

Evaluating Communication Campaigns

2007 Research and Evaluation Conference
September 27–28, 2007



Robert Wood Johnson Foundation

Table of Contents

3	Background
3	Evaluation of Public Health Communication Programs Robert Hornik, Ph.D.
3	Examples of Communication Campaign Evaluations
4	Special Problems of Communication Evaluation
5	Alternative Evaluation Designs for Communication Programs
6	Major Principles of Communication Evaluation
7	Discussion
7	Evaluating Policy Communications: Triangulation and Collaboration Lawrence R. Jacobs, Ph.D.
8	An Overview of Policy-Oriented Communications
9	Examples of Policy Communication Evaluations
9	<i>Americans Discuss Social Security</i>
10	<i>Environmental Public Health Education Campaign</i>
11	Lessons Learned
12	Discussion
13	Evaluating Advocacy Julia Coffman
13	Differences Between Advocacy Evaluation and Other Evaluations
14	Steps in Designing an Advocacy Evaluation
14	<i>Utilization Decisions</i>
15	<i>Strategic Focus Decisions</i>
16	<i>Methodology Decisions</i>
17	Summary
18	Discussion
20	Speaker List

Reported by **Lori DeMillo**

Background

The material in this paper is derived from a part of the Robert Wood Johnson Foundation's Research and Evaluation Conference held at the Foundation's campus in Princeton, N.J., on September 27–28, 2007. As RWJF has expanded the use of communication campaigns to create change, effective evaluation of these campaigns has become crucial. Recent campaigns include Cover the Uninsured Week and the Campaign for Tobacco-Free Kids and the Back to School Campaign (to enroll children in SCHIP and Medicaid). A communication campaign also will be part of the childhood obesity strategy.

The purpose of the conference was to convene evaluators, RWJF staff and staff of other foundations to learn about and discuss evaluating communication campaigns. Participants included evaluators who are grantees of RWJF, staff of RWJF's Research and Evaluation and Communications Departments, and staff of the Annie E. Casey Foundation, Lumina Foundation and the Pew Charitable Trusts.

Evaluation of Public Health Communication Programs

Robert Hornik, Ph.D., Wilbur Schramm Professor of Communication at the University of Pennsylvania's Annenberg School of Communication, highlighted the special problems in evaluating communication programs. He presented evaluation designs that are well suited to specific types of communication programs and outlined major principles of communication evaluation. See his **Power Point**[®] Presentation for more details.

Laura C. Leviton, Ph.D., senior program officer at RWJF, served as commentator. Leviton noted the importance of understanding the mechanisms by which communication programs achieve change.

Although Hornik has conducted many summative evaluations (which are designed to present conclusions about the merit or worth of a program at the end of the program activities) for communication programs, he suggested that there are better ways to use research resources. Determining causality for communication programs is expensive and difficult.

Most communication programs need an evaluation that will show how to improve the program or adapt it for a different audience. Formative research to test message strategies or choose communication channels and monitoring research to test message recall and movement of cognition are often a better use of research money than a summative evaluation.

Examples of Communication Campaign Evaluations

Hornik contrasted two evaluations to illustrate that a scientifically well-designed evaluation may not teach us as much about communication programs as an evaluation where claims of effect are less certain: the Commit Anti-Tobacco Campaign and the National High Blood Pressure Education Program.

The Commit Anti-Tobacco Campaign used local communication sources to target heavy smokers in 11 cities. The randomized control design evaluation was comprised of the 11 experimental cities (with the campaign) and 11 control cities (without the campaign). In order to ensure that communication remained local, program managers could not use national media.

The evaluation showed almost no difference between the percentage of heavy smokers who quit in the experimental cities (18 percent) and those who quit in the control cities (18.7 percent).

Hornik noted that although the evaluation design was excellent from a scientific view, it constrained what could be learned from the campaign, because the need to keep the control cities unexposed constrained the way in which communication was used in the campaign—in particular, by leaving it unable to use national media and take advantage of the sort of buzz that may derive from national coverage.

Unlike the limited Commit Anti-Tobacco Campaign, the National High Blood Pressure Education Program used a “kitchen sink” approach, using mass media, professional and community organizations, and direct education to bring attention to the issue of blood pressure control. There was no control group; the evaluation had to depend on inferences from changes occurring over time.

The National High Blood Pressure Education Program coincided with a very large change in the number of people who had their high blood pressure under control and a sharp decline in stroke mortality. The evaluation design did not permit an unequivocal inference that the campaign caused these changes, although examination of other likely historical forces did not suggest an alternative explanation. Hornik noted that this evaluation was clearly responsive to the “messy kitchen sink way” that communication campaigns can have effects.

