Three Tools for Evaluating Participation: Focus Groups, Q Method, and Surveys

Stentor Danielson, Seth P. Tuler, Susan L. Santos, Thomas Webler, Caron Chess

To enable successful public participation in environmental decision making, practitioners need to know what works, but evaluation of proposed and existing processes is often lacking. We tested three tools for evaluation—focus groups, Q method, and surveys—at two contaminated sites with extensive public participation. Each tool is evaluated based on its requirements for implementation, the information it produces, and its acceptability to stakeholders. Which tool is most appropriate depends heavily on the available resources, what is happening at the site, and the evaluator's goals.

Environmental Practice 14:101–109 (2012)

The past few decades have seen a major surge in interest from academics and practitioners in public participation in environmental decision making. Important practical and moral benefits are to be had from giving affected communities a greater voice in policy making (Dietz and Stern, 2008). Numerous studies have probed the conceptual underpinnings of participation (e.g., Fiorino, 1989; Renn, Webler, and Wiedemann, 1995), as well as the utility of particular features of participation processes (e.g., Rowe and Frewer, 2005; Webler and Tuler, 2006).

Though we now know much about what works and doesn't work as a general rule in public participation, it remains vital for those organizing a public participation process to conduct evaluation of their specific site to examine how their approach is working and seek solutions to particular problems that have arisen. Consider the following different scenarios:

- Outreach to various sectors of the affected public has been uneven, with some (social, occupational, ethnic, geographic) groups not responding to the responsible agency's efforts. How can this barrier be overcome?
- A public participation process has been successful so far, but a new phase of activity (such as cleanup of a new portion of a contaminated site) is about to begin. What hidden cleavages among participants might come to the fore now?
- A new situation has generated relatively little proactive public interest and demand for participation so far. Are there concerns that the responsible agency should nevertheless be aware of that could generate a need for a more extensive participation process?

Site-specific evaluation is sometimes recommended by agency policy—for example, the United States Environmental Protection Agency (USEPA, 2003) requires evaluation of all community involvement activities (see also Dietz and Stern, 2008). Unfortunately, formal reviews of evaluation have reported that it is often unsystematic and based on project leaders' impressions and complaints from the most vocal citizens (Chess, 2000; National Research Council, 2001; USEPA, 2001). This is not surprising given the challenges—time and money costs, staff skills, administrative constraints, etc.—to systematic evaluation. It is therefore important to develop flexible evaluation tools in order to foster more widespread use of formal evaluation.

This article compares three tools that have shown promise in evaluating public participation: focus groups, Q method,

Affiliation of authors: Stentor Danielson, Assistant Professor, Department of Geography, Geology, and the Environment, Slippery Rock University, Slippery Rock, Pennsylvania. Seth P. Tuler, Research Fellow, Social and Environmental Research Institute, Greenfield, Massachusetts. Susan L. Santos, Assistant Professor, Department of Health Education and Behavioral Science, School of Public Health, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey. Thomas Webler, Research Fellow, Social and Environmental Research Institute, Greenfield, Massachusetts. Caron Chess, Assistant Professor, Department of Human Ecology, Rutgers University, New Brunswick, New Jersey.

Address correspondence to: Stentor Danielson, Department of Geography, Geology, and the Environment, Slippery Rock University, 1 Morrow Way, Slippery Rock, PA 16057; (phone) 724-738-2564; (fax) 724-738-4217; (e-mail) stentor. danielson@sru.edu.

doi:10.1017/S1466046612000026

and surveys. On the basis of applying these three tools at two case-study sites, we compare their strengths and weaknesses and suggest for which situations professional evaluators and public participation coordinators might find each tool appropriate.

The Case Studies

Our first case study was the Ciba-Geigy Chemical Corporation Superfund site in Toms River, New Jersey (USEPA, 2009a). This site is the former home of a chemical plant that manufactured dyes, resins, and other products from the 1950s through the 1990s. Dangerous levels of volatile organic compounds were found in the water and soil. Extensive public involvement shaped the choice and design of the two remediation phases significantly: a water purification plant that came on line in 1996 and a soil bioremediation scheme that operated from 2003 to 2009. We conducted test evaluations using the three tools during 2004.

