A study suggested that audience feedback plays a central role in determining the types of television programs shown on the air and that two major ways series are adopted are through imitation of popular series and through a gradual, evolutionary process. Through an analysis of the prevalence of programs in six genres, these hypotheses were supported. Results indicated that programmers are more likely to air shows of a given genre in years immediately following a year in which there was a highly rated program of that genre. Strongest support for the hypotheses was found in prediction of the number of programs of each type returning to the air. The study concluded that the audience does influence the content and variety of programming offered by networks. While the nature of the feedback process is not yet clear, the study concludes that there is evidence that both imitation and evolutionary processes operate in formulation of network schedules. (Figures and tables of supporting data are appended.) (SRT)
The Audience Role in the Evolution of Fictional Television Content

Daniel G. McDonald
Russell Schechter
Department of Communication Arts
Cornell University
(607) 255-2603/255-6500

The authors express their appreciation to the A.C. Nielsen Company and the CBS television network for permission to access their data archives. Omissions or errors in interpretation are strictly the fault of the authors.

The Audience Role in the Evolution of Fictional Television Content

Abstract

Most models of the communication process include feedback as a necessary component in the functioning of the system. While the feedback notion is very easy to understand in an interpersonal context, the role and extent of feedback in mass communication is less obvious.

This study examines audience ratings as a component of the feedback processes in the evolution of program types. Strong relationships were found between audience ratings and the number of programs of a given type subsequently aired.
The Audience Role in the Evolution of Fictional Television Content

Most models of the communication process include feedback as a necessary component in the functioning of the system (Schramm, 1954; Westley and Maclean, 1957; Kline, 1972). Schramm (1954) suggests that there is little direct feedback from the receiver to the sender in mass communication. Accordingly, an elaborate and expensive array of audience research techniques has been constructed to serve as a substitute for direct feedback in mass media industries. The most elaborate and organized of these is the audience rating industry established to aid broadcasters and advertisers (Beville, 1985; Meehan, 1984).

In spite of the vast array of audience measurement devices and the millions of dollars spent annually in maintaining the ratings system, little theoretical or empirical work has examined the role ratings play as a feedback mechanism. Gans (1972) suggests that audience feedback does have some influence on media content, with past audience choices playing a major role in the determination of new content. More recently, Gans (1980) noted that while a broadcast feedback system based primarily on ratings is well-institutionalized, little is known of the relationship between ratings and what actually appears on television.

Himmelweit (1980), Hirsch (1980) and Gans (1972;1980) all suggest interplay and interdependence between broadcasting and society at large. Himmelweit (1980) includes the interaction of audience and mass media organizations as a feedback component of her model of broadcasting and society, but with little specification of the process in operation.

Schramm (1954) has noted that whereas the individual communicator is relatively free to experiment with types and styles of messages, the mass communication organization must, by necessity, maintain proven formulas.
changing the details but not the essential components of the message. In the broadcast industry in particular, if one organization is rewarded with ratings success for a given message, others tend to copy it. This type of derivative programming is not as much a function of lack of originality as it is of the feedback mechanism employed. Because they are one of the few kinds of reliable feedback available, ratings data are used as the central indicator of audience taste.

Several researchers have developed economic models to explain what kinds of programs are adopted (Boebe, 1977; Crandall, 1974; Greenberg and Barnett, 1971). Several studies point out that the television industry follows an economic model of an oligopoly - with a small number of production formats, genres and plot outlines, produced by a few select production companies, aimed at a single mass audience. Hirsch (1980) notes that the oligopolistic model is not unique to television, but has held true for each of the previously dominant mass media (e.g., motion pictures in the 1920s and 1930s, radio in the 1930s and 1940s). Kellner (1981) makes the point that, for all the much-vaulted competitiveness between the networks, their similarities vastly outweigh their differences.

Others have focused specifically on the broadcasting industry and the relationship between economic factors and the diversity of program offerings, with the assumption that a diversity of offerings is a "good" thing for society, while less diversity is inherently bad (Bates, 1983, 1985; Beebe, 1977; Dominick and Pearce, 1977 Crandall, 1974; Greenberg and Barnett, 1971; Hall and Batlivala, 1971; Levin, 1971; Litman, 1979).