Special Problems of Communication Evaluation

Communication programs can have effects because they directly persuade individuals and/or because their messages are diffused through social networks and/or because they influence policy decisions in institutions. Knowing whether the model of effect is individual, social or institutional is the most important problem in communication evaluation, according to Hornik. Each model of effect requires measurement of different outcomes, and different strategies for showing that influence has occurred.

Other special problems in communication evaluation include:

- ***Multiple components of communication programs that interact in a complex way and evolve over time.*** Components include multiple channels, messages and audiences with varying intensity. Communication programs are dynamic. Evaluation designs which are fixed when the program is first designed may not capture effects of complex and evolving programs.
- ***Discrete versus continuing exposure.*** A one-time communication program to produce quick behavior change may be well suited to a conventional

evaluation approach. However, a continuing effort to produce a long-term effect requires measurement over time and the ability to link inputs and outcomes that vary with time.

- ***Time lag for effects.*** The expected timing (e.g., whether effects are expected immediately or years later) will affect the evaluation approach.
- ***Magnitude of effects.*** Communication programs are generally designed to create small changes among many people. It is, however, difficult to detect these effects (e.g., a 2 percent decrease in cigarette use each year over many years). Small-scale interventions (e.g., in one school) require big effects to justify their cost, but mass media programs can be justified even if they produce small effects on many people. Evaluation designs must then be able to detect small effects.
- ***Matching study and target populations.*** A mass media campaign may reach large audiences but may expect to produce change only among the subgroup at risk (e.g., for a condom campaign, largely among single people with casual sex partners). Evaluators must focus the analysis on the number of people who changed among those who were at risk, rather than the total audience.
- ***Matching unit of analysis and unit of treatment.*** If the model of effect is the individual, it is sensible to compare individuals. With a social or institutional model of effect, however, comparing individuals may underestimate the effects.

Alternative Evaluation Designs for Communication Programs

Due to the special problems of communication evaluation, standard randomized controlled trials or individual-difference designs may be inappropriate. Hornik presented examples of alternative evaluation designs, but noted that their usefulness largely depends upon the model of effect and the possible threats to inference in a concrete context.

Alternative designs for communication evaluation include:

- ***Pre-post design.*** It is difficult to make claims of effects when evaluating change over time. However, if there are no reasonable alternative explanations for change other than the communication campaign, the pre-post design is suitable.

For example, the National Urban Immunization Campaign in the Philippines used mass media advertising to educate people about the need for measles vaccination. The pre-post evaluation showed substantial change in on-time vaccination from the previous year (from 32 percent to 56 percent), and no evidence that anything else was changing generally, or in the clinics providing vaccination.

- ***Cohorts with lagged exposure effects.*** This evaluation approach may be used to show effects over an extended time period. For example, the evaluation of the National Youth Anti-Drug Media Campaign found that youth who reported more exposure to the campaign showed lower anti-marijuana norms 12 to 18 months later, after controlling for factors that could distort or mask the true effect of the intervention.

- ***Interrupted time series.*** The Kentucky Anti-Drug Campaign was done twice in Fayette County. After each four-month campaign, spaced nine months apart, evaluators saw a decline in the use of marijuana, among high sensation seekers. In the comparison county, where the campaign was not used, evaluators did not see this effect.
- ***Continuous time series.*** To determine whether news coverage of cancer would influence people to look for more information, an evaluator counted the number of articles that appeared on the AP wire each week, and then compared this to the proportion of people who said (in a weekly survey) that they had searched for cancer information. Among people who said they paid a lot of attention to health information in the media, there was a direct relation between the number of articles that appeared in a given week and whether people looked for cancer information.

This approach might work in evaluating a campaign focused on getting media attention on a policy issue.

- ***Natural experiments.*** The evaluation of the California Tobacco Campaign, which included anti-tobacco advertising, was a natural experiment. Evaluators compared California to the rest of the country. They found that smoking prevalence decreased at a greater rate in California than in the rest of the nation.
- ***Field experiments.*** Formal experiments are sometimes appropriate in communication evaluations. The University of Vermont-evaluated youth tobacco campaign was a four-year intervention in four communities. Some communities received only school-based programs and some also had an ongoing media campaign. The evaluators showed that initiation of smoking was substantially reduced in the communities with the media campaign.

