Public Participation Methods: A Framework for Evaluation
Gene Rowe and Lynn J. Frewer
*Science Technology Human Values* 2000 25: 3
DOI: 10.1177/016224390002500101

The online version of this article can be found at:
http://sth.sagepub.com/content/25/1/3

Published by:
http://www.sagepublications.com

On behalf of:
Society for Social Studies of Science

Additional services and information for *Science, Technology & Human Values* can be found at:

Email Alerts: http://sth.sagepub.com/cgi/alerts
Subscriptions: http://sth.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://sth.sagepub.com/content/25/1/3.refs.html

>> Version of Record - Jan 1, 2000

What is This?
Public Participation Methods:
A Framework for Evaluation

Gene Rowe
Lynn J. Frewer
Institute of Food Research

There is a growing call for greater public involvement in establishing science and technology policy, in line with democratic ideals. A variety of public participation procedures exist that aim to consult and involve the public, ranging from the public hearing to the consensus conference. Unfortunately, a general lack of empirical consideration of the quality of these methods arises from confusion as to the appropriate benchmarks for evaluation. Given that the quality of the output of any participation exercise is difficult to determine, the authors suggest the need to consider which aspects of the process are desirable and then to measure the presence or quality of these process aspects. To this end, a number of theoretical evaluation criteria that are essential for effective public participation are specified. These comprise two types: acceptance criteria, which concern features of a method that make it acceptable to the wider public, and process criteria, which concern features of the process that are liable to ensure that it takes place in an effective manner. Future research needs to develop instruments to measure these criteria more precisely and identify the contextual and environmental factors that will mediate the effectiveness of the different participation methods.

In recent years, there has been increased interest in involving the public in decision making about science and technology policy, such as on issues concerning the management of environmental and health risks. Involvement may be achieved in different ways: at the lowest level, the public may be targeted with enhanced information (e.g., about risks). At higher levels, public views may be actively solicited through such mechanisms as consultation exercises, focus groups, and questionnaires. At still higher levels, members of the public may be selected to take part in exercises that provide them with a degree of decision-making authority. The focus of this article is on public participation methods that aim to include the public in policy making at least

AUTHORS’ NOTE: We would like to thank the U.K. Department of Health for supporting the research reported here.
to the level of gathering their opinions; specifically, our interest is in the
evaluation of such methods. That evaluation is problematic is readily appar-
ent through the paucity of empirical examples in the academic literature. One
reason for this is a lack of appropriate benchmarks against which the quality
of participation exercises might be compared (Lowndes et al. 1998).
Although there have been a number of attempts at specifying criteria against
which effectiveness may be assessed (e.g., Fiorino 1990; Webler 1995), these
have certain limitations, and they have not yet been widely influential in a
practical sense. There is a need for a more comprehensive set of criteria for
determining whether a public participation mechanism is successful. This
article discusses a potential framework for evaluating methods and uses this
to assess a number of the most formalized of these.

The Case for Public Involvement in
Science and Technology Policy

Over the past few decades, there has been growing debate about the role
of the public in determining policy regarding issues of science and technology,
particularly within the context of health and environmental risk management.
There appears to be increasing realization in governmental, scientific, and
industrial bodies that they need to pay greater heed to the public, become
more accountable and responsive to it, and involve it in policy decision mak-
ing when feasible (e.g., Rosener 1978; Renn 1992; Vaughan 1993). Propo-
nents of greater public involvement have ranged from organizations such as
the U.S. Department of Energy and U.S. Environmental Protection Agency
(e.g., Bradbury 1994; Klauenberg and Vermulen 1994) to risk communica-
tion experts in academic institutions and government agencies (Chess, Salo-
mone, and Hance 1995) to the public itself (e.g., Feldman and Hanahan
1996). Indeed, recent legislation in a number of countries (such as the United
States and France) has made it necessary to gain public input and comment
prior to decision making in a number of risk arenas, such as in the siting of
radioactive waste facilities and the prioritization of environmental risk miti-
gation priorities (e.g., Charnigo 1989; Klauenberg and Vermulen 1994;
Barthes and Mays 1998). In any case, public input has arguably been incorpo-
rated into risk management in an informal manner over many years as public
opinion has often driven political and governmental choices about, for exam-
ple, allocating budgets in the face of scientific assessments (e.g., Klauenberg
and Vermulen 1994).
The reasons for the rise in interest in public participation in technical policy matters are bound to be diverse but may generally be considered to derive from either a recognition of basic human rights regarding democracy and procedural justice (e.g., Laird 1993; Perhac 1998) or simply from a practical recognition that implementing unpopular policies may result in widespread protest and reduced trust in governing bodies (e.g., Kasperson, Golding, and Tuler 1992).

