Guidelines for Formative Evaluation Research in Campaign Design

Charles K. Atkin and Vicki Freimuth

Public communication campaigns have achieved a varied record of success in influencing health and prosocial behavior as reflected by the diverse array of cases cited in this book. Designing and implementing effective campaigns requires a disciplined approach where the campaign team performs a thorough situational analysis, develops a theory-based but pragmatic strategic plan, and implements the creation and placement of messages in accordance with principles of effective media campaign practices. Moreover, diligent efforts are needed to enhance the working relationship between campaign designers and evaluators versus creative professionals who translate concepts in messages. A key role of the strategist in the collaborative process is to develop a framework for setting specifications and providing feedback as messages are prepared. Using a research-based approach in the public service domain is challenging when the mind-set of personnel in sponsoring organizations entails rigid advocacy of unpalatable, ideal behavior, devotion to politically correct message content, and self-indulgent artistic expression. Furthermore, specialists in domains such as health, environment, or altruism aren't always conscious of the fact that they differ substantially from their audiences in topical knowledge, values, priorities, and level of involvement, so they lack the perspective of the "average" person. Research data from samples of the intended audiences can help overcome the gulf between sender and receiver (see Dervin & Foreman-Wernet, Chapter 10).

Over the life of a campaign, evaluation research encompasses collection of information about audiences at the *formative* stage, followed by *process* evaluation to assess implementation as the campaign unfolds, and finally *summative* evaluation to track campaign impact (see Valente & Kwan, Chapter 6; Salmon & Murray-Johnson, Chapter 7). Formative evaluation research inputs can enhance campaign effectiveness by guiding the development of sophisticated strategies and effective messages. This preliminary phase of research is useful for determining which approaches are most promising and for revealing whether certain components are ineffective or even counterproductive. According to Palmer (1981),

formative research provides data and perspectives to improve messages during the course of creation. He divides this type of evaluation into two phases. The first involves *preproduction research,* "in which data are accumulated on audience characteristics that relate importantly to the medium, the message, and the situation within which the desired behavior will occur" (p. 227). The second type of formative research is *production testing,* also known as pretesting, in which draft prototype messages are evaluated to obtain audience reactions prior to final production.

Public communication campaigners increasingly utilize elaborate formative evaluation techniques, particularly for major campaigns sponsored by government agencies, foundations, and organizations promoting health and social progress. Reviews of the media health campaign literature conclude that formative evaluation contributes to more successful campaigns (Noar, 2006; Rogers, 1996). Recent exemplars of sophisticated formative research include Bauman, Smith, Maibach, and Reger-Nash (2006); Berkowitz, Huhman, Heitzler, Potter, Nolin, and Banspach (2008); Cho and Witte (2005); Parrott, Steiner, and Goldenhar (2008); Smith, Atkin, Martell, Allen, and Hembroff (2006); and Uhrig, Bann, Wasserman, Guenther-Grey, and Eroglu (2010).

Information about the audience is most often utilized to identify specialized subgroups to be reached, to devise message appeals and presentation styles, and to select sources and channels. Furthermore, the formulation of campaign goals and objectives is increasingly based on research identifying priority areas of concentration, prospects for attaining certain types of impact, and critical stages of the communication process that must be addressed.

The campaign designer must adeptly overcome audience resistance manifested as receivers progress through exposure to processing to learning to yielding to behaving. Perhaps the most elemental problem is reaching the audience and engaging attention to the messages. Other key barriers include underestimating susceptibility to threats, counterarguing persuasive appeals, displaying reactance to compliance attempts, and exhibiting inertia (Knowles & Linn, 2004; McGuire, Chapter 9). Campaign designers are vigilant of unintended side effects that undermine campaign objectives (see Salmon & Murray-Johnson, Chapter 7). Concerns about boomerang effects especially shape the selection and presentation of behavioral recommendations, negative persuasive incentives, and source messengers (see Hornik, Chapter 3). Avoiding counterproductive responses from the audience requires careful formative evaluation inputs, both preproduction research and production testing.

CONCEPTUAL FRAMEWORK FOR CAMPAIGN PERSUASION

Communication–Persuasion Matrix

A fundamental organizing framework for developing campaign strategies is McGuire's matrix, which arrays source, message, channel, receiver, and destination variables as inputs and a series of information-processing and response variables as outputs (2001, Chapter 9). The first three input components are subject to manipulation by the campaign

designer. The *source* concept includes both the organization that sponsors the campaign and the messenger who delivers the message, which can be characterized in terms of demographics (age, sex, socioeconomic status), credibility (expertise, trustworthiness), and attractiveness. Each *message* may feature a variety of content dimensions (themes, appeals, claims, evidence, and recommendations) using various formats of arranging material and styles of packaging, while the overall series of messages in a campaign can vary in volume, repetition, prominence of placement, and scheduling. The *channel* variables comprise the basic medium of transmission (e.g., television, social media sites), the content modes (e.g., news item, PSA), and the particular media vehicles (e.g., specific radio station, magazine title, website).