Major Principles of Communication Evaluation

Hornik outlined the following principles of communication evaluation:

- ***Respect the model of effect.*** Designs are specific to a context and an intervention. Formally weak designs may work in some contexts, while gold standard designs may be misleading.
- ***Measure exposure carefully.*** Most communication projects fail due to lack of sufficient exposure to the message.
- ***Do what is possible and live with uncertainty.*** While it is always better to include more time points, comparison groups and mediating variables, as well as large samples, evaluators must make choices in designing an evaluation based upon what is possible given available resources. Evaluators must consider not whether the evaluation eliminates all uncertainty but rather whether the evaluation is good enough for policy work, compared to the alternatives, including the alternative of not evaluating at all.

In summary, Hornik noted that if summative evaluation is necessary, evaluators should consider a design that is appropriate for the program and the context. Evaluation designs must respect the way in which communication is likely to have effects.

Discussion

Laura Leviton of RWJF led the discussion by noting the challenge of using communication programs to create change in RWJF's childhood obesity work, which focuses on policy and environmental change.

Elaine B. Arkin, a special communications officer at RWJF, noted that communication programs can be difficult and expensive, but they can work. She shared the example of an advertising campaign RWJF funded to raise awareness of SCHIP after its initial passage, when research showed that many parents did not know about the program. The success of this campaign led RWJF to develop the *Covering Kids & Families*[®] program.

A member of the audience asked about grounds for attributing an effect to a communication campaign. Hornik replied that although there are many methods to show that components of a campaign made a difference, communication is a “messy kitchen sink” intervention. Expectations about accountability claims should be realistic.

In response to a question about how to judge the timing of the effect and how long to measure the effect, Hornik suggested monitoring public opinion through a continuous survey. He also noted the need for judgment, that is, “I can't say for sure but it looks good” evaluation. Leviton added that ruling out other influences is a constant challenge in evaluating communication campaigns.

An audience member asked how evaluators adjust to the rapid changes in communication channels (e.g., viral marketing, a Web-based strategy that encourages people to pass on a marketing message to others, creating the potential for exponential growth in the message's exposure and influence).

Hornik responded that thus far campaigns have generally not been able to assure that the largest part of the target audience is receiving the same message through new individually controlled communication technologies. While there is lots of use of such technologies, the capacity to define the messages that individuals receive is quite limited.

On the other hand, there remains a great deal of opportunity to reach large audiences in a controlled way through traditional channels; in that sense technology does not yet make so much of a difference. For example, one can be confident that an anti-drug advertisement on a television program will reach most teenagers in a predictable way; it is much more difficult to assure that a specific message or Web site reaches the mass of teenagers on the Internet.

Evaluating Policy Communications: Triangulation and Collaboration

Lawrence R. Jacobs, Ph.D., Walter F. and Joan Mondale chair for political studies at the University of Minnesota, provided an overview of policy communication evaluations, including effective strategies and lessons learned. He presented several policy communication evaluations that he conducted for the Pew Charitable Trusts to illustrate this type of evaluation. His **Power Point**[®] presentation provides more detail.

Mark A. Peterson, Ph.D., professor of public policy and political science at the University of California, Los Angeles, served as the commentator.

An Overview of Policy-Oriented Communications

Jacobs began by describing the two strategies for policy-oriented communications: informing the citizenry and changing government policy—although these two strategies are not mutually exclusive; here are many examples of communication campaigns that use both approaches.

In a campaign to inform the citizenry, the evaluation focuses on changes in public opinion. When changing policy is the strategy, the evaluation focuses on government policy-makers. Whichever strategy evaluators use, evaluation of policy-oriented communication is often a one-time event.

Propaganda, one-sided indoctrination, rarely works in policy-oriented communication because the public is well educated and has access to diverse information. A more effective approach is to use credible information as a catalyst for change. With this tactic, the keys to success are:

- Credibility—the information and its sources must be trustworthy;
- Communications that are widespread and recurrent;
- Collaboration, which multiplies the effort.

Jacobs noted that the policy communication environment is complex, with many competing forces: the media, real-world events, government statements or actions, and an overloaded and already opinionated public. Analyzing causal connections is difficult and will require a combination of approaches as well as targeted analyses of distinct stages of the policy process from agenda setting to policy adoption, implementation and evaluation itself.

Special problems in policy communication are:

- *Spuriousness*. A seemingly efficacious strategy may reflect the impact of alternative influences.
- *False-negatives*. The evaluation can miss the effects or signs of impending effects. Finding the impact short of the ultimate objective is challenging. Even when a grantee's project coincides with the desired outcome, it can be difficult to pinpoint the role of any single actor or communication effort.

