

Program Appeal and Program
Impact: A New Approach to
Qualitative Ratings

by Elizabeth J. Roberts

Do not quote without the permission of the author.
©1984 Columbia Institute for Tele-Information

Columbia Institute for Tele-Information
Graduate School of Business
Columbia University
809 Uris Hall
New York, NY 10027
(212)854-4222

PROGRAM APPEAL AND PROGRAM IMPACT:
A NEW APPROACH
TO QUALITATIVE RATINGS

Elizabeth J. Roberts

President
Television Audience Assessment, Inc.

October 1984

Research Working Paper Series. Not for citation,
quotation, reproduction, or distribution without
written permission.

This paper was presented to the Columbia University
Graduate School of Business Conference: "Beyond
Ratings: New Directions in Audience Measurement
Research"

#44

THE CHANGING TELEVISION/ADVERTISING MARKETPLACE

Traditional television ratings that measure the size and demographic characteristics of a program's audience do not necessarily indicate the degree to which viewers are involved in and receptive to what they are watching. These issues had little relevance in the 1950s when the traditional ratings systems were adopted and television was scoring its first wave of spectacular success as a mass-market advertising medium. The nature of the television medium and the viewing habits of the American public at that time meant that estimating the number of people tuned into a program gave a satisfactory indication of how valuable it was as an advertising vehicle.

Consider the capacity of the television set and its role in the American home thirty years ago. A typical household felt privileged to have one set, with three or four channels available. Even in areas where poor reception reduced the options to only one or two stations, the signals of those stations were beamed to households that had received nothing of the kind ever before. When they watched television, family members gathered around the wonderful invention in rapt attention for a program they had planned ahead to see and would probably watch again the next week. The limited number of program alternatives coupled with the novelty of the medium enabled industry executives to take it for granted that television

viewers were involved with and loyal to the programs they watched. People who had tuned in a program, it could be assumed, would also see the advertiser's message.

In American homes today, television is a commonplace, not a wonder. Most families now have two sets; many have three or four. Furthermore, the television sets themselves have a greatly altered capacity. Within a few years, over half of all American households will be wired for cable, offering viewers a choice of 20 or 30 programs at any one time. Hand in hand with technological changes have come changes in people's viewing patterns--changes that call into question the wisdom of relying heavily on traditional television ratings when making marketplace decisions.

In the first place, the television set has now become as familiar as a member of the family. In the average home, a set is turned on in one room or another for over seven hours a day, every day of the week.¹ With sets playing for so much longer than they once were, television programming and its advertising must compete with dozens of other demands on people's time and attention. Viewers, of necessity, often engage in other activities while programs and commercials play away in the background. Research conducted by Television Audience Assessment (TAA) in 1981 and 1982 established that nearly 40 percent of audience members are reading, talking on the telephone, eating

dinner, washing dishes, paying bills, doing homework, or actively pursuing some other distracting activity while "watching" the average television show. Furthermore, fully half the audience leaves the room at some time during the show, and most of these people leave not once but repeatedly during both the program and its commercial breaks.² But the audience is not always so distracted. For some programs, research makes plain, viewers set aside distractions and pay close attention.

Second, with so many programs to choose among, today's viewers evidence little program loyalty and switch channels often. Our earlier research found that fewer than two-thirds of the audience for an average hour-long program actually stays with it to the end, and only one-third of the viewers of one week's episode are likely to be in the audience for the same program the following week.³ Viewers' restlessness is also apparent when commercials come on. Almost 15 percent of viewers without cable and 40 percent of cable subscribers report that they "always" or "often" change channels during commercial breaks.⁴ The growing use of remote-control units makes "commercial-zapping" easier all the time, and the availability of continuous special-format programs such as CNN, MTV, and The Weather Channel makes it more tempting. The growing array of alternative programming apparently has become so seductive that one out of ten cable subscribers reports frequently watching two programs at once, switching back and forth between them.⁵ But once again, research shows that some

programs are markedly better than others at holding their audiences.

Changes on this scale suggest that new market strategies are needed. There is clearly room for innovative programming, new promotional approaches, and new media planning and advertising practices in the television industry. However, in order to plan, produce, program, and advertise in this market, industry executives need sensitive measures of a television program's ability to perform.

Today program performance is evaluated primarily on the basis of the syndicated ratings data supplied by A.C. Nielsen and Arbitron. These companies count the house. Pure and simple. They answer basic circulation questions: How big is the audience? Who is in it by age and sex? How are viewers divided among the available channels at a particular time on a particular day? But whether viewers are involved with or are enjoying what they see, whether they find the program they are watching worth planning ahead for, whether they will set aside other activities to watch it or remember either the program or the commercials shown in it, are vital pieces of information these traditional ratings do not provide. As a result, traditional approaches to television ratings make no distinction between those programs which have an audience of distracted restless viewers and those programs which are more effective at capturing an attentive, involved, satisfied and loyal audience.

Although in the past it was possible to ignore such variations in the way audiences responded to programs, today treating viewers as an undifferentiated mass is a costly anachronism. All the new factors in the marketplace have led many observers to suggest that traditional quantitative television ratings could be usefully supplemented by more descriptive measures of the audiences reactions to programs.⁶ Such emerging approaches have become known within the television industry as "qualitative" program ratings, although it should be noted that none of the various systems, either proposed or in use, actually measures the "quality" of programming. What they attempt to measure are viewers attitudes toward programming, defining quality solely in terms of the audience's opinion.

Today, almost a dozen countries around the world are involved in qualitative audience research designed to gather information about the degree to which viewers like, appreciate, enjoy or otherwise react to the television programs they view.⁷ In the United States, however, no system has emerged as an agreed-upon standard of qualitative ratings that American television and advertising executives can rely on to supplement the Nielsen and Arbitron data. A review of the available qualitative audience ratings studies, both domestic and international, suggests that there are several reasons.⁸ First, research often is done on a proprietary basis, and the results of such studies are not widely available for industry use. Second, many of the qualitative

ratings research studies suffer from inadequate sample size or other methodological limitations, making it difficult to trust their conclusions. Third, the measures of audience involvement and satisfaction used in many studies have not been adequately tested for their validity and reliability. Fourth, and most important, most research efforts have not examined the relationship between viewers' reactions to programs and commercially relevant viewer behaviors such as planning ahead to watch specific programs, repeat viewing, levels of attention and distraction, commercial exposure, and commercial effectiveness. As a result, the implications of qualitative program ratings for economic decision-making have not been apparent to most industry decision-makers in the United States.