Although *receiver* factors are not subject to manipulation, sensitivity to the background attributes, abilities, and predispositions of individuals enhances the effectiveness of campaign stimuli. Finally, the *destination* encompasses the array of impacts that the campaign aims to produce, such as immediate versus long-term change, prevention versus cessation, direct versus two-step flow of influence, and intermediate responses versus ultimate behavioral outcomes.

The output variables have been conceptualized in a number of ways, typically beginning with exposure and processing, followed by the hierarchy of cognitive, affective, and behavioral consequences of the campaign inputs, such as skill acquisition, attitude change, and decision to act. Formative research can help identify what types of content and style will attract audience attention, facilitate comprehension, elicit emotional reactions and elaborations, impart knowledge gain and skills acquisition, influence the formation or change of affective orientations such as beliefs and attitudes, and affect pertinent behavioral performance.

Key Theoretical Approaches

Formative evaluation draws upon concepts and influence processes from theories, models, and frameworks in communication, social psychology, marketing, and health education. A number of key perspectives described in Chapter 1 provide pertinent guidance regarding the types of information researchers should collect to facilitate media campaign strategies.

Public service campaigns differ from other leading forms of media content, such as news and entertainment, because the messages are purposively focused on knowledge, attitudes, and behavior. In many respects, though, campaigns in the health and prosocial domains are similar to commercial advertising campaigns. Thus, it is useful to adapt concepts from the *social marketing* framework, which emphasizes an audience-centered consumer orientation and calculated attempts to attractively package the social product and utilize the optimum combination of campaign components that will attain pragmatic goals (see Bracht & Rice, Chapter 20; Rice & Robinson, Chapter 16). Social marketing offers a macro perspective, combining numerous components to be assessed at the formative stage, notably the multifaceted conceptions of "product," costs, and benefits (particularly nonsubstantive incentives), as well as audience segmentation, policy change, and competition.

The *uses and gratifications* perspective helps isolate audience motivations for attending media messages and for utilizing learned information in enacting behaviors. Processing concepts are drawn from the *cognitive response* model that emphasizes how audience involvement shapes thought generation and central versus peripheral routes to persuasion. The *instrumental learning* perspective focuses attention on factors such as source credibility, incentives featured in message appeals, and repetition.

Framing theory suggests careful appraisal of packaging message appeals in terms of gain-frame promotion of positive behavior versus loss-frame prevention of negative behavior. The concept of salience is central to the *agenda setting* perspective, particularly related to the prominence of societal problems and importance of issues, and to *media priming,* where certain cognitions associated with a health or social behavior result in short-term activation effects.

Social Cognitive Theory emphasizes the importance of source role models, explicitly demonstrated behaviors, depiction of vicarious reinforcements, and self-efficacy. Expectancy–value formulations, particularly the *Theory of Reasoned Action* and the *Theory of Planned Behavior,* focus on beliefs about the likelihood that performing a behavior leads to certain consequences and highlight the role of subjective norms. Several concepts from the *Health Belief Model* and *Protection Motivation Theory* pertain specifically to health threat appeals: susceptibility multiplied by seriousness of consequences and the self-efficacy and response efficacy of performing behavior. The *Transtheoretical Model* identifies key stages of change for segmenting audiences and for determining readiness to attempt, adopt, or sustain the recommended behavior. The *diffusion of innovations* theory suggests examining the relative advantage and trialability of focal behaviors and allocating media messages to stimulate social influence via multistep flows.

PREPRODUCTION RESEARCH

During the preproduction stage, the strategist seeks to learn as much as possible about the intended audience before articulating goals and developing strategies. Specifically, the research helps identify intended audiences and focal behaviors, specify significant intermediate response variables, ascertain channel exposure patterns, and determine receptivity to potential message components.

To collect pertinent preproduction information, researchers utilize focus groups, surveys, and secondary analysis. *Focus group* sessions are conducted by a moderator who stimulates extensive, open-ended discussions of selected issues in a small group setting (this technique is described in the pretesting section). The *Sense-Making Methodology* (Dervin & Foreman-Wernet, Chapter 10) applies more in-depth interviews and observations. Quantitative data are most often generated via formal surveys of midsize samples representing intended audience segments; standardized items are used to systematically measure a broad array of variables via interviews or questionnaires. In the case of campaigns featuring individually customized message tailoring, researchers utilize a unique type of formative research where the actual audience first completes an online survey instrument measuring numerous variables (Rimer & Kreuter, 2006). The computer

program then matches this individual assessment input with appropriate portions of content drawn from a database to create a tailored message that is then delivered on-screen or on printed pages. Finally, secondary data can be accessed from technical report tables and online databases; these resources are described at the end of the preproduction section.