Our second case study was the Waukegan Harbor Area of Concern, a complex of contaminated sites under the EPA's Great Lakes Program (USEPA, 2009b). In Waukegan, Illinois, cleanup of polychlorinated biphenyls (PCBs) and other toxins in the harbor sediments is tied in with efforts to revitalize the city's downtown and lakefront. An agreement was worked out by the Citizens Advisory Group (CAG)—which brought together city, state, EPA, and citizen interests—to guide cleanup. We evaluated the three tools in 2006–7. Shortly after the completion of our fieldwork, the city surprised other CAG members by backing out of the agreement and proposing an alternate cleanup plan (Zawislak, 2007).

The Three Tools

A focus group is a guided group discussion among 6–12 people about their views of some topic. Participants' own words can be analyzed to identify the major themes both within and across groups. In our case studies, we held focus groups in each location with (a) government officials, (b) highly involved community activists, and (c) a general public group in Toms River, New Jersey, and a Latino community group (conducted in Spanish) in Waukegan. The specifics of how to conduct a focus group can be found in the guidance document we prepared following this project (Santos et al., 2007), as well as in other published guides (Kruger, 1988; Morgan and Kruger, 1998).

Q method is a procedure by which a series of individuals each represent their viewpoint by ranking a set of statements (typically by arranging small cards in a quasi-normal distribution). The resulting rankings—called *Q sorts*—can be factor analyzed to reveal the nature of a smaller number of shared perspectives. We had 19 participants in Toms River, and 17 participants in Waukegan (4 completing their sorts in Spanish). The specifics of how to conduct Q method can be found in the guidance document we prepared following this project (Webler et al., 2007), as well as other published guides (Brown, 1980; McKeown and Thomas, 1988).

A *survey* involves administering a questionnaire to a random sample of people, asking closed-ended questions about some topic. These questions can be analyzed by using statistical procedures to reveal generalizations about the views of a wider population. Our surveys closely followed the *tailored design method* for mail surveys laid out by Dillman (2000). In Toms River, we sent out 562 surveys and had a response rate of 34%. In Waukegan, we sent out 401 surveys (in both Spanish and English versions), with a response rate of 32%. In Toms River, we used the EPA's existing evaluation survey (Charnley and Engelbert, 2005), but, in Waukegan, we designed our own instrument to match the circumstances of the site more closely.

In this article, we review the three evaluation tools in terms of three general families of criteria: implementation, information content, and participant reaction. These cover three major questions that any evaluator would need to ask: How feasible will it be to carry out this evaluation? What information will I actually get out of doing it? Will stakeholders accept and be able to use the results?

Implementation

Time

Each evaluation tool requires a significant time to prepare, as well as to actually implement it and to analyze the results. We omit the time we spent conducting introductory interviews and familiarizing ourselves with the site since that work was common to all three tools and will vary depending on the evaluator's familiarity with the site. Because individual sites differ greatly in their circumstances, and evaluators differ in their skill levels (members of our project team were generally highly skilled), our

estimates should be treated as only ballpark figures for comparison purposes.

For a focus group, the major time investments are preparing a discussion guide (8 hr), contacting participants (5 min per confirmed attendee), securing the logistics (4 hr), the actual focus group meeting ($1\frac{1}{2}$ hr of group time plus 2 hr of setup and teardown per group), transcribing the recordings (8 hr per group), creating the codebook (8 hr total), coding the results (4 hr per coder per group—we had two people code each group), and writing up the analysis (10 hr). In all, conducting three focus groups can involve 82 hr of work spread over two months or more.

Q method requires the following time commitments: selecting Q statements and printing materials (10 hr), contacting participants and arranging times (5 min per confirmed participant), conducting the Q sorts (up to 2 hr per participant), and analyzing the data (10 hr). A workable Q analysis with 15–20 participants can be done in 62 hr spread over two months.

A survey involves writing a questionnaire, if necessary (20 hr), pretesting and translating the survey, if necessary (which will depend greatly on the pretesting method chosen), obtaining and formatting the list of respondents (4 hr), preparing mailings (three days for commercial printing and then 20 hr of packet assembly for a 400-person sample), waiting for responses (approximately six weeks from sending out the initial mailing until the last questionnaires are returned), coding the data (20 hr for 200 surveys), and analysis [10 hr for a simple analysis using *t* tests and analysis of variance (ANOVA)]. A typical survey will involve 54 hr of work over two months, plus additional time for questionnaire preparation.