TELEVISION AUDIENCE ASSESSMENT UNDERTAKES EXTENSIVE RESEARCH PROGRAM TO TEST NEW QUALITATIVE RATINGS SYSTEM

Television Audience Assessment (TAA) was established in 1980 to develop and test a new qualitative ratings system for regular use by the American television and advertising industries. Following extensive discussions with leading executives from all aspects of the television business, Television Audience Assessment launched a major four-year research and development program. A series of studies (including a special prototype ratings study) were undertaken to: (1) gain insight into how the changing

communications environment is affecting the way people use television; (2) explore the variations in viewers' reactions to the programs they watch on both cable and broadcast television; (3) test different methods for collecting qualitative program data; and (4) examine the relationship between viewers qualitative response to programs and their exposure and receptivity to commercial messages. (See Appendix A for full description of all studies.) The extensive data base created through this four year research effort provided information not only on viewers involvement with programs and commercials, but on their selection patterns and their behavior while watching television. It enabled us to analyze different qualitative ratings indexes and to rate hundreds of different programs on the basis of these prototypes.

The new qualitative rating system that emerged from these studies is based on two indexes which measure how enjoyable and involving a program is for its viewers:

The Program Appeal Index measures viewers evaluation of the overall entertainment value of a television program.

The Program Impact Index measures the degree of intellectual and emotional stimulation a program provides its viewers.

Using data from our prototype ratings study we tested the Appeal and Impact Indexes for their reliability, stability and utility in predicting audience behavior. Both held up well under scrutiny and both were found to correlate with all the audience behaviors examined through the different studies--program planning, loyalty, behavioral distraction, attention and commercial exposure and effectiveness.

The implications of this research are extremely significant for the commercial development and long term acceptance of an industry-wide standard of qualitative ratings. If certain programs are more desirable advertising vehicles than others, not simply because of the size of their audience but because they involve viewers more, sustain their attention and loyalty, and improve their response to the commercials placed within it, then in the future, optimal programming and advertising decisions will benefit from the systematic assessment of both the quantitative and the qualitative performance of programming.

The following sections of this report are based on the data that emerged from this research. Taken together the findings demonstrate that a qualitative ratings system that regularly delivers broad-based and reliable statistics on how the audience rates the Appeal and Impact of television programming shows remarkable promise as a source of commercially valuable and methodically sound information that is now available to the

television and advertising industries. The new information provided by Appeal and Impact ratings could greatly enhance general understanding about viewers use of television -- understanding that is critical for efficient media planning, effective use of narrow casting, and the development of programming to suit diverse consumer preferences in the television market of the 1980s and beyond.

PROGRAM APPEAL: A MEASURE OF VIEWER'S ENJOYMENT

All programs are not equally appealing to the audience

Americans enjoy television. Each evening millions of viewers seek relaxation, companionship, and distractions from their everyday problems by watching it. How much enjoyment they find varies from program to program but, overall, most viewers find television an appealing way to pass the time. Rating programs on a scale equivalent to 0-100, the average Appeal Score of over 250 programs rated in our 1982 prototype ratings study was 73, and 90 percent of all shows were scored at 50 or above by their viewers.

While theoretically a program that an audience finds extremely unappealing might score as low as zero on the Appeal Index, in practice most people don't sit through a program they don't enjoy at all. Clearly, unless viewers are watching a show that someone else has chosen, most people will switch channels, leave the

room, or turn off the set rather than continue to watch a show they find totally unsatisfying.

Although people enjoy what they watch on television enough to rate most programs above 50 on the Program Appeal Index, they register a variety of opinions within that upper range of the scale. The highest-rated show from our two-week study in the spring of 1982 received an Appeal score of 91, while the lowest Appeal rating for the same period was 46. Clearly, the audience perceives sharp differences between programs and registers its reactions accordingly.

By and large, the viewers within our studies appeared to enjoy a program for reasons specific to the program itself, not simply because it was a program of a particular type. Public affairs programs, dramas, and nighttime serials all received somewhat higher average Appeal scores than other program types. However, comedy programs, action/adventure shows, and variety shows, were not without their fans. One of the more significant findings in our program analysis was how little a program's type influenced a show's Appeal and different programs, even within the same genre, varied widely in the amount of enjoyment they offered their audience.⁹

Different programs appeal to different audience subgroups

Qualitative ratings are based on a belief that the television viewing audience is not an undifferentiated mass but rather a

composite of individuals who respond to individual programs differently.¹⁰ The appeal of some programs in our studies varied noticeably for different age groups of men and women within the audience, while others had a more universal appeal.

While age and sex are obviously factors in how much subgroups of the audience like a program, their effects changed with each program. By contrast, the educational background of viewers exerted the only systematic influence on Appeal scores that we found in the demographics of the audience. In general, adults with less than a high school education tend to enjoy all the programs they watch somewhat more than viewers with more formal education. While college-educated viewers are clearly capable of deriving significant enjoyment from television programs, it appears, however, that most programs hold somewhat less appeal for these more educated viewers.

We were also interested in what effect access to cable television had on viewers' satisfaction with the programs they watch. Does a greater choice of channels mean that viewers living in cable homes are ultimately more satisfied with the programs they end up watching? Not necessarily. Usually there was little difference between the Appeal scores given by cable subscribers and scores given by broadcast-only viewers who were watching the same program. While cable subscribers, particularly those with access to pay services, watch considerably fewer programs on network

television, their opinions of the network programs they do watch generally are similar to those of the broadcast-only viewers. Whether people have access to eight channels or 30, whether they subscribe to one or more pay services or not, it is the inherent appeal of the individual programs they view, not how many options they have to choose from, that influences their evaluation of television.