This section outlines basic approaches to selecting audience segments, specifying priority behaviors, elaborating cognitive and affective variables, ascertaining channel usage patterns, and evaluating key message elements. For a detailed illustration of how these methods and concepts are applied to the typical health campaign topic of drunk driving prevention, see Atkin and Freimuth (2001).

Identifying Audience Segments

Rather than attempting to reach a broad cross-section of the population, effective campaigns focus on multiple, specialized audience segments. Formative research is useful in identifying high-priority subgroups by gathering data about which categories of individuals are most relevant to the campaign goals, which are most receptive to media persuasion on the topic (through which media), and which are in a position to influence interpersonally the intended audience. Survey measures with representative samples are typically used to segment the audience along multiple dimensions defined in terms of demographic and psychographic characteristics, social role position, topic-related predispositions and behavioral intentions, readiness for change, and media exposure patterns. Segmenting the population into relatively homogeneous subgroups enables strategists to prepare distinctive messages adapted to the specific characteristics and predispositions of each subgroup and to select appropriate media channels to reach the intended audiences.

Most strategies focus on *primary* audiences composed of people to be directly influenced by the campaign. This approach may be supplemented by messages aimed at *secondary* (or gateway) audiences, such as friends, family members, and authority figures who are in positions to exercise influence over the primary audience (and who may have greater receptivity to campaign messages).

Specifying Focal Behaviors

Formative evaluation seeks to provide strategists with a better understanding of the nature of the existing problematic behaviors to be addressed by the campaign and the "product line" of optimal behaviors to be promoted. Basic behavioral practices are typically composed of specific component actions. Preproduction research helps determine audience receptivity to various discrete actions that could be recommended in a campaign. For example, an impaired driving campaign might promote the designated driver arrangement, companion intervention to take away keys from a drunk driver, one drink per hour for the driver, or abstention by the driver. These data enable designers to identify the promising options that are most amenable to change, to isolate certain unpopular but essential behaviors that will require special persuasive emphasis, and to eliminate from the campaign certain peripheral behaviors that will be widely resisted by the audience. Research may also assess reactions to various forms of symbolic branding of products and

of the sponsoring organization. For example, altruistic images may be associated with the designated driver, and tough punishment is the hallmark of the Mothers Against Drunk Driving group.

Elaborating Cognitive and Affective Variables

As a means to attaining the bottom-line behavioral objectives, campaign messages must first have an impact on preliminary or intermediate variables along the response chain, ranging from exposure and processing to learning and yielding to actual utilization. In particular, campaign designers face certain obstacles that must be overcome; these audience resistance points often involve misconceptions, dysfunctional attitudes, and behavioral inhibitions. Isolating the most crucial response stages is facilitated by an understanding of the characteristics and predispositions of the intended audience.

Knowledge and Literacy

Research illuminates the intended audience's entry-level awareness and information holding about the subject matter of the campaign, identifying what is already known, what gaps exist, what confusions must be clarified, and what misinformation must be corrected. It's also helpful to ascertain knowledge about topic-related vocabulary and terminology; research may show that people hold diverse meanings for same concepts and labels or that they lack comprehension of certain claims or evidence. Measuring degree of familiarity with specific behavioral recommendations or awareness of drawbacks of certain behaviors is useful in determining whether to employ one- versus two-sided message strategies or explicit versus implicit conclusions.

Beliefs and Perceptions

Because many campaign message strategies seek to alter subjective conceptions relating to behaviors and expected consequences, it is important to measure precisely the pre-existing cognitive orientations held by individuals. Measures are taken of the audience's beliefs and perceptions pertaining to 1) barriers and opportunities affecting performance of a behavior, 2) the likelihood expectations of experiencing beneficial and harmful outcomes, notably illusions of invulnerability, 3) social support or opposition from the interpersonal network and subgroup norms, and 4) monetary costs and affordability of societal resources.

Attitudes and Values

Affective predispositions are also a significant consideration in message design, particularly evaluations of outcomes associated with practices. Depending on the direction, intensity, and structure of relevant values and attitudes, the campaign may concentrate on creation, conversion, reinforcement, or activation. Attitudinal predispositions may pertain to potential behavioral products and to approval of policy options advanced in advocacy campaigns. Understanding the desirability of promised or threatened outcomes can help formulate strategy, such as emphasizing gain- versus loss-frames and

intensifying the degree of negativity or positivity related to the value component of the expectancy–value equation.

Salience Priorities

Research also provides guidance concerning which cognitive and affective orientations should be made more or less important to the audience. Salience measures assess level of involvement in the campaign topic, agenda ranking of a policy issue, and relative weighting of various outcomes that combine to shape behavioral intentions. More basically, research assesses the degree of concern or interest in the topical domain as well as the top-of-the-mind salience of performing the recommended behavior.

