

Commercials can also be tested in focus groups, shopping center intercepts, and auditorium-type situations. Leading companies and advertising agencies rarely show their commercials on television until they are tested in a variety of situations. The companies and their agencies do not want to communicate the wrong message to the audience.

Music Research

Music is the product of a music radio station, and failing to analyze the product invites ratings disaster. To provide the radio station's listeners with music they like to hear and to avoid the songs that they do not like or are tired of hearing (*burned out*), radio programmers use two primary research procedures: auditorium music testing and call-out research.

Auditorium tests are designed to evaluate *recurrents* (popular songs of the past few years) and *oldies* (songs older than 10 years). **Callout research** is used to test music on the air (*currents*). New music releases cannot be tested adequately in the auditorium or with callout procedures because it is necessary to play the entire song. New music is often tested on the air during programs with titles such as *Smash or Trash*, where listeners call in and voice their opinion about new releases. Sometimes new music is tested in focus group situations where the respondents can listen to the entire song.

Auditorium tests and callout research serve the same purpose: to provide a program director or music director with information about the songs the audience likes, dislikes, is tired of hearing, or is unfamiliar with. This information allows the program director to make decisions based on audience reaction rather than gut feelings or guessing.

Both music testing methods involve playing **hooks** (short segments) of several

songs for a sample of listeners. A hook is a 5- to 10-second (usually 8-second) representative sample of the song—enough for respondents to identify the song if it is already familiar to them and to rate the song on some type of evaluation scale. Our experience indicates that respondents who are familiar with a song can identify it in three seconds or less.

Research companies and program directors have a variety of scales for listeners to use in evaluating the music they hear. For example, respondents can be asked to rate a hook on a 5-, 7-, or 10-point scale, where 1 represents "hate" and 5, 7, or 10 represents "like a lot" or "favorite." Or the scales can be used without labels, and respondents are instructed to rate the songs on the scale, where the higher the number, the more the song is liked. There are also options for "unfamiliar" and "tired of hearing." (A respondent who is unfamiliar with a song does not rate the song.) Which scale is best? Research conducted over several years by the senior author of this book indicates that either a 7-point or 10-point scale provides the most reliable results.

Sometimes researchers ask respondents to rate whether each song "fits" the music they hear on their favorite radio station. This additional question helps program directors determine which of the tested songs might not be appropriate for their station. In addition, some research companies ask listeners whether they would like radio stations in the area to play a particular song more, less, or the same amount as they currently do. This is an inefficient and inaccurate way to determine the frequency with which a song should be played. The reason is that there is no common definition of *more*, *less*, or *same*, and listeners are poor judges of how often a station currently plays the songs.

The bottom line in all music testing is that program directors should use the data as a guide for selecting their radio station's music, not as a music selection "bible."

Auditorium Testing. In this method, between 75 and 200 people are invited to a large room or hall, often a hotel ballroom. Subjects always match specific screening requirements determined by the radio station or the research company, such as listeners between the ages of 25 and 40 who listen to soft rock stations in the client's market. Respondents are usually recruited by a field service that specializes in recruiting people for focus groups or other similar research projects. Respondents are generally paid between \$25 and \$150 for their cooperation. The auditorium setting—usually a comfortable location away from distractions at home—allows researchers to test several hundred hooks in one 90- to 120-minute session. Usually 200 to 600 hooks are tested, although some companies routinely test up to 800 hooks in a single session. However, after 600 songs subject fatigue becomes evident by explicit physical behavior (looking around the room, fidgeting, talking to neighbors), and statistical reliability decreases. It is easy to demonstrate that scores for songs after 600 are not reliable (Wimmer, 2001), specifically in reference to unstable standard deviations for the songs.