Money

For Q method, monetary costs are relatively low. A few dollars' worth of paper, cardstock, and ink must be used to print the Q cards and data sheets. A variable amount of travel cost will be incurred in visiting each participant at their preferred location (typically a home or workplace). We offered all participants a $30 honorarium to thank them for their time (for a total of around $450 at each site).

Focus groups incur a somewhat higher monetary cost. Table 1 lists a tabulation of our typical costs for the variety of items needed for a focus group. As with the Q sort, we offered $30 honorariums to all participants. The cost of room rental may be avoided if a free community space is used. Additional costs may be incurred if the recordings are professionally transcribed, or if the results of a non-English focus group need to be translated (transcription and translation service fees vary widely, but $200 for a focus group may be typical).

Table 1. Focus-group costs

Cost per case	
Recorder	$40 and up
Microphone	$20 and up
Room rental	$300
Cost per focus group	
Snacks and drinks	$25
Cost per participant	
Honorarium	$30
Total for 3 focus groups of 8 people each	$1,140

Surveys are the most expensive evaluation tool of the three we examined. The cost of a survey is highly dependent on the number of surveys sent out. Each respondent may cost as much as $4 when cover letters, survey booklets, mail-out envelopes, return envelopes, postage, and multiple contacts are considered. One significant element in the cost of a survey is a token gift—several dollars inserted into the survey envelope. We were unable to include token gifts in this study because of to EPA regulations, but studies have shown that token gifts impact on response rates significantly (Dillman, 2000). Without the token gifts, our survey at each site cost about $2,000 for an initial sample size of 400.

Human Capital

Focus groups require an experienced facilitator. A neutral, independent facilitator is more likely to gain the trust of the participants and run a process that is viewed as credible. Analyzing focus-group transcripts requires an individual trained in qualitative data analysis, but no specialized software is necessary.

Conducting Q sorts is relatively easy as long as the facilitator understands the basics of the Q procedure. Preparing Q statements and analyzing Q-method results, on the other hand, do take a person trained in the use of statistics and Q software (such as the free PQMethod[1]).

The quality of a survey depends on selecting appropriate questions and responses, which requires training in survey technique if a predesigned survey instrument is not used. The analysis of results does require a trained professional,

but the statistical tests can be simple ones available in common spreadsheet software, for instance *t* tests, chi-squared, and correlation coefficients.

Information Content

From Whom Do We Learn?

At each study site, it proved useful to distinguish among involved residents, uninvolved residents (aka the general public), elected officials, unelected officials, the responsible party, and the regulatory agencies (e.g., the EPA). Some tools were better than others at tapping the opinions of some of these stakeholder groups, as shown in Table 2.

Focus groups can be used to target most kinds of stakeholders, but are less useful for eliciting information from elected officials and the responsible party. Ideally each group would involve 8–12 participants who are broadly similar in terms of their relationship to the environmental decision. At the same time, focus groups are strengthened when group members vary on demographic variables such as ethnicity, gender, age, and length of residence. Focus groups try to avoid mixing participants who are deeply divided on key positions and thus would not be able to have a productive discussion. If there are contending activist groups, it may be necessary to hold two or more focus groups. At our Toms River case study, participants told us directly that the focus groups would not have worked earlier in the remediation process because of hard feelings between participants who would have been in the same focus group.

In our project, unelected officials (such as town planners and EPA staff) and highly involved residents readily agreed to participate in focus groups. To encourage participation by less highly involved residents, we reached out to them through institutions they trusted (such as social clubs) and framed the project in a way that reassured them that they have something to contribute and would not need to "study up" for the session.

We invited a variety of local elected officials to participate in our focus groups, but only one attended—a situation another participant attributed to the need to maintain a political posture, which made the focus-group form of discussion unattractive to them. We did not try to invite representatives of the responsible party (the companies that created the pollution being cleaned up) to a focus group because they do not fit neatly into a large enough group of stakeholders to constitute a focus group on their own. The perspectives of both of these types of stakeholders may be tapped through individual interviews instead.