Efficacy and Skills

For certain practices, many well-intentioned and highly motivated individuals do not carry out appropriate acts because they lack confidence in their ability to perform the behaviors competently (see Rimal and Limaye, Chapter 17). If research shows that this is a barrier, messages can be designed to enhance personal efficacy or provide training for specific implementation skills.

Ascertaining Channel Usage Patterns

In deciding which channels are most efficient and effective for disseminating campaign messages, strategists need to determine intended audience preferences for traditional and new media and their patterns of interpersonal communication. While basic exposure figures are available from commercial audience measurement services (see "Preproduction Database Resources" in this chapter), customized surveys provide a much more elaborate and relevant array of data.

At a general level, it is useful to know the following information about the intended receivers: 1) amount of time spent watching television, listening to radio, reading magazines and newspapers, surfing the Internet, and visiting social media websites, along with exposure to secondary channels such as direct mail, billboards, or kiosks, 2) usage of specific media vehicles such as TV networks, magazine titles, and blogs and attention to various types of media content such as news stories and PSAs, and 3) patterns of community connections and interpersonal communication with pertinent categories of people.

Topic-specific data are especially pertinent to campaign planning: 1) consumption of news stories, product ads, and entertainment portrayals that directly complement or compete with campaign messages, and 2) interpersonal contacts, such as discussions with specialized opinion leaders, exposure to personal influence and peer pressures, and attempts to influence others on the topic.

Beyond exposure, formative researchers can assess credibility and utility of various media channels, vehicles, and content categories and may also measure audience recall and evaluative reactions to pertinent messages disseminated in previous campaigns. Similarly, audience members can be asked to identify interpersonal influence sources and memorable messages relating to the subject matter of the campaign.

Preliminary Evaluation of Message Elements

Before campaign messages are crafted and pretested, strategic and creative development is facilitated by both informal feedback and formal ratings of prospective components of these messages. Respondents typically rate a checklist of promising options displayed in the final section of an online survey instrument or respond to a set of open-ended queries. A comprehensive preproduction questionnaire might feature listings of three to five examples of various components that are under consideration for use in a campaign: spokespersons, headlines, slogans, persuasive incentives, storyline scenarios, stylistic devices, literacy levels, and supporting evidence. Responses such as credibility, comprehensibility, and subjective effectiveness are measured for each option.

Preproduction Database Resources

In addition to gathering customized information via surveys and focus groups, researchers also access a wide variety of previously collected data from governmental and private sectors. The federal health agencies have numerous databases that contain findings from sample surveys and statistical compilations. The CDC houses the National Center for Health Statistics (www.cdc.gov/hchs/index.htm) and CDC Wonder (www.wonder.cdc.gov), hosting databases on dozens of topics that generally feature demographic subgroups (adolescents, women), health behaviors (exercise, smoking), and diseases and conditions (diabetes, obesity). The Partners in Information Access for the Public Health Workforce (www.phpartners.org/health_stats.html) provides links to a massive array of databases from government agencies and public health organizations.

Companies specializing in public opinion surveys, audience measurement, and evaluation provide free or inexpensive data opportunities that are pertinent to campaign design. Harris Interactive offers low-cost omnibus survey options for general public polls and for specialized samples, such as youth, affluent people, and beltway influentials; it also offers access to findings from hundreds of surveys conducted over the past decade. For Gallup and other major U.S. pollsters, the Roper Center Public Opinion Archives (www.ropercenter.uconn.edu) serves as a repository for the most comprehensive collection of survey datasets emphasizing political opinions, along with data dealing with health, environment, social behavior, crime prevention, cultural attitudes, and media usage.

The Nielsen company (www.nielsen.com) regularly disseminates data on the public's exposure to major categories of media—television, radio, online, and mobile—and posts several white paper reports each month describing trends in use of traditional and new media technologies by the overall public and key audience segments, such as youth and minorities. Certain findings are frequently reported in media industry periodicals and websites, notably *Broadcasting & Cable* (www.broadcastingcable.com). The Arbitron company (www.arbitron.com) conducts the most extensive research on radio audiences, using traditional diaries and portable *people meter* devices. Much of Arbitron's listenership data are available to nonsubscribers; findings include market-wide ratings and useful subgroup comparisons by age, sex, income, education, and race.

One of the leading evaluation research firms is Weststat (www.weststat.com), which frequently conducts social science surveys on topics such as health, education, environmental protection, science, and technology. Their formative projects utilize both quantitative and qualitative methods to guide the development of large-scale communication campaigns, including the CDC's VERB physical exercise campaign and NIDA's antidrug media campaign; methodological techniques and evaluation findings are reported in white papers and journal articles.

Thus, there are a variety of ways to compile background information in the preproduction phase of campaign design. Informed by these data, researchers and strategists are in a position to collaborate with creative specialists in formulating and drafting an array of message themes and styles (and specific elements such as headlines, copy points, layouts, artwork, and music) and selecting source talent to appear in draft messages. As this stimulus construction process progresses, further research inputs are provided in the form of message pretesting.