Jacobs presented the following strategies for policy communication evaluations to enable evaluators to deal with the complex environment and special problems of these evaluations:

- *Analytic pluralism*. Combining quantitative analyses, document review and interviews (analytic triangulation).
- *Sanding to wood*. Using an intensive, disaggregated investigation.
- *Modesty*. Assuming uncertainty and considering many sources of evidence.
- *Collaboration*. Engaging in ongoing, iterative communication with the funding foundation and grantees.

Examples of Policy Communication Evaluations

Jacobs presented two public policy evaluations he conducted for the Pew Charitable Trusts:

- An “Opinion Change” Model: Americans Discuss Social Security
- A “Policy Change” Model: Environmental Public Health Education Campaign.

Americans Discuss Social Security

The objective of this campaign was to educate the public about challenges facing Social Security. Campaign tools for informing public opinion were:

- Non-partisan forums about Social Security, with a 10-city interactive video teleconference and forums in five cities (Austin, Texas; Buffalo, N.Y.; Seattle, Wash.; Des Moines, Iowa and Phoenix, Ariz.);
- Encouraging national and local media coverage of the teleconference and forums in order to widen the impact to the general public.

The opinion change model was based on the hypothesis that the forums would drive public opinion and media coverage, and that media coverage would also drive public opinion.

The evaluative task was to:

- Use multidimensional measurements of public opinion to detect the potential impacts of the forums in terms of salience (the level of attention, thinking, talking and reading about Social Security), knowledge and political participation (e.g., contacting a member of Congress).
- Measure the impact of the forums on changes in opinion (comparing opinions before and after the forum).
- Isolate the foundation’s impact from other factors (e.g., newspaper stories, mailings from the Social Security Administration and AARP activities).

Evaluators conducted comparison group analysis using pre- and post-testing to explore:

- Change in the opinions of the forum participants in Phoenix;
- Change in the opinions of non-participants in the Phoenix forum (e.g., people invited to the forum who did not attend and a random sample of adults living in the greater Phoenix area).

If participants’ views changed significantly more than those of non-participants, the evaluators would infer that the forum prompted the opinion change.

The findings illustrate the challenges of this type of research. Salience and knowledge increased more for participants than non-participants. There was, however, little meaningful change in political participation and the impact of the media on public opinion was difficult to evaluate.

These findings raised key questions about the campaign and the evaluation, according to Jacobs:

- Did the evaluation miss the effects?
- Was the grantee over-ambitious about what the campaign could achieve?
- Can the media be effective as a tool to be wielded to manufacture the change in opinion that the grantee seeks?

Environmental Public Health Education Campaign

The objective of this campaign, known as Health Track, was to contribute to establishing a nationwide network funded by the federal government to collect community-level data on chronic diseases and environmental exposure.

The campaign sought to achieve policy change to secure federal funding for the tracking network. It used many tools to mobilize public opinion, the public and policy-makers:

- “Grasstops” (community and organization leadership) communications
- Public opinion surveys
- Newspaper advertisements
- Coalescing the expert community
- Obtaining media coverage
- Stimulating attention and policy formulation in executive branch agencies and Congress.

Evaluators designed the mid-course evaluation to detect progress short of the goal of establishing a tracking network. They used a complicated, multi-stage strategy that involved digging deep and conducting intensive, detailed, parallel content analyses of media and elites.

Evaluators searched for changes short of policy change: discussions within executive agencies, Congressional hearings and floor activities. They also assessed the efficacy of individual steps in the campaign.

The content analyses tracked salience, framing of the issue (e.g., substantive versus adversarial) and policy directionality (receptiveness). Evaluators analyzed government deliberations in Washington about environmental health and health tracking (e.g. content analyses of congressional hearings and executive branch reports), interviewed national and local/state elites and analyzed media reporting on environmental health issues—both before and after the campaign.

The findings of the pre- and post-analysis showed quite a bit of debate in Congress and executive agencies:

- Two bills were introduced in Congress.
- Establishing a tracking network was the fourth most widely discussed issue among all environmental health issues.
- The tone of the discussion was strong and supportive.

The interviews confirmed salience and support, however; state and national elites did not think the media was very interested in the issue. This finding raised key questions:

- Did the evaluation miss the effects?
- How well did the media priming strategy work?