A few studies have examined scheduling and programming practices from a macro perspective. Dominick and Pearce (1976) focused on general trends and cycles in television program types and the relationship between variety
in programming and industry profits, finding a strong (-.81) negative
correlation between program diversity and industry profits. While their
analysis offers little explanation of causal precedence, the correlation
suggests that higher profits are associated with less diversity in program
offerings.

At least two explanations are possible for the relationship between
diversity and profits. The first is that network programmers seek to schedule
programs that represent a "middle ground" - that which is least objectionable
to the most people. This suggests that network programming gets blander and
blander as the seasons progress. However, such an explanation does little to
explain the cycles and trends in dominant program types reported in several
studies (Dominick and Pearce, 1976; Adams and Wakshlag, 1985).

An alternative explanation, and one that helps explain programming
trends, is that programmers seek to maximize ratings or shares of ratings
(Meehan, 1984). If this is true, it should follow that programmers will
experiment with somewhat different types of programs whenever their share
of the audience in a given time period is low. The most obvious strategy
is to attempt to ride the crest of changing audience tastes by programming
to those predilections. Programming innovations usually occur when the
least successful network tries to differentiate its product from the others
by introducing new concepts (e.g., Monday night football). Hirsch (1980)
suggests that successful innovations are "matched," or imitated within a
year by the other networks.

Adams, Eastman, Horney and Popovich (1983) analyze the cancellation
and schedule manipulation of network television programs and conclude that
audience ratings are not strongly related to cancellation. Examining
series from 1974 to 1979, Adams, et.al., (1983) conclude that the audience
has a very limited ability to influence which programs will be retained on the air.

A problem with accepting the Adams, et.al. (1983) conclusion is that it is counterintuitive. In an economic framework, mass communication systems function primarily in response to viewer acceptance of programming, with ratings serving as the chief measure of that success. Audience ratings were established in 1930, within four years of the beginning of network radio (Beville, 1985). If seen in terms of a mass communication system, it is difficult to accept the notion that the past 55 years have not seen a refinement in the process and technique of programming for particular audiences, or that an organized audience feedback process would not permit major audience input.

The present study combines the perspectives of the Dominick and Pearce (1976) and Adams, et.al. (1983) work to investigate the extent to which audience ratings work to effect choices available to viewers. We suggest that examination of cancellation and manipulation of programs within time slots may result in misleading indicators of how ratings work in the formulation of network schedules. Broadcasters deal with a finite number of program slots, and should be continually assessing not only how well any specific program is doing in the ratings, but also which program types are increasing in popularity, and which types are declining. New programs should be primarily a function of growing generic popularity; cancellations should be a function of programmers' perceptions of lost opportunity.

Stressing the audience-industry aspect of the Himmelweit (1980) interdependence model of broadcasting and society, we hypothesize that broadcasters respond to gradual changes in audience tastes and preferences. We hypothesize that the model should operate in two ways: first, through a
slow, building process in which networks air types of programs that allow
the audience specific gratifications relevant to events occurring in the
social system (e.g., economic and social conditions, etc.); and second,
through the imitation of successful new series. That success should alert
network programmers that a large portion of the audience is viewing a program
having very specific traits; the logical tactic would be to provide similar
traits in another program.

The first process, gradual change in content is difficult to document. We see content evolution as very nearly an imperceptible averaging or
"evolutionary" process, and suggest such a process to account for certain
trends in program types noted in previous studies, such as the rise and
decline of westerns, or cycles in action/adventure programs (Dominick and
Pearce, 1976). We see the evolution as taking place over several years, in
which less obvious (i.e., non-structural) characteristics come to the fore
among the networks' major offerings. Here, programs which may not seem
similar on the surface may be geared toward similar beliefs or political
values, or may fulfill similar gratifications (e.g., a police drama with
emphasis on characterization may be similar to a soap opera set in a modern
hospital, in that both stress interpersonal relationships over formula
narrative).