The traditional view is that decisions regarding technical issues should be left in the hands of experts and scientists. Perhac (1996), for example, suggests that environmental policy based on the public’s conceptualization of risk (which has been shown to differ from that used by risk assessors; e.g., Renn 1992) fails to adequately protect fundamental human rights to health and liberty. Moffet (1996) warns that policies involving the public must balance the desire to foster legitimacy and support for decisions (e.g., about risk priority setting) with concerns to avoid priorities being driven by “the crisis of the day.” Others suggest that human inadequacies limit the public’s capacity to be effectively involved in complex decisions (the so-called “deficit model”), expressing doubts about whether the public understands significant concepts such as “uncertainty” and the nature of science as an incremental process (e.g., Brooks and Johnson 1991; but see, e.g., Frewer, Howard, and Shepherd 1998 for counterclaims), or point to deficiencies in the knowledge and reasoning abilities of laypersons (Slovic, Fischhoff, and Lichtenstein 1982; Earle and Cvetkovich 1995). Apart from ignorance, other factors may limit the potential for the public to contribute to complex policy decisions related to their attitudes, beliefs, and motivations (Ravetz 1986; McCallum and Santos 1997).

Counterarguments to these positions point out that there are frequently limitations in the knowledge of experts, who often disagree among themselves (e.g., Kraus, Malmfors, and Slovic 1992; Pollak 1996; Jasanoff 1997). Others have argued that the public is not necessarily irrational in its concerns about risks or in its rejection of experts’ claims, given abundant historical experience of episodes in which risk promoters have concealed or ignored relevant risk data or simply sought to advance their own interests by using such data selectively (e.g., Jasanoff 1993; Leiss 1995; Petts 1997). Perhaps the most persuasive argument for public involvement, however, is that value judgments are made at all stages of the risk management process, such as in deciding which risks to evaluate (Levidow 1994; Kunreuther and Slovic 1996; McCallum and Santos 1997). The implication is that the public is theoretically able to play a role in risk management at most, if not all, stages.
What Do We Mean by Public Involvement?
Communication versus Participation

The public may be involved in science and technology policy in a number of ways or at a number of levels (e.g., Wiedemann and Femers 1993; Smith, Nell, and Prystupa 1997). While the lowest level might involve communication between scientists or regulators and the public (e.g., about how risk estimates were arrived at), higher levels may seek some degree of public input, as in the solicitation of public opinion or the active participation of public representatives in the decision-making process itself. The lowest level involves top-down communication and a one-way flow of information, while the highest level is characterized by dialogue and two-way information exchange.

It is likely that the most appropriate method of public involvement will depend on the specifics of any particular situation and that more knowledge-based decisions (e.g., technical risk assessments) will require lower levels of involvement than more value-based decisions. Proponents of the deficit model, however, tend to view communication as generally the most apt approach, suggesting that public misunderstandings and objections might be overcome by better expressing the views of experts, with public acceptance of policies and decisions following as a result of an alignment process. Indeed, learning how to effectively communicate complex ideas to laypersons is important, and by developing the public understanding of science more generally, the public’s capacity to enter scientific debate about issues such as risk and risk mitigation will be enhanced (Frewer and Shepherd 1998).