PRETESTING RESEARCH

The second basic phase of formative evaluation is pretesting, the process of systematically gathering intended audience reactions to preliminary versions of messages before final production (U.S. Department of Health and Human Services, 2008). Pretesting can help determine which of numerous ideas or rough messages are most effective, and it can identify strengths and weaknesses in specific test messages. Because formative pretesting relies on measures of perceived message effects, Dillard, Weber, and Renata (2007) performed a meta-analysis to empirically validate how well pretest appraisals predict actual effects when campaigns are disseminated. They found a correlation of $+.41$ between perceived and actual message effectiveness, indicating that pretest measures tend to be moderately accurate. Pretesting research is involved in concept development and message creation through gauging sensitive elements.

Developing the Concept

Concepts are partially formulated message ideas consisting of visual sketches and key phrases that convey the main elements to be represented in the finished product. Pretesting at this stage provides direction for eliminating weaker approaches and identifying the most promising concepts, saving considerable time and money during production. Sometimes, entirely new concepts emerge from audience responses; original ideas are revised and refined, as in the case of concept testing feedback that heightened the realism of diverse student groups working together to create a drug prevention video (Freimuth, Plotnick, Ryan, & Schiller, 1997). Another advantage to pretesting rough concepts is the generation of words, phrases, and vernacular used by the intended audience so that appropriate language can be revised. For example, the National Cancer Institute (NCI) conducted concept testing to learn how to present cancer risk information. They learned that the word *risk* raises alarm while *chance* minimizes it. Vague or unfamiliar terms such as *fourfold* and

lifetime risk, gave people reason to discount the information (U.S. Department of Health and Human Services, 2008).

Creating the Test Message

Complete messages can then be created in rough form for the next stage of pretesting: message creation. For print materials, it is best to test a complete prototype of the final material, including the text, layout, typeface, and visuals. Rough executions of video spots used to rely on animatics (a series of detailed drawings filmed in rapid succession and using camera zooms and pans to give the illusion of movement) or photomatics (similar but using a series of photos), but because of digital photography, web access to stock photographs, and nonlinear editing programs, they have been replaced by much more rapid, computer-based techniques, which are referred to as digimatics. Radio messages and print materials can also be prepared for testing in rough form. If music or sound effects will be used in the final audio product, they should also be included in the rough message. Interactive media, such as websites, CD-ROMs, DVDs, or mobile device applications, should be complete enough to allow basic functionality and design to be assessed. At this stage, pretesting can be used to predict how effectively a message will move the intended audience through key types of reaction to campaign stimuli by 1) assessing the attention value of a message, 2) measuring its comprehensibility, 3) determining its relevance to the intended audience, 4) identifying strengths and weaknesses, and 5) gauging any sensitive or controversial elements.

Assessing Attention

An essential ingredient of messages is their ability to attract the intended audience's attention in the context of competing media and messages. Typically, this criterion is assessed by exposing intended audiences to a clutter format of several similar messages placed within an appropriate medium or context and then asking them to list the messages they remember seeing. For posters, for example, the intended audience might be asked to sit in a waiting room for a few minutes where several posters are hung on the walls. After leaving the room, they would be asked to recall the messages contained in the posters. For video messages, five to seven spot messages might be placed within an entertainment program and, afterwards, the audience asked to list all the ads or public service messages they remember seeing. More direct observational methods also have been used to assess attention, especially with very young children. Children's Television Workshop has used the distracter method, which measures attention by observing whether children are focused on the program or on a competing stimulus. Observational data also are gathered about the children's verbal comments, their singing and dancing in response to the show, and off-task activities, such as talking or playing with their friends (Palmer, 1981).

Measuring Comprehension

Messages must be understood before they can be processed and accepted. Procedures for measuring comprehension range from highly structured, closed-ended questions to open-ended requests for recall of main ideas. When developing the United States Department

of Agriculture (USDA) food pyramid, extensive pretesting revealed that the pyramid shape conveyed key concepts more clearly than a bowl or other shapes. In addition, planners learned that representing fats, oils, and sugars as a bottle of salad dressing, a can of soda, and a bowl of sugar created widespread misunderstanding (U.S. Department of Health and Human Services, 2008).

Identifying Strong and Weak Points

Pretesting prior to final production and distribution can help ensure that each element of a message is likely to meet the information needs of the audience. In a test of a booklet on lung cancer, patients could recall on average 2 out of 12 ideas presented; half the patients were not able to recall any, however. Too many technical terms, high density of concepts, concepts unimportant to the audience, and too little differentiation between diagnostic and treatment procedures inhibited intended audience recall (U.S. Department of Health and Human Services, 2008).