The second process, imitation, results in what are known as "rip-
offs," in which successful series are blatantly imitated, (e.g., "227" and
"Growing Pains" in the wake of "The Cosby Show"), and "spin-offs," in which
characters or situations of a particular program are given series of their
own, (e.g., "Laverne and Shirley" and "Mork and Mindy" characters first
appeared on "Happy Days"). We consider both rip-offs and spin-offs
instances of imitation, and expect that imitation (particularly in the form
of rip-offs) occurs most frequently when a new series is a hit in its first season. As Hirsch (1980) suggests, we would expect imitations to appear within a year of a program attaining "hit" status. We would add that imitation usually results in situational and demographic similarities (e.g., a successful show featuring a black doctor being imitated by development of other series featuring black professionals), rather than similarities along the lines of quality of writing, character development, or other substantive elements. This study tests the model of evolution and imitation by examining the relationships between the types of programs available during a given year and average ratings for those programs during the period 1957-1985.

METHODS

Program Coding

All fictional network prime time programs aired between 1946 and 1985 (2446 programs) were coded from written descriptions appearing in Brooks and Marsh (1985). A categorization scheme consisting of 41 non-exclusive program types was constructed based upon extensions of A.C. Nielsen classifications, previous program types studies (e.g., Dominick and Pearce, 1976; Wagshlag and Adams 1985), and a formally organized version of the undefined Brooks and Marsh (1985) scheme. Coding was carried out by volunteers solicited from an introductory mass media course.

Coders were provided with detailed instructions and definitions of categories and randomly assigned packets of program descriptions from Brooks and Marsh (1985). Each program was coded multiple times (mean number of codings per program was 5.5). Final categorization was determined by aggregating across all codings of each program. The resulting percent of intercoder agreement across all types for each program was considered a
measure of the extent to which a program exhibited the characteristics of each category.

The present research examined only fictional network television content by including 19 of the 41 possible categories. Principal components analysis with varimax rotation was used to reduce the number of categories while retaining the integrity of the original coding scheme. The resulting factors also provide a method of classifying programs through elements that should appear in both the imitation of series and the evolution of program type trends.

Ratings

Coding information was matched with average, annual national ratings data provided by A.C. Nielsen, published in Variety, or made available by the CBS television network archives. Despite the varied sources from which the data had to be gathered, all of the original data were collected and published by the A.C. Nielsen company in May of each year, and are directly comparable. Full ratings information was collected only for regular season programming. The unavailability of complete ratings data for the period 1948-1956 led to the selection of 1957 as base year for the study. Ratings were standardized by year to control for shifts in ratings methodology and general ratings trends over the period contained in the study.

Factor Analysis

Analysis of the 13,512 codings of the 2,446 programs yielded 6 factors with eigenvalues above 1.0, accounting for 45.4% of the variance in coding (Table 1). The first five factors were readily interpretable, and are notably similar to those used by the A.C. Nielsen Company in their yearly summaries of television programming (Nielsen, 1985). The sixth factor was
somewhat less intuitive in that it appeared to reflect a mixture of spectacle, novelty, and anthology.

The factors were labeled suspense (with loadings on drama, police, detective and spy/intrigue), comedy (loading on family, situation comedy, and school/education), fantasy (primarily consisting of fantasy/superhero, science fiction, and occult/horror), drama (loading on drama, medical and soap opera), physical conflict (with high loadings for war/military and western), and non-seriality (a mixture of anthology, sports and circus types). For clarity of presentation, the six factors will hereafter be referred to as program genres.

The next step was to classify programs according to the six program genres obtained through the factor analysis. For four of the six genres, programs with standardized factor scores above 1.96 were selected from all other programs. For two of the genres, comedy and drama, less than 100 programs scored above 1.96. Because of this difficulty, a grouping point of 1.65 was chosen for these genres. Five hundred and ninety-eight programs met the criteria and were classified into the six genres.

Number of Programs

For each of the six genres, three dependent variables related to the number of programs per year are used. The first of these is the total number of programs of a particular type (i.e., the count of all shows with factor scores high enough to be included within that type). This count forms a time series illustrating the rise and fall of dominant program
types on network television. Figure 1 illustrates this series with two of
the genres, physical conflict and drama.