The risk communication area has been the focus of much research. Studies have been conducted on aspects such as the best way to present information (e.g., Golding, Krimsky, and Plough 1992), the best medium for transmitting information to a “target” audience (e.g., Chipman et al. 1996), and the best people to impart such information (e.g., Frewer et al. 1996). Less research, however, has been conducted on mechanisms for involving the public at higher levels of input into decision making.

Public Participation Methods

According to Smith (1983), “public participation” encompasses a group of procedures designed to consult, involve, and inform the public to allow those affected by a decision to have an input into that decision. In this
analysis, “input” is the key phrase, differentiating participation methods from other communication strategies.

A consideration of the literature reveals the existence of a variety of methods and guidelines that might come under the public participation categorization, ranging from those that elicit input in the form of opinions (e.g., public opinion surveys and focus groups) to those that elicit judgments and decisions from which actual policy might be derived (e.g., consensus conferences and citizens’ juries). In the risk domain, it is apparent that most of these procedures have been used for gaining public input with regard to more value-laden and policy-oriented aspects of risk management, rather than for acquiring public input regarding the more technical aspects of risk assessment per se (e.g., Moffet 1996).

Of the large variety of approaches used to gain public input into science and technology policy, some are more formalized than others. In Table 1, the key features of eight of these approaches are described (given space limitations, we cannot detail these in the text). Other procedures are less well developed: some are simply putative procedures that have yet to make it “off the paper” (e.g., see Aronoff and Gunter 1994), others have been implemented by perhaps a single group of researchers or consultants in a number of practical projects (e.g., see Soby, Simpson, and Ives 1994; Swallow, Opaluch, and Weaver 1992; Wiedemann and Femers 1993), and still others comprise a variety of elements in wider-scale and invariably unique programs that may entail variations on more standard procedures (see Chen and Mathes 1989; Glicker 1992; Renn et al. 1993; Klauenberg and Vermulen 1994; Ballard and Kuhn 1996; Elder 1997; Petts 1997). It is possible that some of these innovative approaches, particularly those that combine a variety of methods or that provide variants on the more standardized procedures, will ultimately prove to be the most efficient mechanisms for engaging the public. For the purpose of this analysis, however, the absence of multiple uses of precise formulations of these “procedures” means that they tend to suffer from a lack of unbiased evaluation and an absence of any critical literature about their advantages and disadvantages, and hence the focus here is on the more formalized mechanisms.