Thus we see that the audience's reactions to specific television programs, and, therefore, the Appeal scores that various subgroups give these programs, can vary noticeably with certain features of the audience--age and sex, education level, and media access. But perhaps the most interesting feature of these variations is how difficult it is to generalize about them by features of the audience. In all the demographic variables we studied, we found it possible to make only one broad conclusion based on the audience: viewers with more education like television somewhat less than do viewers who have less schooling. All the rest of the variation we find among audience subgroups in their reactions to the same program is idiosyncratic to the show. Such differential appeal suggests that, program by program, industry decision-makers will find Appeal scores useful (in establishing a nightly schedule) by assessing how much a program satisfies different segments of the audience, and in determining the potential for audience flow from one appealing program to the following one. Likewise, Appeal scores offer useful information

to advertisers interested in targeting their commercials to specific demographic groups.

A program's Appeal is stable from episode to episode

Over 70 percent of the 45 weekly series we examined showed no statistically significant variation in the level of enjoyment they provided viewers from one episode to the next. This stability in the Appeal score is particularly striking in light of the high audience turnover we found; an average of 68 percent of the audience for the first episode did not return to view the second, yet the Appeal score remained similar from week to week. We found the same pattern of stable Appeal scores to be equally true of syndicated "strip" programs (programs shown nightly), with one night's viewers of a syndicated series evaluating the appeal of a program at about the same level as the previous night's viewers.

15-100
w/ [unclear]
[unclear]
[unclear]
[unclear]
[unclear]

Furthermore, when we compared the Appeal scores for all programs in our sample currently appearing in both a first-run network and syndicated strip format, we again found little significant difference. This was true despite the fact that the programs were shown in different time periods on different days of the week and had a substantially different audience. (See Figure 1.)

Finally, we pursued this investigation of episode variability in greater detail by looking at another potential source of

variation in appeal: differences in Appeal scores between reruns and new episodes of established series. For only a few programs was there a significant difference between the Appeal score for a new episode of the same series. In the majority of cases, whether an episode was new or a rerun had little effect on the program's perceived appeal.

It appears, then that the appeal of most programs is a relatively stable feature of the show, and the Appeal Index is a consistent, reliable measure of a show's intrinsic entertainment value. This reliability is of more than theoretical interest. It indicates that a functional qualitative ratings system need not sample the audience every week in order to provide useful data for first-run or syndication programming decisions. The rating for one episode generally predicts the rating for the series, at least over a limited period of time.

Both our own evidence and what has been gathered in research abroad thus assures us that it is possible to measure the audience appeal of television programs, that programs differ widely in their appeal, and that Television Audience Assessment's Appeal Index is a reliable measure because it does not vary substantially with program genre or even with different episodes of a series. In other words, the audience perceives that a particular program has a particular level of appeal, regardless of its type and regardless of which episode of the program is

reliable
measure
of what?

1/10/70
reliable
measure
of
entertainment
value -
constant
of
technical
quality

being shown. From this we can infer that the audience likes a program it rates 80 better than it does a program it rates 70 or 60. But the question is, does this make any difference? What might it mean to the industry that one program is more appealing than another?

Program Appeal predicts viewer planning and loyalty

One of the useful features of the Program Appeal Index and, as we see in the next section, the Program Impact Index is how closely they correlate with certain aspects of the audience's behavior. The Appeal Index in particular was found to be useful in gauging how loyal viewers will be to a specific program. It helps predict the likelihood that viewers will plan in advance to watch a particular show and will watch other episodes of the series.

Planning. Most people do not plan in advance to watch television programs. They generally turn on the set and then choose a program for the available option, often settling for what has become known in the industry as the "least objectionable" program choice. In many homes, people appear to turn on the television set the way they turn on a light when they walk into a room, without thought or planning. Frequently, television serves as little more than a talking lamp, carrying on a monologue in one corner of the living room while family members go about the business of their daily lives. When asked to indicate why they selected a particular program to view, approximately 75 percent

of the respondents in our prototype ratings study replied, "The first few moments caught my interest."

Unplanned selection was particularly true for less frequent viewers of television. Our study provides little support for the theory that light viewers of television are more selective viewers. In fact, it is the heaviest users of television who are most likely to plan in advance to watch a program, with 53 percent of people who watch more than two hours of television an evening planning ahead to see a program, compared to 42 percent of viewers who watch less than one hour a night. The more television that people watch, the more likely they are to know ahead of time what they want to see.

*1.30-1.40
Unplanned selection
with the heaviest
viewers*

While viewers are somewhat more likely to plan ahead to watch information programs, including news and public affairs shows, than they are for standard entertainment fare, there was a consistently strong correlation across all the programs studies between planning ahead and viewers' enjoyment as measured by the Appeal Index. (See Figure 2.) In the case of individual ongoing programs, these patterns of advance planning usually lead to increased viewer loyalty and repeat viewing.

*1.40-1.50
Unplanned selection
with the heaviest
viewers*

Loyalty. The far greater choice afforded by cable television has meant a reduction in overall program loyalty. With more programs to choose from, many viewers are less apt to "tune in next week"

to see a subsequent episode of a show they watched the previous week. We found in our prototype ratings study, which intentionally included a large proportion of cable subscribers, that only about 30 percent of the audience of one week's episode returned to watch the next week's episode of the same series. However, programs differed widely in the proportion of the audience that they were able to carry forward from one week to the next. When we compared viewers' Appeal scores with their repeat viewing patterns for a sample of series programs, we found that, in the great majority of cases, viewers who watched both episodes of the series reported finding the show more appealing than viewers who watched only one of the two episodes. The most satisfied viewers of the show were also the most loyal viewers. This finding was further substantiated in a later laboratory study by the finding that those participants who rated a program high on the Appeal Index were more likely to say they "usually made a special effort to watch the show" and that they "had frequently watched the series before". Since series programs succeed by developing a following of people who will tune in to the program week after week, the Appeal Index offers an important tool for gauging a series' potential for building an audience of loyal viewers.