The second dependent variable is a simple count of the number of new
programs of a given type (i.e., the count of all premiering programs with
factor scores high enough to be included within that type). This second
time series forms an indicator of the extent to which the networks are
attempting to provide additional programs of a particular program genre.

The third measure of the number of programs results from the
subtraction of the number of new programs from the total number of programs
of that type for each year in the sample. This time series provides a
measure of the number of programs of a given type that return (i.e., are
not cancelled) each year, an index of stability in the network schedule.

Genre Ratings

Average ratings per year (genre average) were computed for each of the
program genres. This average should be a major source of the kind of
information necessary for the gradual evolutionary process described above.
In addition to the averages, maximum genre ratings for each year (for each
of the six genres) were used to obtain a measure that should be related to
both evolution and imitation programming strategies. Finally, maximum
ratings of programs of a given genre debuting in a particular year were
obtained for each of the six genres. This maximum debut rating should be
most clearly related to imitation programming strategies.

Two major analyses were used to test the relationship between previous
ratings and the type of programs presented each year, discriminant analysis
and time-series regression analysis. Discriminant analysis was used to
test the extent to which maximum genre ratings of the previous year, as
indicators of a mixture of both evolution and imitation strategies, can
predict which type of programs are aired the following year. Time-series regression analysis techniques were used to assess the relationships between the prevalence of the six genres (analyzed separately) and previous year ratings (average and maximum), and the relative importance of direct imitation and gradual evolution in network program offerings.

RESULTS

Discriminant Analysis

To test for the general relation between previous ratings and which programs appear on the air, genre maximums for the previous year were combined to discriminate between each of the six genres. It might be argued that this is a conservative test of the relationships between ratings and program prevalence, as within each year, six measures (the genre maximum ratings) are used to discriminate between all of the programs on the air that year that met the grouping requirement. In addition, more than one genre may be popular at a time, and, with a limited amount of broadcast time available, programmers may find they have to choose between two popular genres, yet the procedure tests for direct relationships between ratings for a type and prevalence for the type.

Two discriminant functions were statistically significant in the classification process (Table 2). The total structure coefficients (Klecka, 1980) are correlations between the discriminant functions and the original variables (i.e., the previous year maximum ratings). With examination of the group centroids, function 1 appears to tap a "physical vs. mental action" discrimination between programs, though function 2 is more difficult to evaluate. Function 2 appears to reflect a "reality-based drama vs. fantasy" orientation. Twenty-nine percent of the programs were
correctly classified into the six program genres.

Table 2 about here

Time-Series Regression

For each of the six genres (factors), and each of the three program count variables (prevalence of genre, returning programs, and new programs of the genre), multiple regression analyses were run to determine the relationship between the average rating per genre, the maximum rating per genre, and the maximum rating for programs debuting within each genre. Results of the Geary and Durbin-Watson tests for serial correlation of residuals were not significant for any of the equations (Ostrom, Jr., 1978).

Average ratings were significant predictors of prevalence of programs in four of the six genres: physical conflict, comedy, nonserial, and drama (Table 3). Overall, results of Table 3 suggest fairly strong relationships between average ratings per program type and the number of programs of that type on the air within one or two years.

Table 3 about here

Somewhat weaker results are found in the analysis of the number of programs of a given type debuting each year (Table 3). Significant relationships exist between the average ratings of physical, nonserial and drama programs, and the number of new shows of those types the following year.
Strongest results were obtained in examination of the number of programs returning to the air (Table 3). In this case, returning programs in 5 of the 6 genres (all but fantasy) were significantly predicted by previous ratings.

DISCUSSION

The present study began with the suggestion that, contrary to the interpretation of several previous studies, audience feedback should play a central role in determining the types of programs on the air. We suggested that two major ways series would get on the air were through imitation of popular series and through a gradual, evolutionary process.

Our analysis of the prevalence of programs of six genre types provides support for this contention, although the exact process may be specific to the genre. The discriminant analysis tested the relationship by attempting to classify programs into genres merely on the basis of the maximum rating for each of the six genres. Results were significant, suggesting that programmers are more likely to air shows of a given genre in years immediately following a year in which there was a highly rated program of that genre. The discriminant functions suggest some spill-over into different genres, however, as years in which physical action programs do well appear to be followed not only by more physical action, fantasy and nonserial programs.