Evaluating Public Participation Methods

In a practical sense, it is important for sponsors and authorities to appreciate the key characteristics of the various available public participation methods and to understand what these have to offer. In particular, the effectiveness
<table>
<thead>
<tr>
<th>Participation Method</th>
<th>Nature of Participants</th>
<th>Time Scale/Duration</th>
<th>Characteristics/Mechanism</th>
<th>Examples/References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referenda</td>
<td>Potentially all members of national or local population; realistically, a significant proportion of these.</td>
<td>Vote cast at single point in time.</td>
<td>Vote is usually choice of one of two options. All participants have equal influence. Final outcome is binding.</td>
<td>Biotechnology in Switzerland (Buchmann 1995); waste repository in Sweden (af Wåhlberg 1997).</td>
</tr>
<tr>
<td>Public hearings/inquiries</td>
<td>Interested citizens, limited in number by size of venue. True participants are experts and politicians making presentations.</td>
<td>May last many weeks/months, even years. Usually held during weekdays/working hours.</td>
<td>Entails presentations by agencies regarding plans in open forum. Public may voice opinions but have no direct impact on recommendation.</td>
<td>Frequent mechanism in, for example, United States (Fiorino 1990), Australia (Davison, Barnes, and Schibeci 1997); review by Middendorf and Busch (1997).</td>
</tr>
<tr>
<td>Public opinion surveys</td>
<td>Large sample (e.g., 100s or 1,000s), usually representative of the population segments of interest.</td>
<td>Single event, usually lasting no more than several minutes.</td>
<td>Often enacted through written questionnaire or telephone survey. May involve variety of questions. Used for information gathering.</td>
<td>Radioactive sites in United States (Feldman and Hanahan 1996); genetically modified food in the United Kingdom (Vidal 1998); biotech surveys (Davison, Barnes, and Schibeci 1997).</td>
</tr>
<tr>
<td>Negotiated rule making</td>
<td>Small number of representatives of stakeholder groups (may include public representatives).</td>
<td>Uncertain: strict deadline usually set: days/weeks/months.</td>
<td>Working committee of stakeholder representatives (and from sponsor). Consensus required on specific question (usually, a regulation).</td>
<td>Used by U.S. Environmental Protection Agency (Hanson 1984); method discussed by Susskind and McMahon (1985) and Fiorino (1990).</td>
</tr>
<tr>
<td>Method Type</td>
<td>Description</td>
<td>Duration</td>
<td>Interaction</td>
<td>Key Use Cases</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Consensus conference</td>
<td>Generally, ten to sixteen members of public selected by steering committee as “representative” of the general public. Preparatory demonstrations and lectures (etc.) to inform panelists about topic, then three-day conference. Lay panel with independent facilitator questions expert witnesses chosen by stakeholder panel. Meetings open to wider public. Conclusions on key questions made via report or press conference.</td>
<td>Used in Denmark and Netherlands on topics from food irradiation to air pollution (Joss and Durant 1994; Grundahl 1995); also used in United Kingdom on plant biotechnology (Ellahi 1995).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citizens’ jury/panel</td>
<td>Generally, twelve to twenty members of public selected by stakeholder panel to be roughly representative of the local population. Not precise but generally involve meetings over a few days (e.g., four to ten). Lay panel with independent facilitator questions expert witnesses chosen by stakeholder panel. Meetings not generally open. Conclusions on key questions made via report or press conference.</td>
<td>Examples in Germany, United States, and United Kingdom (e.g., Crosby, Kelly, and Schaeler 1986; Coote, Kendall, and Stewart 1994; Lenaghan, New, and Mitchell 1996).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Citizen/public advisory committee</td>
<td>Small group selected by sponsor to represent views of various groups or communities (may not comprise members of true public). Takes place over an extended period of time. Group convened by sponsor to examine some significant issue. Interaction with industry representatives.</td>
<td>Particularly evident in United States, for example, in cleanup of waste sites (Lynn and Busenberg 1995; Perhac 1998); see Creighton (1993) for guidelines.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focus groups</td>
<td>Small group of five to twelve selected to be representative of public; several groups may be used for one project (comprising members of subgroups). Single meeting, usually up to two hours. Free discussion on general topic with video/tape recording and little input/direction from facilitator. Used to assess opinions/attitudes.</td>
<td>Guidelines from Morgan (1993); U.K. example to assess food risk (Fife-Schaw and Rowe 1995).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of the various methods needs to be established (e.g., Rosener 1978). But what constitutes “effectiveness,” and how might we determine this, either theoretically or empirically? To answer these questions, it is important to understand what results of a participation exercise constitute “good” outcomes and what processes contribute toward these (and are thus desirable). Unfortunately, there is little comprehensive or systematic consideration of these matters in the academic literature, and hence whether any particular application of a particular method may be considered successful usually remains undetermined. Indeed, participation methods—such as referenda and public hearings—often seem to be employed simply in recognition of a need to involve the public in some way, assuming that involvement is an end in itself, rather than a means to an end (Wiedemann and Femers 1993). This may reflect the intentions of authorities using such techniques, where the appearance of involvement is sufficient, and little genuine interest exists in implementing any recommendations that might arise from the exercise.