Media analysis showed:

- Increased salience: coverage of environmental health increased 60 percent from 1999 to 2000 and coverage of Health Track doubled;
- More positive, constructive coverage of environmental health and Health Track;
- The media framed the story in national terms that highlighted the need for an inclusive system of tracking;
- The media relied upon supportive sources.

Evaluators found little impact of the grassroots communication strategy. One half of the targeted regions failed to draw significant coverage of health tracking and only a few state advocates emerged as sources for local media. This finding raised key questions:

- Did the evaluation miss the effects?
- How effective were the grantees and the campaign strategy?

The public opinion surveys did not attract the attention of policy-makers and the media. Public support for health tracking was not cited or perceived in press coverage, statements of policy-makers or interviews. Staff and grantees argued that the evaluation missed the effects. Jacobs noted that the strategy may not have been cost effective.

Lessons Learned

Jacobs presented lessons learned from these two evaluations:

- *Use a multi-prong approach to evaluation (analytic triangulation).* Combine quantitative analyses with document review and interviews.
- *Strive for payoff.* The evaluation should contribute to the foundation's up/down recommendation as well as a hard-headed assessment of the costs and benefits of particular tools and future policy communication efforts.
- *Consider evaluation a collaborative project:*
 - Evaluators should reach out to communicate with foundation staff and grantees in order to understand objectives and "local knowledge" of the area and players, and to consider alternative approaches to conducting evaluation to track progress.
 - Strive to identify gaps in the research design by listening to the people being evaluated.
- *Ensure that the data are credible:*
 - Check on the subjectivity/biases of the evaluator, especially in relation to difficult analyses.
 - Quantitative data are a critical foundation for judgments by the evaluator and staff about degrees and relative progress or payoffs.

Discussion

Peterson noted that because of the complexities of the policy-making process, and the myriad factors that influence policy outcomes, macro thinking is important in designing and evaluating interventions. There are three basic ways in which policy campaigns affect policy-maker decision-making:

- Direct unmediated policy information received by policy-makers;
- Influence on policy-makers by creating *the appearance* of public support—“Astroturf” mobilization. Influence on policy-makers by mobilizing *actual* public support). They affect different parts of the policy-making process.

It is difficult to know whether a particular policy change has taken place in response to a policy campaign or due to other factors, he said. Even when policy-makers take action as a result of a campaign, it may be because they were influenced by normal communications (e.g., reading the newspaper or surfing the Internet) rather than by communication efforts targeted to influence them through mobilizing the public, their constituents.

In order to sort out these effects, the intervention and the evaluation must both be designed with a clear understanding of the policy-making process. Peterson highlighted John W. Kingdon’s model of the policy-making process, which has three streams:

- Problems (recognition, definition, and framing of problems);
- Policy (the policy alternatives);
- Politics (the relative influence of coalitions in support of or in opposition to various alternatives).

When there is an alignment of the three streams, there is a policy-making window of opportunity. Some policy campaigns (e.g., RWJF “Cover the Uninsured”) try to influence the problem stream—but will affect policy change only if appropriate policy options are available and along with favorable political support. Other campaigns strive to influence the politics stream—but any effect on policy outcomes will depend on the problem already being appropriately understood and viable policy options being readily available.

Lester W. Baxter, deputy director of evaluation at the Pew Charitable Trusts, noted that communication planning and evaluation planning are similar. Both require answering the following questions:

- Who is the audience?
- What do they need to know?
- How best will they take in the information?
- Who would be the most credible messenger?

Baxter also said that good evaluation and communication plans change with the circumstances.

An audience member asked for examples of indicators of progress toward policy change. Jacobs responded that indicators include bills introduced and actions in Congressional committees (e.g., hearings).

Another audience member commented that polling and survey, which can be difficult to interpret without more detailed qualitative data collection—semi-structured interviews with a subset of survey respondents—are not very helpful. Jacobs responded that both types of data are necessary.

Evaluating Advocacy

Julia Coffman, an evaluation and strategy consultant with the Harvard Family Research Project at the Harvard Graduate School of Education, discussed the differences between evaluating advocacy and evaluating programs or services and illustrated the steps in designing an advocacy evaluation. Coffman presented a new tool to facilitate advocacy evaluation design—an advocacy composite logic model developed by multiple funders, evaluators and advocates. Her **Power Point** presentation provides more details.

Elaine B. Arkin, a special communications officer at RWJF, served as the commentator.

Coffman began by noting that in some ways, evaluating advocacy efforts is similar to evaluating anything else. All evaluations involve systematic data-based inquiry and aim to provide high-quality information that has value for who or what is being evaluated. The same evaluation designs, models, methods and tools apply regardless of the subject of the evaluation.