Because Q method is a one-individual-at-a-time tool, it is not subject to the factors that make elected officials and company representatives reluctant to talk in group settings. We completed Q sorts with people in all stakeholder categories except the general public in both case-study locations. Q's crucial limitation is that it requires substantial knowledge about the site in order to complete a meaningful Q sort. Residents with a low level of knowledge or involvement may be able to complete a Q sort, but it may not reveal anything relevant about their attitudes (Danielson, 2007).

Surveys are the most effective tool for getting input from a broad cross section of members of the general public, including those who have had little involvement with the site. The ability of a survey to measure attitudes of uninvolved residents will depend on the questions asked. Detailed questions about specific events or cleanup technologies may result in meaningless answers, nonanswers, or even refusal to complete the survey. Surveys also require less expenditure of time and mental effort if most questions are closed-ended, thus overcoming some of the barriers to participation in a focus group or Q. The EPA, government officials, and responsible company stakeholder groups would generally be too small for any results from surveying them to be subjected to meaningful statistical analysis.

Table 2. Who do we learn from?

	Focus group	Q method	Survey
Involved residents	✓	✓	✓
Uninvolved residents	✓–		✓
Elected officials		✓	✓–
Nonelected officials	✓	✓	✓–
Agency	✓	✓	
Responsible company		✓	

Checkmark (✓) indicates the method is well-suited to tap this group's perspectives. Checkmark with a minus (✓–) indicates the method can be used to tap this group's perspectives but may present challenges.

What We Learn About

A significant advantage to a focus group is that participants can share both their opinions (e.g., that the obser-

vation deck to view cleanup progress is ineffective) and the underlying rationale for them (because a pile of dirt blocks the view from the deck). Furthermore, other group members may hear these opinions and rationales and respond to them. In this way, one can also learn about how firmly views are held by people in the group.

Q can document opinions and preferences, but has more difficulty establishing the reasons behind them. The researcher must verbally probe about the rationale for particular sorting choices. Q is also more limited than either focus groups or surveys in how many different types of information it can cover in one study. For example, in our studies we asked participants to do one Q sort about what their ideal public participation process would be like and another evaluating the cleanup outcomes that had occurred thus far at the site. This meant that we were unable to investigate direct evaluations of the actual participation process or suggestions for future cleanup options. These topics could be explored with Q, but because they would need separate Q sorts, they would require sacrificing the topics we did address (or demanding more time from participants).

Survey information consists of aggregations of individual views across a representative sample of respondents. Like Q, surveys largely elicit opinions and preferences. It may be possible to get some indication of rationales by presenting respondents with a set of justifications and asking them to rate them. Nevertheless, this limits respondents to rationales chosen by the survey designer rather than allowing them to express themselves in their own words.

Richness and Depth of Information

Focus groups provide the most opportunity for developing an in-depth understanding of people's viewpoints. Participants have the opportunity to speak about their views in their own words, and the moderator can probe for more information. Group dynamics can lead to deeper discussion because each person's comments are elaborated on or challenged by others. Thus, for example, we not only learned that people in Toms River are concerned about the toxic drums in the landfill and dissatisfied with the EPA's responsiveness, but we also heard it in "Julianne's" words:

> Well, I guess at the meeting when we talked about removing the drums, and they just said "they're fine." That was our formal feedback. They don't have to remove them. . . . I learned more about how those drums are contained sitting next to this gentleman [another participant who had earlier explained how a landfill liner might leak] here. You know, when they say contained, it's contained in a sealed, lined thing. I thought it was lined on the top.

Q provides a moderate amount of depth, though it gives relatively thin information about any single issue, since participants are merely ranking short statements. But the resulting shared perspectives have the advantage of bringing together people's evaluations of all of the statements in a holistic way so that the rankings of one statement can shape the interpretation of others. Q can demonstrate the relative importance of various issues more easily than focus groups or surveys. Comments made by participants explaining why they chose to sort the statements in the way that they did can enrich the description of the perspectives.

