

Television Sound and Viewer Perceptions

W. Russell Neuman, MIT
Ann N. Crigler, USC
V. Michael Bove, MIT

This study examines the impact of changes in audio quality on viewer perceptions of the television viewing experience. A sample of 367 mass audience viewers recruited from a shopping mall were randomly assigned to conditions which included low fidelity vs. high fidelity, monaural vs. stereo and three types of television programming. Viewers watched a series of short video clips with varying audio quality and asked to rate their liking, level of interest and psychological involvement in the programming and asked to evaluate picture and audio quality.

Our mass audience subjects did not prove to be very discriminating. They had a difficult time distinguishing mono from stereo and even low fidelity from high fidelity sound. In a mono-stereo comparison 41% preferred stereo, 34% preferred mono and 24% heard no difference. In the low fidelity-high fidelity test 43% preferred high fidelity, 27% preferred low fidelity and 30% heard no difference. The strongest differences in preference were apparent in comparing low fidelity mono to high fidelity stereo television audio. We conclude that high fidelity television audio and stereo imaging need to be introduced in tandem in order to have significant psychological impact on untrained audience members.

As an interesting sidelight, we discovered that video with better quality and stereo sound were consistently rated as more likeable, interesting and involving. Viewers also rated programming with better audio as having higher picture quality, but this occurred in only one of the three test programs we ran.

The purpose of this study is to examine the impact of changes in audio quality on viewer perceptions of television, particularly on viewers' evaluations of content and technical quality. Low-fidelity, high-fidelity, monaural, and stereo audio were examined for three kinds of television programming: sports, situation comedy, and action-adventure.

The research questions focus on evaluations of programming content as well as on evaluations of differences in the technical quality characteristics. How strong are preferences, if any, for television with high-fidelity and stereo sound? How do audio fidelity, stereo separation, and programming content interact? Does improved quality audio affect perceptions of video quality? Does improved audio quality affect evaluations of liking, interest and involvement with the programming content?

The audio signal that accompanies an NTSC television image is transmitted on an FM subcarrier 250 kHz from the top of the broadcast channel. The characteristics of the audio signal have changed very little since the standard was set, despite the improvements that have been made in radio audio quality in the same period. Stereo radio broadcasts became widespread in the 1950s, at about the same time that television adoption was growing rapidly.

In December of 1983, television broadcasters and equipment manufacturers agreed on a standard for multi-channel television audio. With FCC acquiescence, the agreement has become a de facto standard. The new standard includes a companded L-R difference signal on an audio subcarrier of twice the horizontal line rate. The dbx companded difference signal permits the audio channel to achieve a higher signal-to-noise ratio than would otherwise be possible with the limited frequency deviation permitted in the television audio subcarrier channel. The

addition of the stereo difference signal does not affect the audio signal received by audience members with mono-audio sets. With the help of modern semiconductor technology, the new standard yields an inexpensive and compatible potential improvement of the television watching experience. Is that potential being realized and does the viewing/listening audience notice or appreciate the differences in audio quality?

RESEARCH DESIGN

This section details the research design issues that were important in the implementation of the television audio quality comparison tests. Both the technical and physical set-up of the experiment and the logistics of recruiting and conducting the experiment are explained below.

Physical Set-up

The Audio comparison tests were performed using a single 19 inch NEC consumer color monitor with hi- and lo-fidelity speaker systems. Figure 2 illustrates the setup. Acoustically-transparent black cloth masked the speaker systems from the subjects' view. Research staff members in a control room behind the subjects controlled the image and the audio quality delivered to the audio room. The program clips were recorded on one-inch tape and played on an Ampex VTR. The "low-fidelity" speakers were inexpensive Radio Shack models (\$20 each); ADS model L880 speakers (retailing at about \$400 each) provided the "high-fidelity" audio.

Research Method

This study examined the impact of variations in audio quality on viewer perceptions of television

programming. Viewers were asked to evaluate segments of Miami Vice, Cheers, and the 1986 All-Star Baseball Game. We gathered data as well on music programming, a live performance of Tina Turner produced for HBO, but because of sampling problems we have set that data aside. The patterns of response for the musical programming were quite similar to those reported here. Audience members were recruited at our Mall facility and brought into the audio room one at a time. Each subject was told only that the study was sponsored by the television industry and that he or she would be asked to evaluate a few programming clips. There were approximately 120 subjects in each of the three content conditions (about 30 per audio condition for each tape) for a grand total of $N = 367$.

The study consisted of two parts: first a conservative test in which the viewers were asked to watch and evaluate a program segment and second, a test in which respondents were asked to compare two otherwise identical 30-second clips for which the audio quality was systematically varied.

The Conservative Test

First, each respondent was asked to watch a 2 1/2 minute television clip of one content type in one of the four audio conditions. They were then asked to respond to three standard seven-point evaluation scales tapping the dimensions of liking, interest, and involvement. No mention of audio quality was made at this point. We have dubbed this the "conservative test" because any differences in evaluations would be a result of uncued respondents focusing primarily on program content rather than on technical variables. We then asked respondents to rate the picture quality, audio quality and "overall quality of the TV set" for the same clip on 5-point scales.

The Comparison Tests

In the second part of the study, each respondent was asked to compare two otherwise identical 30-second clips for which the audio quality differed. A sequential A/B comparison test was used in order to vary as few set characteristics as possible. This procedure included the explicit comparisons of all combinations of the four audio conditions. Respondents were then asked if the first clip was better, the second clip was better, or whether there were no differences. The question was repeated again with specific reference to picture and sound quality. In addition, each subject filled out a background questionnaire.