Differences Between Advocacy Evaluation and Other Evaluations

In other ways, evaluating advocacy efforts is different than evaluating programs or services, and can be more challenging. In advocacy work:

- ***Strategies evolve and activities and outcomes can shift quickly.*** Traditional evaluation designs prefer static conditions and aim to keep the intervention from changing over time.
- ***Advocates need real-time data to make informed decisions about the strategic questions they regularly face.*** If the evaluation aims to inform ongoing advocacy strategy, data must be collected and analyzed quickly (e.g., a survey). Traditional evaluations may not prioritize ongoing and rapid feedback.
- ***Contextual factors weigh in heavily.*** It is not possible to control for the many extraneous variables involved with the advocacy and the policy process.
- ***Many factors contribute to policy outcomes.*** It is difficult to isolate the unique contribution of a particular strategy or organization to a policy outcome.

Although evaluating advocacy is a relatively new field, “norms” are starting to emerge:

- ***Integrate evaluation.*** Advocacy evaluation tends to be most beneficial when it is integrated into—and informs—advocacy strategies. Evaluators should stay connected to the advocacy effort to keep abreast of strategy changes and facilitate real-time reporting and feedback.
- ***Examine contribution, not attribution.*** Demonstrating an advocacy effort’s

contribution to policy change is more important than proving attribution. Evaluators should seek to make a plausible and defensible case that the advocacy effort contributed to the policy process.

- **Value interim outcomes.** The end goal—policy change—is not the only important outcome to measure. Other outcomes related to the broader advocacy strategy, such as whether new advocates emerge, can also be important as the policy change itself.
- **Define rigor broadly.** Rigor applied to advocacy means methodological clarity, not necessarily using experimental designs, which are not always applicable to advocacy efforts. (Experimental designs, however, can be used to assess certain tactics, particularly in the communication arena, e.g., public-service announcements and message placements.)
- **Realize that less can be more.** Many advocacy organizations cannot manage highly involved evaluations. A “less is more approach” can be wise when identifying what to evaluate and how to do it.

Steps in Designing an Advocacy Evaluation

Coffman illustrated the steps in designing an advocacy evaluation using a theory of change approach. Theories of change illustrate the pathways by which change is expected to occur and the role that initiatives—or advocacy efforts in this case—play in producing that change. They show how strategies or activities connect to interim outcomes that then set the stage for long-term goals and impacts. Theories of change directly inform and help frame evaluation planning.

The steps and decisions Coffman described are those generally involved in planning any evaluation around a theory of change. For each step, however, she offered thoughts about current advocacy evaluation trends and how to address the unique challenges and opportunities that advocacy presents.

NOTE: Coffman used a hypothetical case study in her presentation. In the case study a statewide health advocacy organization is leading an advocacy effort with the policy goal of universal health care coverage. The effort is comprised of legislative advocacy and grassroots organizing. The desired impact is for the state to have a policy under which all residents have access to the health care they need. PowerPoint™ slides are available [online](#).

This report provides information in a more theoretical way.

Utilization Decisions

These decisions define the evaluation’s parameters and should be made first. Evaluators must answer three questions:

- Who is the evaluation’s audience or its intended users?
- How will the audience use the evaluation?
- What is the evaluation’s timeframe and resources?

1. *Who is the evaluation’s audience or intended users?*

Advocacy evaluations can have many audiences or users. The most common

are advocates or their funders. Evaluators should determine the primary audience(s), focus the evaluation on that audience, and involve audience members in developing the evaluation.

2. How will the audience use the evaluation?

Audiences can use the evaluation for learning or accountability. Learning relates to advocates' need for real-time data to inform ongoing strategies and adjust advocacy efforts while they are in progress. It also can mean using the evaluation to find out which strategies or tactics are effective. Accountability means using the evaluation to hold the advocates responsible for their efforts to funders or other stakeholders and to themselves.

3. What is the evaluation's timeframe and resources?

The evaluation's timeframe and resources help set expectations. While many advocacy efforts are long-term since policy goals can take many years to achieve, the evaluation's timeframe may be shorter. Evaluators must be realistic about what the advocates and the evaluation can accomplish in the available timeframe.

Expectations should also be adjusted based on the available resources for the advocacy effort and the evaluation. The evaluation resources greatly affect methodological choices and the level of evidence expected.