Surveys provide the least depth of understanding of rationales behind opinions. Most of the information obtained from a survey comes in the form of simple numerical responses to Likert scales and yes/no responses to checkbox questions. This austerity of the data makes aggregation and comparison of the responses of a large set of participants relatively easy. But it also leaves the results fairly thin and very dependent on the question design.

An example illustrating these differences is the question of the best way to reach out to the Latino community in Waukegan. At the focus group, Latino community members identified social clubs as one of the best ways to provide them with information about the cleanup. They did not mention churches. This was a surprise to the focus groups with mainly non-Latino participants, including highly involved residents and government officials, who asserted strongly that churches were the best way to reach the Latino community.

This kind of information was not revealed by the survey or Q because there were no questions in those instruments that anticipated this issue. The survey asked respondents to rate "Presentations at local clubs and organizations" in response to "How would you prefer to receive information about cleanup activities?" This item was not significant overall or when responses were broken down by ethnicity (Latino vs. white non-Latino). Notice, however, that it is not asking about ethnic social clubs specifically or about the relative benefits of social clubs versus churches.

The same was the case for the Q study. We did not include statements about outreach through social clubs or churches, so this issue could not be assessed. Even if such a statement

had been included, participants may not have ranked the statement high enough relative to all of the other statements. In the case of the focus group, it is not that the idea was very important, but rather that role of social clubs was not consistent with the prior beliefs of highly involved residents and officials.

An important caveat, however, is that social clubs were a key avenue through which we recruited people for the Latino focus group, whereas churches were not. Thus, we cannot make a generalization to the entire Latino community that social clubs are better for outreach than are churches (whereas we could say such a thing had our survey asked separately about these two avenues of outreach). The focus-group results should be considered merely a suggestion that should be examined.

When evaluation is conducted in order to improve the process, one important outcome ought to be specific suggestions for how the process should be changed. Focus groups can give the richest input on suggestions for change. The moderator can engage participants in extensive discussion or brainstorming about possibilities. People can debate the particulars of implementation, trade-offs, etc. However, it is important not to let participants start to view the focus group as a planning workshop or charette, lest they feel betrayed if their suggestions are not implemented.

Q method presents more difficulties for producing suggestions. Depending on the nature of the statements and sorting instructions, it is possible for participants to outline a general vision for how the process should be set up. For example, our Q sorts revealed a difference between a vision of the public as watchdogs versus the EPA as educator in Toms River. A key advantage of Q is that it covers all topics within each individual. If there is disagreement among participants at a focus group, it may be difficult to clearly separate the distinct agendas of different camps. (For example, were the people who favored off-site disposal the same as the ones who wanted deed restrictions?) Q method, on the other hand, allows you to see the full agenda of each person distinctly.

Surveys face many of the same limits as Q because surveys involve individuals responding to a set of prompts chosen by the researcher, although the level of detail is highly dependent on what questions are chosen. A general-purpose survey suitable to multiple sites, such as the EPA currently uses (Charnley and Engelbert, 2005), cannot elicit context-specific opinions.

Comparability, Robustness, and Generalizability of Elicited Information

Depending on the goals of the evaluation, it may be important to have comparable, robust, and generalizable information elicited from the participants. Data gathered in a survey are the most comparable across respondents. In a survey, all respondents answer the same questions. The principles of questionnaire design emphasize ensuring that all respondents read and understand the question in the same way (Presser et al., 2004). The responses to closed-ended questions produce numerical data that can be easily summed, averaged, or otherwise combined.

In a focus group, comparability depends to a great extent on the moderator, who can direct the participants to respond to the same topics from one group to the next. For example, in our Toms River focus groups, officials gave themselves high grades for responsiveness, but the highly involved group—and to some degree the general public group—felt that the EPA had not been as responsive to their concerns as they wanted. Members of the highly involved group credited themselves with having "pushed" the EPA and viewed themselves as advocates for the public. The general public credited the EPA for being responsive to concerns about remediation. In these cases, asking about the same topic enabled comparison between the responses of different groups. However, it is often the case that different groups will discuss different content, even if the same general topics are raised by the moderator. Note, as well, that just because each *group* discussed the same topic does not mean each *participant* weighed in on it.