The time-series regression provided more detailed support for the hypothesized model of evolution and imitation. The total number of programs appears to be a function of imitation for fantasy, comedy and drama (because it is predicted by a combination of genre and debut maximums of the previous year). However, the total number of nonserial programs appear
to be best explained by a process of evolution (predicted by genre averages and debut maximums of two years before). Prevalence of the physical conflict genre is best explained by a mixture of evolution and imitation.

For the number of programs making their debut each year, the regression equations were significant for three of the six genres, suggesting the difficulty in predicting how many of each genre will debut in any given year. On the other hand, three statistically significant equations (out of six) is well above the number expected by chance, again providing support for our hypothesized model.

Strongest support for our model was found in prediction of the number of programs of each type returning to the air. Equations for five of the six genres were statistically significant, although the particular combinations of significant variables were unique for each equation.

The present study thus suggests that the audience does influence the content and variety of programming offered by the broadcast networks. While the nature of the feedback process is still not clearly specified, we have found evidence that both imitation and evolutionary processes operate in formulation of network schedules. Further research needs to expand on the program types analyzed here by including a range of fictional and nonfictional types. A wider range would enable clearer specification of the feedback processes and how the extent and method of feedback differs for various program genres.

One confounding factor may be trends that develop which have manifestations across genre types. So, to provide a recent example, the phenomenon of music videos has triggered a formal response in production, leading to programs like "Dreams" and "Miami Vice." Success of the latter brings about imitation, such as "The Insiders," which may be only
targentially related to the imitated source.

Finally, our model should not be interpreted as suggesting that ratings are the sole consideration in programming. Numbers alone may be misleading in the case of programs generating a key demographic audience, and a myriad of non-ratings factors probably influence programmers, including genre success as exemplified by popular feature films, relationships between network executives and particular production companies, pressures from corporations and advertisers, and personal quirks and preferences among programmers. More detailed research into the function of ratings conjoined with analyses of issues such as those suggested above are necessary before any complete picture of the intricacies of network scheduling can be drawn.
References


Figure 1. The prevalence of physical conflict and drama programs, 1957-1984.
<table>
<thead>
<tr>
<th>Table 1</th>
<th>Factor Loadings from Principal Components Factor Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1 Suspense</td>
</tr>
<tr>
<td>Anthology</td>
<td>.002</td>
</tr>
<tr>
<td>Drama</td>
<td>.748</td>
</tr>
<tr>
<td>Situation Comedy</td>
<td>-.131</td>
</tr>
<tr>
<td>Circus</td>
<td>-.047</td>
</tr>
<tr>
<td>Family</td>
<td>-.048</td>
</tr>
<tr>
<td>Fantasy</td>
<td>.055</td>
</tr>
<tr>
<td>Lawyer</td>
<td>.239</td>
</tr>
<tr>
<td>Medical</td>
<td>.018</td>
</tr>
<tr>
<td>Media</td>
<td>.126</td>
</tr>
<tr>
<td>Police</td>
<td>.716</td>
</tr>
<tr>
<td>Detective</td>
<td>.702</td>
</tr>
<tr>
<td>School</td>
<td>.038</td>
</tr>
<tr>
<td>Science Fiction</td>
<td>-.036</td>
</tr>
<tr>
<td>Soap Opera</td>
<td>-.006</td>
</tr>
<tr>
<td>Sports</td>
<td>-.020</td>
</tr>
<tr>
<td>Spy/Intrigue</td>
<td>.341</td>
</tr>
<tr>
<td>Occult/Horror</td>
<td>-.065</td>
</tr>
<tr>
<td>War/Military</td>
<td>-.095</td>
</tr>
<tr>
<td>Western</td>
<td>.076</td>
</tr>
</tbody>
</table>

| Eigenvalue | 2.138 | 1.663 | 1.480 | 1.163 | 1.093 | 1.081 |
| Variance Explained | .113 | .088 | .078 | .061 | .058 | .057 |