Strategic Focus Decisions

This is the point at which a theory of change is developed and used to help frame the evaluation. Theories of change clarify what advocacy efforts are doing to achieve their intended outcomes and impacts. Once a theory of change is developed, evaluators seek empirical evidence that the theory's components are in place and that the links between them exist.

Coffman presented a new tool developed to facilitate theory of change development—an “advocacy composite logic model” that was developed by more than 50 funders, evaluators and advocates. The composite model offers a menu of possible ingredients that might go into an advocacy theory of change. It offers a full range of advocacy inputs, activities, outcomes, policy goals and impacts—and it allows users to select the components that are most relevant to their work. After selecting the relevant components from the model, evaluators can build a more fleshed-out theory of change that shows the relationships between components.

The model also provides direction about the types of outcomes that can or should be measured for advocacy efforts, beyond achievement of the policy goal. For example, interim outcomes can include advocacy capacity (e.g., increasing organizational capacity or getting new donors) and policy actions (e.g., getting media coverage or building political will).

Once a theory of change is developed, this next set of decisions helps to focus the evaluation, as it is rarely possible (or useful) to measure *everything* about an advocacy effort or every component in the theory of change. Evaluators make strategic decisions by answering these questions:

- Given the evaluation’s audience and use, which outcomes are most important?
- Are there outcomes for which the strategy should not be directly accountable?
- Given the evaluation’s timeframe, which outcomes are achievable?

1. Given the evaluation’s audience and use, which outcomes are most important?

Certain theory of change components or relationships between the components will be more important for some evaluation audiences than others. For example, if the primary audience is the organization leading the advocacy effort, and that organization wants to use the evaluation to get real-time data that will help to improve the advocacy strategy, then the evaluation may want to focus on assessing the activities and interim outcomes that come earlier in the policy change process.

2. Are there outcomes for which the strategy should not be directly accountable?

For some advocacy efforts, Coffman said that certain outcomes or impacts may be so long-term or hinge on so many external or contextual factors that it is appropriate to focus the evaluation more on shorter-term outcomes that are connected more directly to the advocacy effort. Capturing the organization’s unique contribution to the outcomes to which it has the closest links may be more meaningful than capturing outcomes that many organizations or other factors will affect.

3. Given the evaluation’s timeframe, which outcomes are achievable?

Policy goals can take a long time to accomplish. Evaluations, however, may take place on a shorter timeline. With advocacy evaluations it is important to make sure that the evaluation has realistic expectations about what both advocates and the evaluation can accomplish within the available timeframe. Applying a timeline to the theory of change helps ensure that too much is not expected too soon.

Methodology Decisions

Once the strategic decisions have been made, evaluators can determine the evaluation’s methodology. Coffman noted two core questions related to methodology:

- Which design is best?
- Which methods can be used to capture the measures?

1. Which design is best?

Evaluations feature three types of designs—experimental, quasi-experimental, and non-experimental. With advocacy efforts, experiments are rarely, if ever, used; quasi-experiments are more common; non-experimental designs are the most common.

2. Which methods can be used to capture the measures?

Advocacy evaluation can use a range of methods—some traditional, some developed specifically for advocacy evaluation. Traditional methods may

include: public polling, media tracking, policy tracking, interviews, surveys, focus groups, observation and document review. Methods particularly well suited for advocacy evaluation include:

- **Advocacy capacity assessment.** Evaluators can use surveys or an advocacy capacity assessment tool (several have been developed) to determine whether organizations involved in the advocacy effort have increased their advocacy capacity over time.
- **Bellwether methodology.** Through structured interviews with “bellwethers” (influential people who know about and track a broad set of policy issues), evaluators determine an issue’s position on the policy agenda. Coffman’s team at Harvard developed this method. The bellwether methodology can be used to gauge political will.
- **Social network analysis.** Evaluators identify people, groups or institutions that make up the network and use mathematical algorithms to examine the relationships between them and to address the development of partnerships or alliances.
- **Policy-maker ratings.** Evaluators gauge policy-maker support for issues or proposals by asking advocates to rate individual policy-maker’s level of support for the issue and their political influence. Advocates also rate their confidence in policy-maker’s support and influence ratings. Coffman’s team at Harvard also developed this method.
- **Intense period debriefs.** Right after an intense period of activity, evaluators debrief advocates.
- **Blog tracking.** Evaluators use blog search engines to track whether issues or proposals are generating “buzz.” Coffman feels this method will be more useful when blog tracking tools are more reliable.