The size of the audience does not indicate the Appeal of a program

Having found the Appeal Index useful in predicting what proportion of an audience will plan ahead to watch a program and

remain loyal to it, we might expect that programs with the most enthusiastic viewers would also develop the largest audiences. However, the size of a program's audience is affected by many circumstances other than a show's appeal -- factors such as the time of day, the season of the year, the availability of viewers, and the appeal of competitive offerings. A relatively unappealing program can still receive a high share of the total audience if it is aired on a night with weak competition. Conversely, a strong program up against strong competition may register a small share despite its high audience appeal.

We found no significant relationship between the size of a program's audience and the amount of satisfaction viewers reported deriving from the show. When we plot the relationship between the size of the audience and the average Appeal scores for the prime-time programs in our prototype rating study, we can see how widely Program Appeal scores varied among programs with similar size audiences. (The size ratings from our study are based on "person" rather than "household" viewing estimates.) (See Figure 3.)

For example, programs with audience size ratings of 12 had Appeal scores that ranged from the low 50s to above 80, and some of the programs with the smallest audiences in our study received the highest Appeal scores. If audience size were a reflection of Program Appeal, we would see almost all points on this chart

clustering around a rising diagonal that showed higher scores correlating with higher ratings. Instead we see a dispersed mass, with neither factor related to the other.

The traditional size ratings are not intended as measures of audience satisfaction. The oft-repeated adage "If they watch it, they like it" ignores the remarkable amount of variation in viewers' appreciation of programs. The Appeal Index, on the other hand, because it is not subject to external variables such as the competing menu of programs, provides a more reliable measure of a programs' intrinsic entertainment value. This index in turn is a more useful predictor of how likely audiences are to watch the program again and how likely they are to plan ahead to do so.

PROGRAM IMPACT: A MEASURE OF VIEWERS' INVOLVEMENT

Our second audience response measure, the Program Impact Index, indicates the degree of emotional and intellectual stimulation that a program provides its audience. As the other major component of our qualitative ratings system, it offers a means of assessing how involved or distracted viewers are while they watch television.

It is not necessary to discuss the Impact Index at the length devoted to Appeal because Impact shares many of the same characteristics. Impact scores also vary by audience subgroup; they discriminate among programs regardless of genre, and they remain stable from episode to episode for the majority of programs.

However, the Impact Index is more sensitive than the Appeal Index in assessing the effect a program has on its viewers and correlates more strongly with peoples' behavior while viewing and with commercial exposure and effectiveness measures. Before we consider the relationship between Program Impact and viewer behavior, however, we look at the derivation of the Impact Index and then at how Impact and Appeal are related.

Basis of Program Impact scale: "I learned something from it" and "It touched my feelings"

Different television programs appeal to their viewers for different reasons, but the most commonly stated reason in our research was that they offered a "pleasant way to pass the time." Seventy percent of the participants in one of our preliminary studies placed the program they were watching in this category. Like relaxing with an old friend, people watch certain shows confident about what to expect and generally pleased with what they find.

We learned that other programs, however, satisfy their viewers because they give them "something more" than this; these programs stimulate, compel, inform, or attract their audience more intensely. Thirty percent of the viewers in our preliminary study said that the program they had just seen was more than a way to pass the time. It was this indication of an added dimension to some people's viewing experience that led us to explore a second characteristic of a program's qualitative effect -- its emotional and intellectual impact on its audience.¹¹

In our prototype ratings study we asked our panelists to rate each program they viewed on the basis of how strongly it touched their feelings and how much they learned from it.¹² The vast majority of programs in our sample, either scored high on both of these valuation scales -- or they received relatively low marks on both scales. Occasionally, a program scored higher in one dimension than the other, but by and large, we found the emotional and intellectual impact scales sufficiently correlated for all programs to permit us to combine them into a single 0-100 index.¹³

Handwritten notes:
more...
...
2...
...
diff...

High Impact programs are highly appealing

When we compare a program's Impact score with its Appeal score we find a strong relationship between the two for most programs. In general, the higher a show's emotional and intellectual impact, the more the audience reports enjoying it, and programs with

Handwritten notes:
...
...
...

relatively low Impact scores generally receive lower Appeal scores as well.

However, Program Impact has a considerably wider range than Program Appeal, making it a more sensitive measure of the diversity of viewers' responses to different programs. Impact scores in our prototype ratings study ranged from a low of 15 to a high of 86 and the Impact scores for all programs we analyzed averaged only 44 compared with the much higher average Appeal score of 73. In general, then, we can state that most of the time people watch programs they enjoy -- hence, the high average Appeal score, but they don't always demand that those programs be involving -- hence, the wide range and lower average of Impact scores.

*Still
people like
watching
that kind
of material
but
good mean*

The Impact Index, like Appeal, distinguishes among programs regardless of genre

The Impact Index, we found in our analysis, holds up well as a measure of the inherent qualities of individual programs. While certain genres may be intrinsically more involving than others -- dramas, for instance, often rate above comedies and variety shows on the Impact scale -- the scale is fully capable of registering wide-ranging distinctions in the audience response to programs regardless of type. In many cases Impact scores had a wider range within a program type than did Appeal scores. Clearly, programs of many different types can involve as well as capture viewers.

The wider range of Impact scores is an indication that the Impact Index is more sensitive than the Appeal Index to differences in programs' content. This sensitivity caused the index to vary somewhat more often than Appeal scores did from episode to episode, as well. As with the Program Appeal Index, these differences were not caused by differing reactions to new episodes versus reruns, nor were they significantly affected by whether the program was an established series or a newly introduced one. Program Impact seems to reflect specific features of the program itself.