Comparability of results gathered via Q method is high because, as in a survey, all participants respond to the same statements that are all analyzed equivalently. Shared perspectives are identified based on differences in people's responses to the entire slate of statements. For example, one statement in the Waukegan Q study asserted that the process should ask members of the public what they think. This principle is ranked very high by the one perspective, but fairly low in a second perspective. Based on the statistical analysis of the results and qualitative information gathered by the researcher, we concluded that the second perspective feels that spending too much time reaching out and listening to the public may delay the process and make it more costly.

Robustness refers to how well the responses authentically reflect the beliefs and opinions of the respondent. If a person has thought a great deal about the topic at hand,

then they are likely to voice equally robust responses no matter which method is used. However, if the person has not reflected on their opinions or develops them over time in discussions with colleagues and peers, then they may respond in ways that do not reflect any stable beliefs (Lichtenstein and Slovic, 2006). Surveys are notoriously vulnerable to producing nonrobust results (Bishop, 2005), because they often reach people who are not deeply involved. Q method is also vulnerable to this (Danielson, 2007), but less so because the evaluator is there and can ask the participant to explain apparent inconsistencies in their Q sort, which may also help prevent participants from giving spurious answers. Focus groups are the best for ensuring robust results because they equip the opinion-collection process with an automatic peer review. If someone says something spurious or nonsensical, others at the table can request clarification and justification.

Generalizability addresses the question of whose opinions are represented by the small number of respondents. Surveys produce the most generalizable results, if the sample to which the survey was mailed was properly composed, for example by being randomly selected, and the response rate is sufficient. Rates below 50% are common for surveys of the uninvolved public (Connelly, Brown, and Decker, 2003). Even a low response rate can produce acceptable results if the responses are not biased by one type of person being more likely to return the survey.

Neither focus groups nor Q draw on a random sample of the population. Even if participants in a focus group or Q study were chosen randomly from a larger population, the number who participate is too small to enable statistically valid generalizations with any reasonable confidence interval. Nevertheless, focus groups and Q method often achieve a near census of individuals in certain smaller stakeholder categories such as leading community activists whose distinctive views would be lost in the aggregate if they were included in the survey sample.

Participant Reaction

In order to result in meaningful change in a process, evaluations need to be acceptable to participants and decision makers. We administered short questionnaires after the focus groups and Q sorts to obtain participants' views about the evaluation tool itself, and held a review meeting with key stakeholders following the writing of our reports, to elicit feedback about how acceptable each tool was to them.

Evaluation Process

During the focus-group process, participants were generally comfortable expressing themselves. Participants consistently reported that participation in the focus group stimulated their thinking about their own views and others' views, and was a good use of their time.

Participants were generally happy to complete the Q sorts, and some found the process interesting or stimulating. Nevertheless, several people said they felt constrained in their ability to express themselves.

Participants' views of the survey process are difficult to discern because our only data are the surveys themselves. The response rates were somewhat lower than we might have hoped [though comparable to the response rates reported by Charnley and Engelbert (2005)], but it is unclear whether this was due to lack of time, lack of familiarity with and interest in the issues, perceived lack of useful comments, misplaced or lost surveys, strong desire for privacy, or other reasons.

Evaluation Outcomes

Our focus-group report got the most enthusiastic response, as members of our feedback group found its results the easiest to understand, the most useful to them, and the most trustworthy as a representation of the situation. The narrative style and use of illustrative quotes from participants were cited as reasons to favor the focus-group report.

Q method, on the other hand, generated a significant amount of resistance. Q raised some concerns about trust and manipulation of the results. It was easier to see how data were transformed into results in the case of a focus group or survey. Respondents were also confused at times as to what precisely the Q-method results were supposed to be telling them. Presentation of Q results should focus on the narrative descriptions of the perspectives that it reveals. Nevertheless, a number of participants did find the Q results interesting and informative. For example, in Waukegan we were told that our Q results (more so than the focus group or survey) predicted the fault lines along which the process broke down when the city backed out of the cleanup agreement.

The biggest hurdle with respect to the survey report was many stakeholders' discomfort with quantitative data. Ju-

dicious use of charts and moving numerical data to appendices may help make survey results more accessible.

Conclusions

The three tools for evaluation that we tested have complex patterns of strengths and weaknesses. Moreover, whether a certain feature is a strength or weakness depends on context.

