boundaries.

3. Parts of the program selected for testing

Segments of the "Sesame Street" program were trialed as the program was in short segments. Whole pilot shows were trialed to reduce the risk that the series would not hold its audience once it reached them (Cook et al, 1975:48). Only whole programs of "Finders Seekers" were tested, partly because of the dramatic format.

4. Purpose of the trial test

The "Sesame Street" trial tested interest holding, comprehension and learning. The interest holding capacity of "Sesame Street" was a vital part of the segment testing. Comprehension and learning were considered part of the trials as the aim of the series was to stimulate "cognitive growth" (Cook et al, 1975:38).

The main focus of the "Finders Seekers" trials was testing interest, in line with the Liaison Officer's stated aim of giving short term feedback to the Producer. While its stated aim was not to test learning, one question asked the children to write what they learned.

5. Methods

The original "Sesame Street" trials had observation as the main method using "a team of in-house researchers" (Cook et al, 1975:49).

Observation, using only two observers, was the main method of "Finders Seekers".

6. Reporting students' responses

The original "Sesame Street" trials used observations as responses as the preschool children were too young to write. Responses to "Finders Seekers" included both observation and self-report as the focus of the trial was on interest holding and because there could be a great variety in the literacy skills of children in years four and five.

7. Recording

The "Sesame Street" log book was kept for researchers in the future as research on the series was part of a nationally important project which needed to be carefully recorded to justify its funding. "Finders Seekers" trials were not seen as a research project but as a method of getting short term feedback to the producer. Official Reports were compiled and then filed but as no one had asked to see them before, the recording of data was not a high priority.

8. Statement of aims of the trial

A statement of aims was known to all the "Sesame Street" research team because it was seen as a research project. The aims were stated in terms of specific learning goals at one stage because the aim of the series was to teach or instruct.

The teachers in classrooms were not presented with the aims of the "Finders Seekers" trial (described earlier as a "special agenda") so as not to pre-dispose their thinking. Because the trial aimed at testing interest, aims were not stated in terms of learning goals.

9. Responses of adults to the trial program

In "Sesame Street" originally there were no records of adult responses as children were considered the vital audience. Later models tested

adults in two roles, as teacher and as learner. In theory the responses of adults as teachers took up about one third of the trial of "Finders Seekers". In practice it appeared to yield a relatively small amount of the written feedback, as I was able to examine the responses of only one teacher out of fifteen, possibly because of limitations on teachers' time in the school day.

10. Effect of the style of the series

The segmented format of "Sesame Street" was seen as easy to trial, and to change (Cook and Curtin, 1985:108), if the trials indicated a need, compared to the dramatic format used in "Finders Seekers".

Conclusion

Formative evaluation allows changes to be made to a program, so that the series will be successful. Both "Sesame Street" and "Finders Seekers" are successful series. Although an effective trial should lead to a "successful" series, this direct link cannot be guaranteed (Cook and Curtin, 1985:96).

However, the Children's Television Workshop trials were considered responsible for the enormous success of "Sesame Street" and were used as justification for funding of subsequent Children's Television Workshop series (Cook and Curtin, 1985:92).

In Formative evaluation a number of practical and theoretical problems need to be considered in the design of a trial. Practical considerations include the effect of the team work of the people involved and the "halo effect" of interest and enthusiasm for something new. Some factors make a trial easy, for example, the segmented format of "Sesame Street" is easier to trial and change than a drama.

While research continues (Clark, 1990), lack of clear theory on how children react to programs and learn from them in various settings is a concern in trials. There is some research to show that classroom influences encourage learning from television (Krendl and Watkins, 1983). Palmer's *The Lively Audience* (1986) could serve as a model for large Australian studies of how people watch television at school and more work could be done on learning theory and television.

As there are scarce resources in Australia it is important that Australian trials use these resources effectively. Some significant effects of the low level of funding seemed to be influential in the trials of "Finders Seekers" in: the lack of any controlled conditions, the trialing for interest holding alone, the classroom trial having a large oral element which went unrecorded at the time, recording of information for the future being low priority, there being no written statement of the specific areas of feedback desired by the producer and, lastly, that there were few written responses from adults, especially teachers.

While Australian television producers of educational programs have relatively small budgets, a model Australian trial could be more rigorous and should take the following factors into account. The settings would be varied, although research on the effects of this is not yet conclusive.

Trials should reflect the similarities and/or differences of its target audiences. While testing segments may be useful, pilots should be ready early to use in the trial process. Observation as a methods needs careful analysis if numbers of children are high, especially in a school situation. Self-report needs to reflect the literacy skills of respondents. Reporting needs to take into account editing which could obscure the message. Stating the aims of a trial is problematical. It can predispose an audience or can be helpful. Responses of adults can be easily overlooked and need to be dealt with, even for a children's program.

The formative evaluation of both series aimed to trial for interest holding. "Sesame Street" also trialed for learning and comprehension. Perhaps, given limitations on resources, interest holding is the most important factor. An educational program would be counter productive if the program did not hold interest. This is similar to a view accepted by the makers of "Sesame Street":

If we are going to attract children to quality children's programming, then it must have the production values (meaning pace, humour, professional performing talent, film inserts, animation and so forth) to which today's young children have become accustomed (Polsky, 1974:11).

Recording of aims and results of trials (and if possible publication) is important for future formative research.

The references list for this article is not available because it was not provided in the original print version.

Author: Marianne Treuen has been a secondary teacher for ten years in NSW and ACT. She has been a curriculum writer for the Queensland Department of Education. She is a member of the executive of the Australian Teachers of Media (Queensland Chapter) and the Film and Television Subcommittee of the Queensland Board of Senior Secondary School Studies. She is now a lecturer in the Faculty of Education, Queensland University of Technology.

Please cite as: Treuen, M. (1993). Formative evaluation of educational children's programs: A comparison of Australian and American models. *Australian Journal of Educational Technology*, 9(1), 87-97.
<http://www.ascilite.org.au/ajet/ajet9/treuen.html>