Summary

Evaluation approaches for advocacy efforts must be adjusted to make the evaluation relevant and useful within an advocacy and policy context. Coffman suggested that advocacy evaluators:

- **Be as real-time as possible.** Evaluating advocacy is a learning process for most evaluators. Evaluation is most useful when it can be used to inform strategy regularly. Coffman combines rapid response research with longer-term evaluation.
- **Be strategic and prioritize.** Most advocates have limited resources and must do some data collection on their own. Be strategic about what is evaluated to make sure the evaluation is useful and not overly burdensome.
- **Focus on progress, not just the policy goal.** Assess interim outcomes to capture contributions to policy change. What happens during the process of achieving policy change should be the focus of most advocacy evaluations.
- **Stay current and be innovative.** Advocacy tactics are constantly changing and growing. Evaluators must keep pace and develop new methods to capture interim outcomes.

Discussion

Arkin began the discussion by presenting a RWJF advocacy case study, Cover the Uninsured Week. This annual event includes grassroots organizing, prominent national partnerships and events to engage many sectors of society. The interim goal of Cover the Uninsured Week is to put the issue on the nation's agenda and build a broad base of support for it. Cover the Uninsured Week involves issue reframing, public will, media coverage and constituency or support base growth.

The evaluation focuses on:

- What was changing?
- What works?
- What does not work?
- What needs to happen next?

It includes many methods:

- Public polling to determine whether the effort was making progress in public understanding of the issue and its importance;
- Media tracking to show which strategies work, which issues the media want to cover and what people are seeing and hearing about the issue;
- Policy tracking;
- Personal interviews with representatives of partners that co-sponsor and financially support Cover the Uninsured Week¹ and selected coalitions about what worked and did not work during the campaign, and future plans;
- A survey of other national organizations (such as professional organizations and national associations) that participated in Cover the Uninsured Week on what worked and did not work during the campaign, and future plans;
- Focus groups, used to plan for the next year;
- Observation to track the type of events held and whether they are likely to move policy.

Arkin noted that RWJF sometimes changes evaluation plans during Cover the Uninsured Week. The evaluation focuses on producing immediate results and interim policy goals. Evaluation has been an integral part of the event and is necessary to keep people engaged.

The discussion then continued more broadly. David Morse, vice president for Communications at RWJF, noted that many RWJF-funded campaigns seek to change political will, which is difficult to measure.

Coffman agreed that political will is difficult to measure and said that Harvard is trying to use two methods they developed—the bellwether methodology and the policy-maker rating methodology—to address this issue. Jacobs added that he interviews elites, reads reports and tracks floor debate and bills coming up for votes.

¹ Partners are national organizations such as the U.S. Chamber of Commerce, AFL-CIO, United Way of America and foundations.

An audience member suggested that a good advocate would know through relationships with Capitol Hill staffers whether an issue will move. Jacobs replied that staffers have limited time and that coalitions are more effective than advocates in presenting issues. He also noted that environmental factors—such as partisan and ideological composition of legislative and executive branches, public mood (liberal or conservative) and major shift in party control of government—affect issues. Peterson added that advocacy campaigns often last many years, but staffers come and go.

Another audience member asked how free the evaluators feel to tell a funder that a program should not be evaluated or that the proposed evaluation will not yield meaningful findings. Jacobs replied that he has done this several times and believes it enhances his credibility.

Dwayne C. Proctor, Ph.D., a senior program officer at RWJF, asked whether Coffman addresses counter advocacy efforts in evaluations. Coffman has not incorporated assessments of counter advocacy in her evaluations, but she expects to do so in the future.

An audience member commented that to achieve change, it is more important to get close to people on the ground than Capitol Hill staffers. Jacobs responded that while it is important to get close to people on the ground, who to focus on depends upon the issue at hand. Lester Baxter of Pew added that people with a constituency, such as those representing a grassroots organization, are some of the most effective spokespeople he has seen.

Speaker List**Julia Coffman**

Consultant
Harvard Family Research Project
Harvard Graduate School of Education
Cambridge, Mass.

David Colby, Ph.D.

Vice President
Research & Evaluation
Robert Wood Johnson Foundation
Princeton, N.J.

Robert Hornik, Ph.D.

Wilbur Schramm Professor of Communication
University of Pennsylvania, Annenberg School of Communication
Philadelphia, Pa.

Lawrence R. Jacobs, Ph.D.

Walter F. and Joan Mondale Chair for Political Studies
University of Minnesota
Minneapolis, Minn.

Laura C. Leviton, Ph.D.

Special Advisor for Evaluation
Research & Evaluation
Robert Wood Johnson Foundation
Princeton, N.J.