Determining Personal Relevance

Intended audiences must perceive that a message personally applies to them for the message to be effective. When the National Bone Health Campaign research team explored the credibility of a spokesperson among girls 9 to 12 years old, the sample said they wanted to hear the message from a girl who was strong, bold, confident, active, healthy, and popular. This advice guided the development of Carla, a cartoon spokes-character who speaks to girls as a peer working to build powerful bones. In the pretesting of a website featuring Carla, girls described her as powerful and fun—someone they'd like to know and be friends with (U.S. Department of Health and Human Services, 2008).

Gauging Sensitive or Controversial Elements

Pretesting can help in determining whether messages may alienate or offend intended audiences, often rejecting sponsors' or interest groups' assumptions about the general public's responses. Pretest results showed that using vernacular language to discuss diarrhea in a booklet on chemotherapy was regarded as offensive by the low literacy-intended audience, who preferred a more technical discussion (U.S. Department of Health and Human Services, 2008).

TYPES OF PRETESTING

The following sections summarize a variety of pretesting techniques in developing public service messages: 1) focus group interviews, 2) individual in-depth interviews, 3) central-location intercept interviews, 4) self-administered questionnaires, 5) theater testing, and 6) other techniques, such as readability testing, usability testing, gatekeeper review, participatory rapid appraisals, and multimethod combinations.

Focus Group Interviews

Focus group interviews are a form of qualitative research adapted by marketing researchers from group therapy (Krueger & Casey, 2009). They are conducted with a group of approximately 5 to 10 respondents simultaneously. Using a discussion outline, a moderator builds rapport and trust and keeps the session on track while allowing respondents to talk freely and spontaneously. As new topics related to the outline emerge, the moderator probes further to gain useful insights. An experienced, capable moderator should lead the groups. The moderator must be well informed about the subject and the purpose of the group sessions.

Focus groups are a very flexible formative research method. Many focus group sessions are conducted in sophisticated facilities with one-way mirrors and video cameras, but they are also carried out in workplaces, in outdoor settings, via telephone, and over the Internet. Online focus groups can be held in modified chat rooms. Participants are recruited by phone or e-mail, and they log onto a designated website at a prearranged time. Once in the "room," the participant's screen is usually divided into two sides: One side has the text of the discussion, and the other shows messages and materials. Although the online setting reduces the in-depth emotional information obtained by personally observing the participants, this method lowers the costs (by eliminating expenses for travel, food, and renting facilities), and discussion threads create an instant transcript (Heckman, 2000). Turney and Pocknee (2005) argue that these virtual focus groups are particularly useful when participants are difficult to recruit or access in central locations and that this method enables safe, secure, and anonymous environments in which to share ideas and experiences.

Focus group interviews provide insights into intended audience beliefs on an issue, allow program planners to obtain perceptions of message concepts, and help trigger the creative thinking of communication professionals. The group discussion stimulates respondents to talk freely, providing valuable clues for developing materials in the consumers' language.

As with any qualitative research approach, care must be taken not to interpret or generalize focus group interview results quantitatively; the testing is indicative and not definitive. Focus group respondents should be typical of the intended audience. Subgroups within the intended audience representing relevant positions on the issues should be included, usually in separate focus groups. For example, in testing message concepts on smoking aimed at a general audience of smokers, it would be helpful to conduct focus groups with some key subgroups, such as males and females, heavy and light smokers, and older and younger ages to determine if the messages were effective across all these groups.

Individual In-Depth Interviews

Individual, in-depth interviews are used for pretesting issues that are very sensitive or must be probed very deeply and for respondents who are difficult to recruit for focus group interviews, such as physicians, dentists, and chief executive officers. Such interviews can be quite long, lasting from 30 minutes to 1 hour, and are used to assess comprehension as well as feelings, emotions, attitudes, and prejudices. Although in-depth interviews are

very costly and time-consuming, they may be the most appropriate form of pretesting for sensitive subjects (e.g., breast reconstruction).

Central-Location Intercept Interviews

Central-location intercept interviews involve stationing interviewers at a location frequented by individuals from desired intended audiences and, after asking a few screening questions, inviting qualified respondents to participate in the pretest. If they are willing to participate, each respondent is taken to the interviewing station, shown the pretest messages, and asked a series of questions to assess their reactions to the message concepts or executions. One advantage to this type of pretesting is that a high-traffic area can yield many interviews in a reasonably short time. The second advantage is that using a central location for hard-to-reach intended audiences can be a cost-effective means of gathering data. As with focus groups, sampling is not random, and the results cannot be generalized to a larger population. A significant disadvantage of this method is the obtrusiveness of the study; because respondents know they are participating in a test, their responses may be less valid.