The fact that Program Impact is both more sensitive to episode content than Program Appeal and still strongly related to Appeal suggests that the measure may be useful in predicting the potential long-term appeal of new series. For example, Canadian and British researchers have found that it usually takes many weeks before the appeal of newly introduced series can be accurately estimated, because new shows generally have a heterogeneous audience comprised of many "experimenters" who are trying the show out for the first time. Some of these people undoubtedly will decide that the show is not for them and will not watch future episodes, but their presence in the audience for the first few telecasts may bring down a program's average score on an Appeal index. Several weeks usually are required for a new program to find its audience; the same period is needed for the program's appeal measure to find its level. Judging how much

emotional and intellectual impact a program has on its viewers requires no such wait, however; a programmer can measure this from the Impact score on the first showing. Since Impact correlates with Appeal for established series, new series episodes that are perceived as offering unusually high emotional or intellectual stimulation should generate increasing levels of audience enjoyment over time -- provided, of course, that they maintain their high levels of impact.

It seems clear from our analysis that despite the strong correlation between a program's Impact and its audience Appeal, the Impact Index measures a qualitative aspect of television programming that is distinct and separate from its entertainment value. It registers the level of a show's emotional and intellectual impact and offers important diagnostic information to cable and broadcast programmers interested in determining or improving the basis of a program's appeal.

As with the Appeal Index, however, the question then arises, what makes this separate measure worth recording? What useful information is there in the Impact Index of the television and advertising industries? Again the answer lies in the correlation of the measure with viewers' behaviors while watching. In addition, the Impact Index provides critical new information about the value of different programs as environments for advertising.

Handwritten notes:
1st
2nd

Handwritten notes:
3rd
4th
5th
6th
7th
8th
9th
10th

High-Impact programs capture more of viewers' attention

Today's television viewers are notoriously restless. Busy with home, school, or job and bombarded with an increasing array of media stimuli, many viewers go right on with their chores or amusements while a television program is on. In our prototype ratings study, we asked viewers to report how much attention they paid to the programs they watched. We found that, by their own report, television commanded peoples' undivided attention only 37 percent of the time, and almost 25 percent of our respondents reported giving "hardly any" or only "some" attention to the programs they were watching. Comparable findings were evident from all our other studies.

For an industry that spends millions of dollars developing advertising to persuade viewers to buy something, these low attention rates are cause for concern. However, even these self-reported attention rates may underestimate the distractions that the audience is subject to. In order to gain a clearer picture of viewers' actual attention while viewing, we interviewed participants in our prototype ratings study by telephone, asking them about their behavior during the course of the program they had just watched--specifically, whether they had left the room during the program or its commercials and what else they were doing while the set was on. The results of this survey confirmed findings from our earlier preliminary studies: over 50 percent of the respondents "watching" television reported that

they left the room at least once during the course of the show, and most of these people reported leaving repeatedly during both the show and the commercials. Furthermore, during the time when they were in front of the set, many "viewers" were engaged in conversation, household tasks, school work, reading and other leisure activities; in fact, many were doing several of these things at once. In all, approximately one-half of the viewing audience was engaged in other activities while watching television, and of these people over 40 percent said they were substantially distracted from television viewing by the other things they were doing.

Such high levels of viewer distraction tell us a great deal about television and how most people use it. Think how hard it is to read the sports page in the morning newspaper and carry on a conversation with your spouse about vacation plans; how difficult it would be to audit the business accounts while seated in a theater; how virtually impossible it is to read a novel or listen to a concert while talking to your mother on the phone. But apparently television is suffering from this kind of competition most the time. The question is, what, if anything, can the industry do about it? Viewers' attention to programs is influenced by a number of factors outside advertisers' and producers' control. Our study showed that the length of the program, the presence of young children in the home, and the age of the respondent all had a direct bearing on how much attention

a viewer could give to a specific show. Attention levels also varied somewhat according to the time of evening when the show was telecast. Programs scheduled early in the evening tended to receive less attention than those shown later at night in the hours when younger children are finally in bed and the demands of family life have slowed down.

But viewers' attentiveness to television programs is also influenced by the programs themselves. Despite the general findings of typically high levels of distraction and room-leaving, we also found that these behaviors were consistently lower among viewers who reported on their Television Viewing Monitors, for the same time period, substantial levels of emotional or intellectual stimulation as measured by the Impact Index. The higher the Impact score given to a program, the more likely the viewers of that program are to set aside distracting activities, to remain in the room, and to give the show their undivided attention.

Figure 4 shows notable differences in the behavior of people who find a program stimulating. While 44 percent of the viewers who scored the program they were watching low in Impact were doing something distracting while it was on, only 18 percent of those who rated a program high in Impact were distracted by something else. In the laboratory study, viewers who rated the program they watched high in Impact reported watching closely for 3 times

as many minutes as viewers who rated the show low in Impact. Furthermore, viewers rating a program high in Impact are more likely to remain in the room throughout the entire show. The significance of these findings for marketplace decisions is substantial, for not only do high-Impact programs capture more attentive and less distracted viewers, they deliver more of these viewers to the advertiser.

*Not
measured*

High Program Impact means greater commercial exposure

Standard television ratings tell advertisers and programmers how many sets are tuned to their program, but they do not tell whether anyone is in the room to watch or, if they are, whether anyone is actually looking and listening. Advertisers naturally want a wide audience, and the Nielsen and Arbitron ratings of audience size provide the basic information as to how large a viewership they are -- potentially -- reaching. Within that basic estimate of the audience, however, there exists for every television show another, smaller, more realistic and more effective market -- people who are actually present when the commercial is broadcast.

The Program Impact Index provides advertisers with a basis for estimating the size and composition of this second market. When external influences are controlled for, we find a strong correlation consistent across all program types between how emotionally and intellectually involved viewers are and the

Handwritten notes:
1. Not
2. Not
3. Not
4. Not
5. Not
6. Not
7. Not
8. Not
9. Not
10. Not

probability of their being present in the room during the commercial. (See Figure 5.)

These findings suggest that estimates of audience size alone do not give a true picture of a program's ability to deliver viewers for the commercial messages and that commercials placed within programs offering greater emotional or intellectual stimulation, as measured by the Program Impact Index, will be seen by a larger fraction of their nominal audiences.

Handwritten note:
W. J. ...