In lieu of a simple summary or reductionist chart, we conclude with three vignettes revisiting the three scenarios described in the introduction, each of which illustrates a type of situation where one tool may be the most useful. These vignettes are meant to show how the many features described above may factor into a choice of how to evaluate a public participation process. The vignettes are based on contaminated site cleanups, but the principles can apply more broadly.

1. Seeking Solutions: Focus Groups

The public participation process at a site is stuck. Despite the agency's best efforts, outreach to certain (social, occupational, ethnic, geographic) populations has not produced much engagement or interest. Here, focus groups may be a good way to diagnose the problem and elicit suggestions. The agency and highly involved stakeholders can be brought together to discuss their perceptions of the problem and outline what efforts they have made. Another set of focus groups can promote feedback from the hard-to-reach populations. As these people talk to one another, the moderator can probe their reactions to existing outreach efforts and encourage them to brainstorm solutions that would be more suitable for their situation and needs. The group discussion dynamics among similarly situated individuals would encourage sharing and evaluating new possibilities.

2. Hidden Agendas: Q Method

A site has been operating for some time, and all of the major stakeholders seem to have staked out their positions. Yet there is another phase of the project on the horizon—a phase that will require extensive planning. This raises two dangers: First, deep disagreements that were hidden by the compromise on the current phase may erupt again as a new project is confronted. Second, entrenched battle lines may inhibit stakeholders from seeing common ground between those who reject the current plan and those who endorse it. Here, Q method could be useful. Meeting with stakeholders one on one can reduce posturing and reveal what stakeholders believe are key issues addressed (or not) in the current situation. A factor analysis of a consistent set of statements can reveal unexpected correspondences and divisions that may run perpendicular to existing coalitions. The resulting descriptions of the discourses can help stakeholders predict, and manage, fault lines that were obscured in the current phase.

3. Assessing Concerns: Survey

A new source of contamination has recently been identified, and the agency has stepped in to investigate cleanup options. A few vocal individuals have spoken out, but it is unclear how many people they speak for, and no well-defined interest groups have formed yet around the cleanup issue. In this case, the agency's primary interests are in assessing how widespread knowledge and concern is about the site, what kind of outreach options will get the best response, and whether there are particularly vulnerable communities whose situation the agency needs to address. Two key elements of a survey make it the most suitable evaluation tool. First, it can tap the perspectives of a broad, statistically representative cross section of the community. This enables it to establish the prevalence of concerns and knowledge in the larger population, and give a more equal voice to disadvantaged groups. Second, the low time commitment and knowledge demand mean that people will not be put off by embarrassment at their own ignorance.

The results of this study illustrate that all three evaluation tools show promise for better evaluation of public participation. Formal evaluation is critical for decision makers to understand what works and doesn't work in particular cases. The choice of evaluation tool ought to be driven not by convenience or habit, but by a careful consideration of the strengths and weaknesses of the available tools.

Acknowledgments

We would like to thank the people of Toms River and Waukegan, as well as the government and corporate staff at those sites, for contributing their time to this project. This research was funded by cooperative agreement 831219-01-3 with the US EPA.

Note

1. Available at http://www.lrz.de/~schmolck/qmethod/.

References

Bishop, G.F. 2005. *The Illusion of Public Opinion: Fact and Artifact in American Public Opinion Polls*. Rowman and Littlefield, Lanham, MD, 248 pp.

Brown, S.R. 1980. *Political Subjectivity: Applications of Q Methodology in Political Science*. Yale University Press, New Haven, CT, 369 pp.

Charnley, S., and B. Engelbert. 2005. Evaluating Public Participation in Environmental Decision-Making: EPA's Superfund Community Involvement Program. *Journal of Environmental Management* 77(3):165–182.

Chess, C. 2000. Evaluating Environmental Public Participation: Methodological Questions. *Journal of Environmental Planning and Management* 43(6):769–784.

Connelly, N.A., T.L. Brown, and D.J. Decker. 2003. Factors Affecting Response Rates to Natural Resource–Focused Mail Surveys: Empirical Evidence of Declining Rates over Time. *Society & Natural Resources* 16(6):541–549.