The National Youth Antidrug Media Campaign used the central intercept method in a unique way. This campaign developed a rigorous, quantitative copy testing system to assess near-final ads. The research used an experimental design with respondents recruited in a mall intercept and randomly assigned to two conditions. The test group saw the message, and the control group did not, and both groups filled out brief questionnaires to assess effects. The results had to show significantly greater antimarijuana attitudes and intentions relative to the control group in order for the ad to be aired (Foley & Pechmann, 2004).

Self-Administered Questionnaires

Self-administered questionnaires can also be used to pretest concepts and rough messages. These questionnaires can be mailed to respondents along with pretest materials or distributed at a central location. Each respondent is asked to review the materials, complete the questionnaire, and return it by a certain date. The Internet has enhanced the use of this method. It is now possible to inexpensively conduct these tests on a website, which speeds up the data collection process and encourages broader participation. Disadvantages include a low overall response rate, tendency toward self-selection of individuals who have either strongly positive or negative responses to the pretest materials, and sample biases relating to respondents being Internet users.

An efficient approach for assessing audience responses to messages is a fairly brief set of questionnaire items that can be administered following exposure to a specimen message. The instrument measures the perceived effectiveness of the message for producing an impact on an intended audience (e.g., "How effective is this message in influencing college students to avoid driving drunk?" with response categories such as *Very Effective, Fairly Effective, Slightly Effective,* and *Not Effective*). Then, respondents evaluate the message along, perhaps, a dozen quality dimensions using a numerical scale ("What is your personal reaction to the message? Give ratings using a scale from 0 to 10 on each of these

factors."). Typical factors (and accompanying definitions) include *informative* (tells you something new, increases your knowledge), *sensible* (presents wise advice that seems reasonable), *memorable* (vivid image, fascinating fact, and catchy slogan), *enjoyable* (interesting, entertaining, and stimulating message), *useful* (valuable information, helpful advice worth remembering), *imaginative* (style is refreshing, novel, unique, and clever), *believable* (accurate information, sincere and trustworthy characters), *professional* (high production quality), *motivating* (presents influential reasons to prompt changes in behavior), and *relevant* (personally meaningful content, identifiable with characters and situations). Next, respondents provide assessments of whether or not the message has any of these negative features using a simple Yes or No response to a series of questions on factors such as the following: *preachy* (tone of message too moralistic or righteous), *disturbing* (turned off because it is too emotional or threatening), *confusing* (vague or difficult to understand), *irritating* (offensive or annoying), *dull* (boring, stale, or trite style), or misleading (biased arguments or exaggerated claims). These standardized evaluations may be supplemented with open-ended questions soliciting positive or critical comments as well as suggestions for improving the message.

Theater Testing

Theater testing uses forced exposure to test rough television message executions in controlled settings. Testing takes place with several hundred randomly recruited respondents representative of the message's intended audience; they are seated in groups of approximately 25 around large TV monitors. The test spots are embedded among other commercials in TV program material to camouflage the intent of the testing situation and simulate the home viewing context, and all questions are prerecorded and administered over TV monitors.

Respondents are asked to recall, on an unaided basis, all the messages they remember by brand name, product type, or public service (the attention measure). They are then asked to write down the central point each message was trying to convey (the main idea communication measure). Subsequently, respondents are exposed to the embedded test messages for a second time, followed by diagnostic questions that probe respondent reactions, including personal relevance and a believability measure. Theater testing also provides an opportunity to use electronic devices to record and display moment-to-moment evaluations of messages, which can later be overlaid on the actual messages to identify particularly positive or negative components.

Other Techniques: Readability, Usability, and Gatekeeper Review

Readability testing is critical when producing print materials because it estimates the educational level required for intended populations to adequately comprehend written text. Readability tests are available on many standard word processing packages, or a score can easily be computed by hand (U.S. Department of Health and Human Services, 2008).

To ensure that users will find a website or mobile phone app well designed and easy to use, *usability testing* is performed during the development of the site or app. People representing the intended audience actually sit down and use the site to complete tasks,

while a research aide observes how they interact with the site and asks them questions following completion of the tasks.

Public communication campaigns often rely on *gatekeepers* to disseminate materials, such as PSA directors or station managers who select and schedule broadcast messages and health professionals who hand out or display print messages. In order to enhance cooperation, gatekeepers may be asked to review and evaluate rough materials concurrently with audience pretesting.

The traditional formative research methods described in the previous sections have been criticized because the audiences are seldom involved in problem identification that sensitively meets their information needs. Some campaigns employ participatory rapid appraisal as a semistructured process of learning from, with, and by underserved or remote population segments such as residents of rural or foreign locales (Clift & Freimuth, 1997) or the sense-making methodology (Dervin & Foreman-Wernet, Chapter 10).