High Program Impact increases commercial effectiveness

Enhanced exposure of the audience to advertising is not the only commercial benefit of programs that are highly involving. Using a special experimental laboratory design we tested the Impact and Appeal Indexes for their influence on how viewers perceive commercials. We found that the Impact measure, in particular, was a good predictor of viewers' reactions to a program's commercial. Significant correlations show that viewers generally find commercials more memorable, likable, credible and persuasive when placed in a show they rate high in Impact than when placed in a show they rate low in Impact.

Handwritten notes:
1. Not
2. Not
3. Not
4. Not
5. Not
6. Not
7. Not
8. Not
9. Not
10. Not

-- Deleted Material --

This section of the paper draws on materials from Television Audience Assessment's forthcoming publication, Program Impact and Program Appeal: Qualitative Ratings and Commercial Effectiveness, which will be publicly released on October 29, 1984.

Figures 6-10 have similarly been deleted from this draft.

Handwritten notes:
The No. of views
seen during the
commercial
70

probability of their being present in the room during the commercial. (See Figure 5.)

These findings suggest that estimates of audience size alone do not give a true picture of a program's ability to deliver viewers for the commercial messages and that commercials placed within programs offering greater emotional or intellectual stimulation, as measured by the Program Impact Index, will be seen by a larger fraction of their nominal audiences.

Handwritten note:
W. J. ...

High Program Impact increases commercial effectiveness

Enhanced exposure of the audience to advertising is not the only commercial benefit of programs that are highly involving. Using a special experimental laboratory design we tested the Impact and Appeal Indexes for their influence on how viewers perceive commercials. We found that the Impact measure, in particular, was a good predictor of viewers' reactions to a program's commercial. Significant correlations show that viewers generally find commercials more memorable, likable, credible and persuasive when placed in a show they rate high in Impact than when placed in a show they rate low in Impact.

Handwritten notes:
doesn't viewer
...
What if this
...
in ad?

THE USES OF QUALITATIVE RATINGS

With the development and testing of the Appeal and Impact Indexes, Television Audience Assessment concludes its research and developmental work in creating a new qualitative ratings system for the American television industry. This research has established a number of propositions basic to the system:

- People do hold identifiable opinions about the television programs they watch, and their opinions vary from program to program because of perceived differences in how much enjoyment and stimulation programs provide.
- These audience opinions can be systematically measured and quantified in valid and reliable indexes, permitting an observer to rate and compare programs on the basis of their appeal and impact.
- Contrasts in the audience's response to television programs translate into differences in audience behaviors. The audience is more involved with and more loyal to some shows than to others, and the programs most likely to inspire viewers to plan ahead, to watch again, to set aside distracting activities, and to remain in the room and respond positively to the commercials can be identified by their Appeal and Impact ratings.

*as listed
document
list*

- These qualitative program ratings can be collected in an economical and methodologically sound manner.

These are provocative observations. They tell us that a new qualitative ratings system based on measures of Program Appeal and Program Impact is sound and viable. And they tell us that it holds substantial commercial promise.

A new tool for the changing television market

One simple rule applies to the marketing of consumer products: know your customers -- what they want, what they like, how and when they use your product. Our research into what is going on in front of American television sets underlines the importance of applying this principle to today's television consumer.

The syndicated ratings of audience size, which at present are the primary tool for consumer analysis in the television business, give the industry sales data on how much television is consumed and by whom, but they tell nothing about the consumers' evaluation of the product or their real use of it. As any good marketer knows, sales data alone are not sufficient to track and maintain the viability of a product in an increasingly competitive marketplace.

The financial success of the television industry over the past decades demonstrates that traditional size ratings were a sufficient marketplace tool in an era of a mass audience choosing among a few channels, watching a modest amount of television, and paying close attention. With only a few program offerings for the audience to choose from at any time, gradations in the audience's taste and behavioral response were largely irrelevant but today, choices for the television consumer are multiplying dramatically, and understanding the diversity of audience tastes and preferences has become more than pertinent; it has become essential for survival.

comment
for Impact
as supplemental
on the side
sales basis
is the
negative
correlation
with
positive
more with
comparing
D. 110

Qualitative television ratings begin to provide this understanding. They tell us in considerable detail what the audience thinks and feels about programs; and give us insight into differences in people's viewing behavior. The new information provided by Appeal and Impact ratings will greatly enhance general understanding about viewers' use of television -- understanding that is critical for efficient media planning, effective use of narrowcasting, and the development of programming to suit diverse consumer preferences in the television market of the 1980s and beyond.

Potential applications of Appeal and Impact Indexes

The full import of qualitative ratings for commercial decision-making will be apparent only when Appeal and Impact

ratings are regularly collected, reported, and used, but the broad outline of their expected utility has been suggested by the many industry executives we have spoken with over the course of our developmental work.

- Broadcast and cable programmers, for example, should find in viewers' evaluations of a show's appeal and impact information useful for their development, selection, scheduling, and promotional decisions. In general, the system will tell them what programs are most satisfying to different audience segments and will enhance their ability to develop competitive counterprogramming. Likewise, the use of qualitative ratings should help in the vertical scheduling of programming by identifying unserved audiences or served audiences that can be retained from one program to the next. The ability to understand the elements of Program Appeal and Impact and to locate audiences elsewhere in the schedule who are most likely to find a given program highly enjoyable and stimulating can also aid with decisions about program promotion.

Extensive research from England, where qualitative ratings have been collected and used for many years, suggests other potential applications for these data. In analyzing qualitative and quantitative program ratings, British researchers have found that an upward-trending appeal score

generally predicts a rise in that show's audience share over time. Conversely, a drop in appeal score is an early warning indicating "wearout" and eventual audience decline.¹⁵ These same researchers have found that for new shows, several weeks usually are required before a true reading of a program's appeal can be judged. Our research suggest that, unlike Appeal, the Impact Index may require no such wait, so an unusually high Impact rating for an otherwise slow starter could help programmers make more reliable tactical decisions.