Danielson, S. 2007. *Discourses about Wildfire in New Jersey and New South Wales* (unpublished PhD dissertation). Clark University, Worcester, MA, 328 pp. Available at http://debitage.net/academic/DissFull.pdf (accessed June 8, 2010).

Dietz, T., and P.C. Stern. 2008. *Public Participation in Environmental Assessment and Decision Making*. National Academies Press, Washington, DC, 322 pp.

Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*, 2nd edition. Wiley, New York, 480 pp.

Fiorino, D.J. 1989. Technical and Democratic Values in Risk Analysis. *Risk Analysis* 9(3):293–299.

Kruger, R.A. 1988. *Focus Groups: A Practical Guide for Applied Research*. Sage, Beverly Hills, CA, 197 pp.

Lichtenstein, S., and P. Slovic, eds. 2006. *The Construction of Preference*. Cambridge University Press, Cambridge, 808 pp.

McKeown, B., and D. Thomas. 1988. *Q Methodology*. Sage, Beverly Hills, CA, 83 pp.

Morgan, D.L., and R.L. Kruger, 1998. *The Focus Group Guidebook*, volume 1: *The Focus Group Kit*. Sage, Thousand Oaks, CA, 103 pp.

National Research Council. 2001. *A Risk Management Strategy for PCB-Contaminated Sediments*. National Academies Press, Washington, DC, 452 pp.

Presser, S., J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, and J. Martin, eds. 2004. *Methods for Testing and Evaluating Survey Questionnaires*. Wiley Series in Survey Methodology. Wiley, Hoboken, NJ, 624 pp.

Renn, O., T. Webler, and P.M. Wiedemann. 1995. *Fairness and Competence in Citizen Participation: Evaluating Models for Environmental Discourse*. Kluwer, Dordrecht, The Netherlands, 381 pp.

Rowe, G., and L. Frewer. 2005. A Typology of Public Engagement Mechanisms. *Science, Technology, & Human Values* 30(2):251–290.

Santos, S.L., S. Danielson, and C. Chess. 2007. *Guidance on the Use of Focus Groups for Evaluation of Public Involvement Programs at Contaminated Sites*. Social and Environmental Research Institute, Greenfield, MA, 40 pp. Available at http://www.seri-us.org/sites/default/files/FGGuidanceSuperfund.pdf (accessed April 9, 2012).

US Environmental Protection Agency (USEPA). 2001. *Stakeholder Involvement & Public Participation at the U.S. EPA: Lessons Learned, Barriers, & Innovative Approaches*. USEPA, Washington, DC, 31 pp. Available at http://www.epa.gov/stakeholders/pdf/sipp.pdf (accessed June 8, 2010).

US Environmental Protection Agency (USEPA). 2003. *Public Involvement Policy of the U.S. Environmental Protection Agency*. USEPA, Washington, DC, 33 pp. Available at http://www.epa.gov/publicinvolvement/pdf/policy2003.pdf (accessed June 8, 2010).

US Environmental Protection Agency (USEPA). 2009a. *Ciba-Geigy Superfund Site Remediation*. USEPA, Washington, DC. http://www.cibageigysuperfund.org/ (accessed June 8, 2010).

US Environmental Protection Agency (USEPA). 2009b. *Waukegan Harbor River Area of Concern*. USEPA, Washington, DC. http://www.epa.gov/glnpo/aoc/waukegan.html (accessed June 8, 2010).

Webler, T., S. Danielson, and S.P. Tuler. 2007. *Guidance on the Use of Q Method for Evaluation of Public Involvement Programs at Contaminated Sites*. Social and Environmental Research Institute, Greenfield, MA, 43 pp. http://www.seri-us.org/sites/default/files/QMethodGuidanceSuperfund.pdf (accessed April 9, 2012).

Webler, T., and S.P. Tuler. 2006. Four Perspectives on Public Participation Process in Environmental Assessment and Decision Making: Combined Results from 10 Case Studies. *Policy Studies Journal* 34(4):699–722.

Zawislak, M. 2007. County Holds Off on Harbor Cleanup Funds. *Chicago Daily Herald*, June 27, p. 4.

Submitted September 14, 2011; revised January 17, 2012; accepted January 24, 2012.