Finally, multiple formative evaluation methods are frequently combined as the campaign is developed. The VERB campaign, the CDC effort to increase physical activity among *tweens* (9- to 13-year-olds), made extensive use of formative evaluation methods to develop the brand, test messages, and materials and decide on which media outlets to place campaign advertising (Berkowitz et al., 2008). The brand development process used interviews with triads of tweens, focus groups with parents of tweens, and adults who work with this age group and in-depth interviews with industry professionals. In addition, this campaign conducted both semiotic analyses of ads and brands targeting tweens and hedonic analyses of these brands to reveal the elements that inspire tween affinity from a pleasure-seeking perspective. Another innovative formative technique used in developing the VERB brand was diaries kept by tweens and their parents depicting tweens' after-school experiences, expressing ideas, feelings, hopes, and fears through words and pictures.

Conclusion

By collecting preproduction information and feedback reactions to pretest theoretically derived versions of the message concepts and executions, campaign designers are in a much better position to devise more effective campaign plans and messages before final production and full-scale dissemination. Formative evaluation facilitates the development of more sophisticated campaign strategies, helps avoid pitfalls, and improves the quality and effectiveness of the created messages.

References

Atkin, C., & Freimuth, V. (2001). Formative evaluation research in campaign design. In R. E. Rice & C. Atkin (Eds.), *Public communication campaigns* (pp. 125–145). Thousand Oaks, CA: Sage.

Bauman, A., Smith, B. J., Maibach, E. W., & Reger-Nash, B. (2006). Evaluation of mass media campaigns for physical activity. *Evaluation and Program Planning, 29,* 312–322.

Berkowitz, J., Huhman, M., Heitzler, C., Potter, L., Nolin, M., & Banspach, S. (2008). Overview of formative, process, and outcome evaluation methods used in the VERB™ campaign. *American Journal of Preventive Medicine, 34*(6), S222–S229.

Cho, H., & Witte, K. (2005). Managing fear in public health campaigns: A theory-based formative evaluation process. *Health Promotion Practice, 6*(4), 482–490.

Clift, E., & Freimuth, V. S. (1997). Changing women's lives: A communication perspective on participatory qualitative research techniques for gender equality. *Journal of Gender Studies, 6*(3), 289–296.

Dillard, J. P., Weber, K. M., & Renata, G. V. (2007). The relationship between the perceived and actual effectiveness of persuasive messages: A meta-analysis with implications for formative campaign research. *Journal of Communication, 57,* 613–631.

Foley, D., & Pechmann, C. (2004). National youth antidrug campaign copy testing system. *Social Marketing Quarterly, 10*(2), 34–42.

Freimuth, V. S., Plotnick, C. A., Ryan, C. E., & Schiller, S. (1997). Right turns only: An evaluation of a video-based, multicultural drug education series for seventh graders. *Health Education & Behavior, 24*(5), 555–567.

Heckman, J. (2000). Turning the focus online. *Marketing News, 34*(5), 15.

Knowles, E. S., & Linn, J. A. (2004). *Resistance and persuasion.* Mahwah, NJ: Lawrence Erlbaum.

Krueger, R. A., & Casey, M. A. (2009). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks, CA: Sage.

McGuire, W. (2001). Input and output variables currently promising for constructing persuasive communications. In R. E. Rice & C. Atkin (Eds.), *Public communication campaigns* (pp. 22–48). Thousand Oaks, CA: Sage.

Noar, S. M. (2006). A 10-year retrospective of research in health mass media campaigns: Where do we go from here? *Journal of Health Communication, 11*(1), 21–42.

Palmer, E. (1981). Shaping persuasive messages with formative research. In R. E. Rice & W. Paisley (Eds.), *Public communication campaigns* (pp. 227–242). Beverly Hills, CA: Sage.

Parrott, R., Steiner, C., & Goldenhar, L. (2008). Georgia's harvesting healthy habits: A formative evaluation. *The Journal of Rural Health, 12,* 291–300.

Rimer, B., & Kreuter, M. (2006). Advancing tailored health communication: A persuasion and message effects perspective. *Journal of Communication, 56,* S184–S201.

Rogers, E. M. (1996). Up-to-date report. *Journal of Health Communication, 1,* 15–24.

Smith, S. W., Atkin, C., Martell, D., Allen, R., & Hembroff, L. (2006). A social judgment theory approach to conducting formative research in a social norms campaign. *Communication Theory, 16,* 141–152.

Turney, L., & Pocknee, C. (2005). Virtual focus groups: New frontiers in research. *International Journal of Qualitative Methods, 4*(2), 1–10.

Uhrig, J. D., Bann, C. M., Wasserman, J. W., Guenther-Grey, C., & Eroglu, D. (2010). Audience reactions and receptivity to HIV prevention message concepts for people living with HIV. *AIDS Education and Prevention, 22,* 110–112.

U.S. Department of Health and Human Services. (2008). *Making health communication programs work: A planner's guide.* Bethesda, MD: National Cancer Institute. Retrieved September 29, 2010, from the national Cancer Institute: http://www.cancer.gov/pinkbook