A number of authors have made specific suggestions about the criteria that need to be satisfied for effective public participation in policy decisions or have discussed methods in terms of whether they incorporate certain elements or characteristics that are believed to be important— and hence to have implications for effectiveness (e.g., Crosby, Kelly, and Schaefer 1986; Fiorino 1990; Lynn and Busenberg 1995; Webler 1995; Smith, Nell, and Prystupa 1997). Most of the criteria discussed in the literature are procedural rather than substantive (Middendorf and Busch 1997), in that they relate to what makes for an effective process, rather than how to measure effective outcomes (e.g., quality of final decisions). Essentially, these various elements, mechanisms, and characteristics may be regarded as evaluation criteria because they provide benchmarks against which participation methods may be assessed. Otherwise, the evaluation of methods tends to be limited to ad hoc suggestions and criticisms about the advantages and disadvantages of the various techniques, and the lack of a clear framework for criticism makes it difficult to compare and contrast their relative merits. It is our aim to provide such a framework by specifying a number of evaluation criteria.

The results from research on “evaluation” will be amalgamated to produce a relatively tight set of criteria that can be used to assess the effectiveness of participation methods, and these will then be used to theoretically assess a variety of the most formalized methods. It should be noted, however, that many of the “results” used as the source for the criteria are no more than suggestions from academics and practitioners, rather than findings from empirical studies. The paucity of experimental results (e.g., from systematic comparisons of methods using validated methodologies to see which is the most
“effective”) reflects the difficulties in implementing controlled experimental studies in this domain. This arises as a consequence of the great number of variables that need manipulation and control—from design aspects of the procedures to contextual or environmental aspects of the situation in which the participation exercise takes place. Indeed, the contextual/environmental factors will interact with method type, such that there will be no one universally effective method (Smith, Nell, and Prystupa 1997). Difficulties also arise from the sheer variety of ways any one method may be implemented (partly a consequence of loose procedural definitions), which means that a particular method might prove either effective or ineffective, depending on how it is formulated and conducted. A further problem in evaluation comes from the lack of standardized measurement instruments (Crosby, Kelly, and Schaefer 1986). For example, the use of questionnaires on panelist attitudes might be useful in some cases, while for others there may be quantifiable outcomes. The development of such instruments is a practical, empirical issue.

Although basing evaluation criteria on the flimsy foundations of “opinion” lacks rigor, the condensing of considerable practical experience from researchers and practitioners does have some value. The definition of evaluation criteria here will provide a framework for discussion on the basis of which empirical studies might be designed, leading to either factual support or refutation and reformulation of the criteria.

**Evaluation Criteria**

Evaluation criteria may be divided into acceptance criteria, which are related to the effective construction and implementation of a procedure, and process criteria, which are related to the potential public acceptance of a procedure. Discussions in the literature have generally focused on one type or the other—for example, Fiorino (1990) assesses a number of procedures on the basis of “democratic criteria” (to which our acceptance criteria are conceptually similar)—but we believe that both types are required for method evaluation. If a procedure is effectively constituted but perceived by the public to be in some sense unfair or undemocratic, then the procedure may fail in alleviating public concerns. On the other hand, if a procedure and its recommendations are accepted by the public but the ultimate decision is attained in an ineffective manner, then its implementation could prove objectively damaging for sponsors and public. Furthermore, recognizing only democratic or acceptance criteria may lead to the advocacy of methods that are liable to lead to sponsor dissatisfaction. Since public participation exercises are normally
funded by a sponsor, sponsor dissatisfaction might lead to either a rejection of public participation in decision making or, if an exercise is enacted, a refusal to implement its output/decision.

One evaluation framework to which ours bears some parallels is that of Webler (1995), who discusses criteria of “fairness” and “competence” in citizen participation. However, Webler’s concerns largely lie with the attributes of the discourse within a participation exercise, whereas our concern is more general and not focused solely on exercises that involve interaction within a group. Webler’s framework is perhaps the most thoughtful and comprehensive consideration of the question of evaluation to date and is used by a number of authors to evaluate a variety of methods in a book edited by Renn, Webler, and Wiedemann (1995).