- Cable and broadcast operators will be able to use qualitative data to demonstrate the relative value of their audiences to advertisers, due to the positive relationship between qualitative ratings and audience attentiveness and commercial exposure. Sales forces may find in these ratings a new and useful framework for pricing specific parts of their schedule based on each program's ability to satisfy and involve its audiences. And with this new resource, cable operators will have a reliable basis for determining the relative value of specific programs or program services in attracting and keeping subscribers.
- Program syndicators and distributors should find the Impact and Appeal data useful in forecasting the syndication performance of off-network programming. The data could help establish the value of syndication rights and help in selling

both off-network and first-run syndication properties to stations and cable operators. To the extent that syndicators and distributors produce first-run syndication programs, qualitative data will also help them to tailor their products to maximize audience enjoyment and involvement.

- Finally, advertisers and agencies seeking greater effectiveness for their advertising will find Television Audience Assessment's qualitative ratings system helpful. Since two programs with audiences of equal size and composition can differ in their Appeal and Impact and in the subsequent loyalty they inspire and attention they command, they can also vary in their efficiency as vehicles for commercial messages. Use of the new system will aid advertisers and agencies in determining which program environments hold the most promise for ensuring commercial exposure and improving effectiveness among target market segments. Careful examination of each program's Appeal and Impact scores can produce a road map for advertisers leading to those programs that can better retain viewership and heighten receptivity during commercial breaks. Qualitative ratings will also enable them to evaluate the efficiency of placing commercials in programs with relatively small or narrowly defined audiences. Furthermore, since qualitative ratings can aid in predicting audience planning and loyalty, advertisers may find it useful to advertise in highly

appealing programs regardless of their audience size when the goal is repeated exposures. Only qualitative ratings can provide the systematic means for assessing these differences in programs' performance.

In a highly competitive media environment where it is no longer possible to tell from the traditional ratings numbers who is actually watching and noticing and reacting to advertisers' messages, reliable and consistent qualitative program ratings can make an important contribution to sound market decisions. Program Impact and Program Appeal ratings, having been rigorously tested and consistently tied to commercially relevant audience behaviors, offer new promise for more incisive and better-informed judgments about how different programs perform in today's changing communications marketplace.

ENDNOTES

1. A.C. Nielsen Company reported in the Nielsen Television Index that daily television viewing for the 1983 calendar year was 7 hours, 2 minutes.
2. Television Audience Assessment, Inc., The Audience Rates Television (Cambridge, MA: Television Audience Assessment, Inc., 1983), p. 4.
3. Ibid.
4. Ibid.
5. Television Audience Assessment, Inc., The Multichannel Environment (Cambridge, MA: Television Audience Assessment, Inc., 1983), p. 13.
6. Setlow, C.E., "TV Ratings-There Just Might Be a Better Way," New York Times, December 31, 1978.
7. See, for example:

Corporation for Public Broadcasting, Boston (WGBH) Field Testing of a Qualitative Television Ratings System for Public Broadcasting, CPB, Office of Communications Research, 1981.

National Citizens Committee on Broadcasting, "Qualitative Ratings Project: Pilot Test Report," NCCB, March 1978.

Holt, S.A., Television Ratings: An Overview of Domestic and International Approaches (Cambridge, MA: Television Audience Assessment, Inc., 1980).
8. See Commercial Effectiveness and Viewers' Involvement with Programs: A Literature Review, prepared by Clancy, Shulman & Associates, Inc. (Boston: Television Audience Assessment, Inc., 1984.)
9. For a further discussion of the relationship between audience appreciation and program type, see: "A Box for All Seasons-The Ebb and Flow of Television Viewing and Appreciation Over a Whole Year, 1980," Audience Research Department of the Independent Broadcast Authority, ND.
10. See, for example:

Barwise, T.P., Ehrenberg, A.S.C., Goodhardt, G.J., "Viewers' Average Appreciation Scores-An Explanatory Study," London Business School, Thames Polytechnic and Aske Research



Limited, 1980. This study, undertaken in Great Britain, concluded that differences in their Appreciation Index scores largely reflect differing responses to programs by viewers, rather than some kind of audience segmentation with high and low scoring viewers watching different programs.

- 11. Other audience researchers also attempt to measure this added dimension of programming. For example, the Canadian Broadcasting Corporation uses a seven-item scale as the basis for a "relaxing/demanding" index on which they rate all television programs broadcast in Canadian markets. See, for example: "Some Insights from Canada," Studies in Television, prepared for the Independent Broadcasting Authority by Aske Research Ltd., November 1981. With few exceptions, in this system, information programs score high in the "demanding" direction, while entertainment programs receive higher "relaxing" scores. While this relaxing/demanding index may be useful in gauging the level of effort different programs require from their viewers, it does not permit analysis of the variations in the impact of the whole range of entertainment programs. Our specially designed Impact Index, however, measures differences in audience involvement regardless of what kind of program is being considered.
- 12. The Impact Index is based on summing viewers' responses to two four-point scales in the Television Viewing Monitor: "This program touched my feelings" (not at all, only a little, some, and very much); and "I learned something from this program" (not at all, only a little, some, and very much). For purposes of this analysis, the scale points were treated as equal intervals and values were assigned to result in another 0-100 scale.
- 13. There is a strong positive correlation between viewer's ratings of a program's intellectual and emotional impact; for example, of all the respondents who gave a scale rating of "not at all" when asked if a program touched their feelings, 70 percent gave the same rating when asked if they learned something from that program. Similarly, of all those who reported a program "touched my feelings very much," 61 percent reported they also "learned very much" from that program.

		"Program touched my feelings"			
		Not at all	Only a little	Some	Very much
"Learned something from program"	Not at all	70%	32%	12%	3%
	Only a little	18	41	24	6
	Some	10	24	53	30
	Very much	2	3	11	61
		100%	100%	100%	100%

14. The constant-sum method asks respondents to estimate the proportion of past and future purchases given to specific brands. In our study, the pre-exposure question was: "Out of the last 10 times you made a purchase in each of the product categories, please tell how many times you purchased each brand. Note that the numbers have to add up to 10." The post-exposure question was: "Thinking about the next 10 times you will make a purchase in each of these categories, please tell how many times you think you might purchase each brand. Note that the numbers have to add up to 10."