ANALYSIS

Difference of means tests (both analysis of variance and t-tests) were used to assess respondents' evaluations of program content and technical characteristics. Analyses were conducted on three combinations of the sound conditions, first, for the entire sample across the four conservative test audio conditions

(lo-fi monaural, hi-fi monaural, lo-fi stereo and hi-fi stereo,) second, between monaural and stereo sound (mono lo-fi and mono hi-fi vs. stereo lo-fi and stereo hi-fi,) and third, between low fidelity and high fidelity sound (lo-fi mono and lo-fi stereo vs. hi-fi mono and hi-fi stereo.)

In addition to examining the differences across audio conditions for the entire sample, several audience member and programming content characteristics that might influence subjects' perceptions and evaluations of stereo TV were tested. For example, the particular show watched might be more or less suited to stereo separation. A baseball game, for instance, may be more involving in high-fidelity stereo than in low-fidelity monaural sound because of the ambient sounds of the ball park crowds.

We also took into account the age of the respondent and his or her television preferences and usage habits. Are people who like television and heavy viewers more or less likely to appreciate improvements in audio quality than people who dislike or seldom view television? If subjects own new or high quality television sets or other high quality home media equipment are they more likely to prefer or notice differences with the enhanced audio condition?

Findings

Significant differences on the conservative tests in mean evaluation scores of content and technical characteristic variables were strongest when comparing extreme audio conditions. Respondents in the Stereo-High Fidelity condition liked the program content significantly more and found it significantly more involving than respondents in the Mono-Low Fidelity condition. In addition, respondents in the Stereo-High Fidelity condition rated the sound quality significantly higher than respondents in the Mono-Low Fidelity condition.

When only one aspect of the audio was varied -- either fidelity or separation -- respondents showed less dramatic powers of discrimination. There were few significant differences among respondents across the four audio conditions on variables evaluating either content or technical characteristics. Similarly, respondents who shared behavioral or demographic characteristics, such as frequency of TV viewing or age, also generally failed to distinguish among audio conditions. It appears that, overall, the impact of audio variables are consistent but small enough that it did not reach appropriate levels of statistical significance except when aggregated over all three types of programming.

Thus we note cautiously that the overall mean score of the high fidelity audio was consistently higher than the mean score of the low fidelity audio on both content and technical evaluations. This pattern holds among all sub-groups of respondents. Respondents consistently evaluated television images presented with stereo or high fidelity sound as more interesting, more involving and better liked than the same images presented with mono or low fidelity sound. Likewise, respondents

rated the sound quality, picture quality and overall quality of television presented in stereo or high fidelity as superior to mono or low fidelity television.

One group of respondents who demonstrated particularly strong (and significant) patterns were those subjects who own high quality television sets. These respondents liked programming with stereo sound more and found it both more involving and more interesting than programming with mono sound. This is particularly interesting because it suggests a pattern of particularly discriminating tastes among early adopters. The question is raised whether such a group represents a small outlier subset or whether their tastes and discriminatory capacities will influence other consumers and ultimately help to "educate" their friends, relatives and neighbors.

Results from the comparison tests, in which side-by-side comparisons were made between identical visual and varied audio conditions, indicate that

respondents are willing to report the overall quality and sound quality of one set as better than another, but do not provide clear support for the hypothesis that respondents prefer "high" quality over "low" quality sound. Less than one-third of the sample reported no difference between the two clips when asked "Which was better?", or when asked to compare the sound quality of the two sets. Thirty-two to 44 percent of the sample reported a difference in picture quality between the two sets.

Across the entire sample, the mean score for sound quality was significantly different from zero only when comparing low to high fidelity, both overall and within stereo and mono conditions. Means on other comparisons and on other variables were not found to be significantly different from zero, though they tended to be positive rather than negative. As with the conservative test, there was not significant variation in the ratings for different groups of respondents.

Figure 1
The Conservative Test

	<u>Mono</u>	<u>Vs.</u>	<u>Stereo</u>	<u>Lo Fi</u>	<u>Vs.</u>	<u>Hi Fi</u>
Like Programming	4.6		4.8	4.5		4.8
Program Involving	4.1		4.4	4.0*		4.4*
Picture Quality Evaluation	4.2		4.3	4.2		4.3
Sound Quality Evaluation	4.3		4.4	4.3*		4.5*

	<u>Lo Fi Mono</u>	<u>Vs.</u>	<u>Hi Fi Stereo</u>
Like Programming	4.4*		5.0*
Program Involving	3.8**		4.6**
Picture Quality Evaluation	4.2		4.4
Sound Quality Evaluation	4.3*		4.6*

Means on 7-point evaluation scales, higher score = positive evaluation

* T-Test Significant at .05

** T-Test Significant at .01

Figure 2
Comparison Test

	<u>Mono Preferred</u>	<u>No Difference</u>	<u>Stereo Preferred</u>
Overall	34%	24%	41%
Miami Vice	28%	23%	49%
Cheers	33%	35%	31%
Baseball	31%	16%	53%

	<u>Lo Fi Preferred</u>	<u>No Difference</u>	<u>Hi Fi Preferred</u>
Overall	27%	30%	43%**
Miami Vice	26%	22%	52%*
Cheers	24%	29%	47%*
Baseball	31%	40%	29%

* Significant at .05
** Significant at .01