In the subsequent section, we subdivide the main criteria into others that address particular key aspects of public acceptance and good process in participation exercises. Although we suggest that it is important for participation methods to score well on all of the evaluation criteria described below, no claims at present are made about the relative importance of these.

Acceptance Criteria

Criterion of representativeness: The public participants should comprise a broadly representative sample of the population of the affected public. One concern frequently expressed in the literature is the need for participants to be representative of the broader public (or the affected subgroups within the population), rather than simply representing some self-selected subset (e.g., Nelkin and Pollak 1979; Crosby, Kelly, and Schaefer 1986; Kasper, Golding, and Tuler 1992; Webler 1995; Middendorf and Busch 1997). Particular caution should be exercised with regard to disenfranchising poorer groups or segments of society (Vaughan 1993) or employing an intelligent, motivated, self-interested, and unrepresentative elite that might intensify existing tendencies to place high-risk or undesirable projects in the communities of those least able to handle resultant disruption (Freudenberg and Olsen 1983). Further caution is needed in transboundary disputes, where one nation needs to make a decision about some risk issue with implications for another (such as in the siting of a nuclear power station or waste repository close to a national border or the construction of a dam along a major river to which other nations have access). For true representativeness to be achieved, members of all affected communities, including those in other nations, should be canvassed. There are likely to be political difficulties in implementing this level of representativeness, but the alternative is liable to be national
dispute. Representation also should take into account the relative distribution of views: in a small sample, the use of participants who represent each and every viewpoint may lead to a relative diminution of influence of those whose views are held by the majority (Rahl 1996).

One approach to achieving good representativeness is to select a random stratified sample of the affected population. Another might involve the use of questionnaires to determine the spread of attitudes with regard to a certain issue, using this as a basis for the proportionate selection of members (e.g., Crosby, Kelly, and Schaefer 1986). Once a panel is selected, the task environment should be sufficiently friendly to allow ease of attendance and not disadvantage some members so that they drop out of the process. Methodologically, representativeness is important if one genuinely wishes to gauge the opinions of the general public. Practically, the appearance of any bias in sampling may undermine the credibility of the exercise.

Although representativeness is an important criterion, practical constraints may limit its implementation. To fairly represent all stakeholders in the general public, a large sample is required, but groups cannot function efficiently with a large number of members. Therefore, some bias seems likely—it is just a question of how much. Financial limitations might also hinder attempts at gaining a representative sample.

Criterion of independence: The participation process should be conducted in an independent, unbiased way. Management of the participation process should be unbiased (e.g., Nelkin and Pollak 1979), such that managers and facilitators are not only independent in actuality but are seen to be independent. Likewise, public representatives should be independent of any affiliation to the sponsoring body.

Independence might be obtained and shown through the appointment of a steering committee or management team that incorporates members from diverse bodies or neutral organizations, such as university academics (e.g., Crowfoot and Wondolleck 1990). Disclosure from participants of any relationship to the sponsoring body might help confirm independence in the minds of the wider public. The use of a respected facilitator—such as a newscaster—might also prove beneficial (e.g., Ellahi 1995).

A possible objection to this criterion is that its implementation might drastically reduce the control and influence of the sponsoring organization. For example, it has been suggested that agency representatives should be included as collaborative participants throughout the negotiation process (e.g., Crowfoot and Wondolleck 1990; Aronoff and Gunter 1994). In a sense, however, a willingness to accept independent participants and facilitators
might serve to differentiate true efforts at gaining public input from those instances in which a sponsor simply seeks legitimization of a decision already made. The sponsor might allow clear independence of the process and still retain control by defining how the end conclusions of the process will be used (see the criterion of influence, below).