For a discussion of the advantages of the constant-sum technique in predicting purchasing behavior, see Axelrod, Joel N., "Attitude Measures that Predict Purchase," Journal of Advertising Research, Vol. 8, No. 1, pp. 3-17.

15. Maneer, P., "Towards TV Qualitative Research: The BBC Viewpoint," BBC Broadcasting Research Department, March 1981.

APPENDIX B

Figure 1

Appeal of programs is consistent whether shown in prime-time or "strip" format.

Program	Format	Appeal Score
Barney Miller	strip (12 midnight, Monday-Friday)	77
	prime (9pm, Thursday)	73
Happy Days	strip (6pm, Monday-Friday)	72
	prime (8pm, Tuesday)	68
Jeffersons	strip (6pm, Monday-Friday)	76
	(7pm, Monday-Friday)	73
	prime (9:30pm, Sunday)	74
Laverne & Shirley	strip (7pm, Monday-Friday)	57
	prime (8:30pm, Tuesday)	64
M*A*S*H	strip (7pm, Monday-Friday)	79
	(11:30pm, Monday-Friday)	81
	prime (9pm, Monday)	81

Figure 2

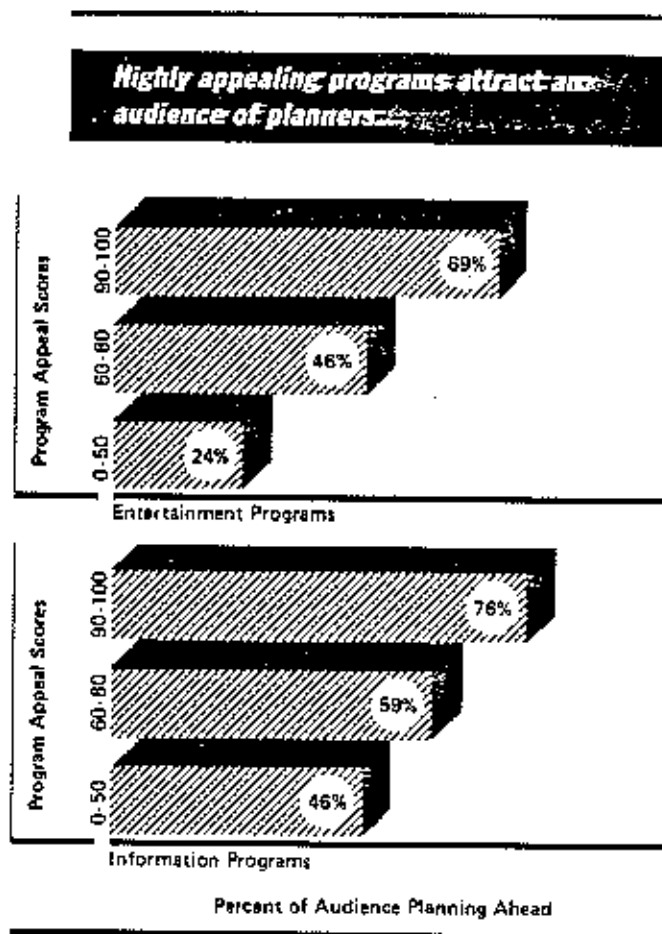
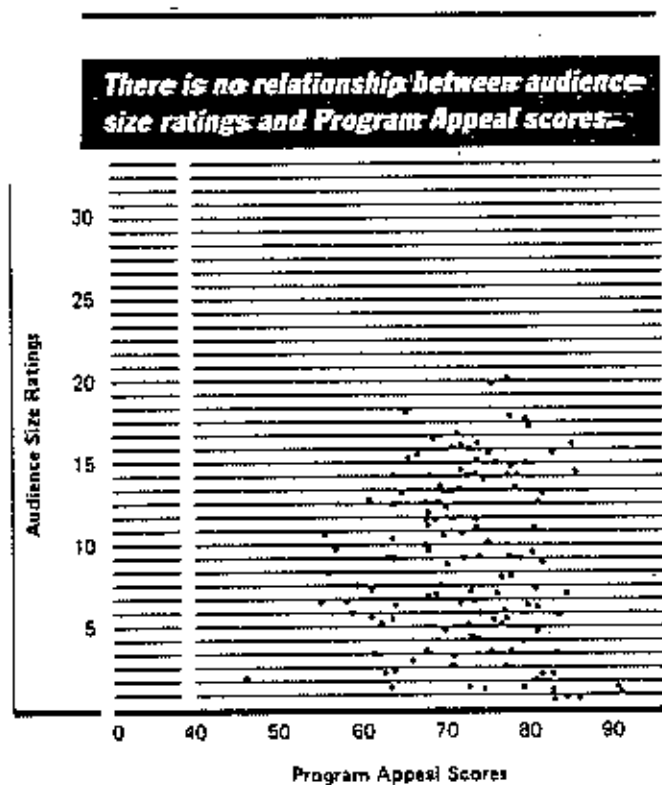


Figure 3



Includes regular prime-time programs; excludes sports, movies, news.

Figure 4

High-Impact programming reduces viewer distractions.

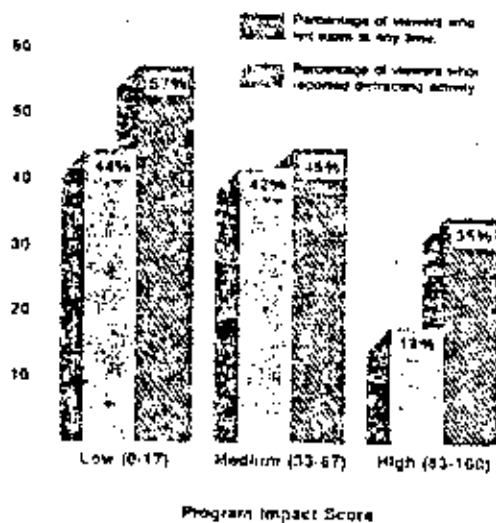


Figure 5

2. High-Impact programming inspires viewers to remain in the room during commercials.