While auditorium music testing is a relatively simple project, there are many things to consider to ensure that the test is successful. Among numerous procedures and steps, some basic procedures to follow when conducting an auditorium music test include the following:

The key to a successful test is a good introduction that explains the purpose of the test and how important the respondents' answers are to hearing good music on the radio. It is important to stress that there are no right or wrong answers in the test and that the goal is to collect a variety of opinions.

The moderator must be in total control over the situation to ensure that

respondents do not talk amongst themselves or try to influence other respondents' answers.

Adequate breaks must be taken during the session. Respondents shouldn't listen to more than 200 songs without a break.

The moderator must make sure that all respondents understand the scoring system. After the test begins, the moderator should check to see that each person is rating the songs correctly.

The moderator should not allow the respondents to sing along with the songs. This disrupts the other respondents.

The moderator must always expect the unexpected, including such things as electrical outages, sick respondents, unruly respondents, equipment failures, or problems with the hotel arrangements.

In the late 1990s, a few radio research companies developed alternative approaches to the firmly established (reliable and valid) auditorium music testing methodology. One method is to test music hooks over the telephone; the other tests hooks via the Internet. Both methods face the same problem of relinquishing control over the testing situation, and there is no way to know who is actually rating the songs. While the authors' experience shows that telephone and Internet music testing should not be used because there is no publicly available research evidence to support that it is reliable and valid, we know that many program directors and general managers use them anyway. The only advice we have is *caveat emptor*.

Callout Research. The purpose of callout research is the same as that of auditorium testing; only the procedure for collecting the

data is changed. Instead of people being invited to a large hall or ballroom, randomly selected or prerecruited subjects are called on the telephone. Subjects are given the same rating instructions as in the auditorium test; they listen to the hook and provide a verbal response to the researcher making the telephone call. Callout research is used to test only newer music releases. (The reliability and validity of using the telephone for call-out research is better than using the telephone for auditorium testing because only a small number of songs are tested in callout.)

While callout methodology is adequate because only a few songs are tested, the limitation on the songs tested is also the methodology's major fault. Well-designed callout research involves testing a maximum of 20 songs because subject fatigue sets in quickly over the telephone. Other problems include the distractions that are often present in the home, the poor quality of sound transmission created by the telephone equipment, and the fact that there is no way to determine exactly who is answering the questions.

Even with such limitations, many radio stations throughout the country use callout research. Since callout research is inexpensive compared with the auditorium method, the research can be conducted frequently to track the performance of songs in a particular market. Auditorium research, which can cost between \$20,000 and \$40,000 to test approximately 600–800 songs, is generally conducted only once or twice per year.

Programming Research and Consulting

Several companies conduct mass media research. Although each company specializes in specific areas of broadcasting and uses different procedures, they all have a common goal: to provide management with data to use in decision making. These companies offer custom research in almost any area of

broadcasting—from testing call letters and slogans to air talent, commercials, music, importance of programming elements, and the overall sound of a station.

Broadcast consultants can be equally versatile. The leading consultants have experience in broadcasting and offer their services to radio and television stations. Although some of their recommendations are based on research, many are based on past experience. A good consultant can literally “make or break” a broadcast station, and the task of a consultant is probably best described by E. Karl (2009), a former leading international radio consultant who was asked to describe what a consultant does for a radio station. He said:

A consultant works with research data to help a station's management team design its overall strategy. A consultant puts research information into a package that will make sure the target audience's most important programming elements are on the air, and that the station is positioned correctly in listeners' minds. A consultant also helps market the station to attract listeners to try out the station. The consultant does anything from designing music rotations, creating “clock hours” on the station, and selecting air talent . . . to developing television commercials to advertise the station, executing direct marketing campaigns to ask listeners to listen, designing website content, and working with the station staff to make sure the “promise” of the station's position stays on track.

Performer Q

Producers and directors in broadcasting naturally want to have an indication of the popularity of various performers and entertainers. A basic question in the planning stage of any program is this: What performer or group of performers should be used to give the show the greatest appeal? Not unreasonably, producers prefer to use