Note. N=2446 programs.
Table 2
Discriminant Analysis Predicting Prevalence of Genres

<table>
<thead>
<tr>
<th>Previous Year Genre Rating</th>
<th>Total Structure Coefficients</th>
<th>Functions Evaluated at Genre Centroids</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Function</td>
<td>Genre</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Physical</td>
<td>.853</td>
<td>.234</td>
</tr>
<tr>
<td>Fantasy</td>
<td>.672</td>
<td>-.257</td>
</tr>
<tr>
<td>Nonserial</td>
<td>.473</td>
<td>.805</td>
</tr>
<tr>
<td>Drama</td>
<td>-.166</td>
<td>.461</td>
</tr>
<tr>
<td>Suspense</td>
<td>-.516</td>
<td>.569</td>
</tr>
<tr>
<td>Comedy</td>
<td>-.464</td>
<td>-.092</td>
</tr>
</tbody>
</table>

Explained Variance .662 .236

Canonical Correlation .303 .187

Note. N=598 programs.
## Table 3
Hierarchical Regression Analysis of the Prevalence of Six Program Genres

<table>
<thead>
<tr>
<th>Rating</th>
<th>Suspension</th>
<th>Physical</th>
<th>Fantasy</th>
<th>Comedy</th>
<th>Nonserial</th>
<th>Drama</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Number of Programs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genre Average $t-1$</td>
<td>-.258</td>
<td>.709*</td>
<td>.287</td>
<td>-.066</td>
<td>.811*</td>
<td>-.131</td>
</tr>
<tr>
<td>Genre Maximum $t-1$</td>
<td>.506</td>
<td>.613*</td>
<td>.603*</td>
<td>-.400*</td>
<td>.172</td>
<td>.852*</td>
</tr>
<tr>
<td>Debut Maximum $t-1$</td>
<td>.078</td>
<td>.499*</td>
<td>-.149*</td>
<td>.511*</td>
<td>.169</td>
<td>.371*</td>
</tr>
<tr>
<td>Debut Maximum $t-2$</td>
<td>.088</td>
<td>.352</td>
<td>.097</td>
<td>.031</td>
<td>.672*</td>
<td>-.292</td>
</tr>
<tr>
<td>Total $R^2$</td>
<td>.193</td>
<td>.864*</td>
<td>.332</td>
<td>.378*</td>
<td>.899*</td>
<td>.664*</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>.053</td>
<td>.840*</td>
<td>.216</td>
<td>.270</td>
<td>.881*</td>
<td>.606*</td>
</tr>
</tbody>
</table>

|                      |            |          |         |        |           |       |
| **Number of Programs Debuting** |
| Genre Average $t-1$ | -.221      | .583*    | .113    | -.093  | .811*     | -.401*|
| Genre Maximum $t-1$ | .165       | .178     | .313    | -.331  | .172*     | .563* |
| Debut Maximum $t-1$ | -.094      | .590*    | -.323   | .250   | .169      | .249  |
| Debut Maximum $t-2$ | .128       | -.046    | -.018   | .319   | .672*     | -.318 |
| Total $R^2$          | .080       | .683*    | .107    | .261   | .899*     | .486* |
| Adjusted $R^2$       | .000       | .628*    | .000    | .132   | .881*     | .397* |

|                      |            |          |         |        |           |       |
| **Number of Programs Returning** |
| Genre Average $t-1$ | .489*      | .642     | .302    | -.016  | .721*     | .167  |
| Genre Maximum $t-1$ | .696       | .729*    | .538*   | -.330  | .904*     | .874* |
| Debut Maximum $t-1$ | .167       | .355*    | .184    | .593*  | .235      | .377* |
| Debut Maximum $t-2$ | .014       | .498*    | .176    | .367*  | .179      | -.179 |
| Total $R^2$          | .495*      | .794*    | .309    | .525*  | .873*     | .670* |
| Adjusted $R^2$       | .407*      | .758*    | .189    | .442*  | .851*     | .613* |

*p < .05

Note: Unless otherwise noted, table entries are beta weights. Analysis is based on a series of 28 years, 1957-1984.