Criterion of early involvement: The public should be involved as early as possible in the process as soon as value judgments become salient. One aspect that is much discussed in the literature is the stage at which the public should become involved in policy matters. The consensus seems to be that public participation should occur as soon as is reasonably practical (e.g., Ng and Hamby 1997; Middendorf and Busch 1997). For example, it may not be sensible to have public participation in making decisions about highly technical issues, such as in the scientific assessment of risk. But at the stage when value judgments become important, it is necessary to consider psychological and sociological understandings of risk, and the public should be consulted (e.g., Swallow, Opaluch, and Weaver 1992; Chakraborty and Stratton 1993; Renn et al. 1993; Moffet 1996). An instance in which involvement might be seen as too late is when participation is used to choose among possible sites for a hazardous facility, with the public having been denied the opportunity to consider whether the facility is needed in the first place (e.g., Lake and Disch 1992). Public debate should thus be allowed on underlying assumptions and agenda setting and not just on narrow, predefined problems (Crosby, Kelly, and Schaefer 1986; Moffet 1996). This criterion is important if the credibility of the sponsors is to result from the process.

Caution needs to be exercised in the application of this criterion. Chakraborty and Stratton (1993) suggest that too much involvement of all standpoints (i.e., technical, economic, social, political, ethical, and public) might result in confusion over aims and judgments, hinder decision making, make clarification of issues impossible, and only produce defensive arguments of one standpoint against another. Thus, at each stage in the control of risks, there is an appropriate participation level that may not involve all standpoints equally.

Criterion of influence: The output of the procedure should have a genuine impact on policy. The output of the procedure should have a genuine impact on policy and be seen to do so (e.g., Crosby, Kelly, and Schaefer 1986; Fiorino 1990; Wiedemann and Femers 1993; Smith, Nell, and Prystupa 1997; Ng and Hamby 1997). One of the main complaints about participation methods is that they often have been perceived as ineffectual, simply being used to legitimate decisions or to give an appearance of consultation without
there being any intent of acting on recommendations. This results in public skepticism and distrust concerning the motives of sponsors.

One approach that might lead to fulfilling this criterion is to ensure that there is a clear acceptance beforehand as to how the output will be used and how it might direct policy (see also the criterion of task definition). Afterwards, use of the media to inform the general public about the specific ways in which the output has influenced policy would seem beneficial. Highlighting any areas where public suggestions have been adopted despite sponsor resistance might enhance credibility further. One caveat related to this criterion is that sponsors should probably be wary of accepting binding votes and giving away all of their power to public participants in case this results in the compulsory implementation of a decision based on emotion or prejudice (e.g., Rahl 1996).

Criterion of transparency: The process should be transparent so that the public can see what is going on and how decisions are being made. It is generally accepted that the participation process should be transparent (e.g., Frewer 1999), so that the wider public can see what is going on and how decisions are being made (i.e., it should not be held behind closed doors). By being transparent, it is likely that public suspicions about the sponsors and their motives may be allayed. Transparency might involve releasing information on aspects of the procedure, varying from the manner of the selection of the public participants to the way in which a decision is reached to the minutes of meetings. If any information needs to be withheld from the public, for reasons of sensitivity or security, it would seem important to admit the nature of what is being withheld and why, rather than risking the discovery of such secrecy, with subsequent adverse reactions.

Process Criteria

Criterion of resource accessibility: Public participants should have access to the appropriate resources to enable them to successfully fulfill their brief. It is clear that effective decision making requires access to appropriate and relevant information, but this is only one aspect of the resources that public participants should have to enable them to successfully fulfill their brief in a participation exercise (e.g., Ng and Hamby 1997). Necessary resources include (1) information resources (summaries of the pertinent facts), (2) human resources (e.g., access to scientists, witnesses, decision analysts), (3) material resources (e.g., overhead projectors/whiteboards), and (4) time resources (participants should have sufficient time to make decisions).
Restrictions on any of these resources are liable to have an impact on the quality of the participation process.

It is essential to ensure that there is real commitment to the process by the sponsor and that appropriate finances are available. The precise problem will dictate what resources are required: in some cases, providing information on fundamental aspects of science (e.g., the way science and knowledge advances, the role of uncertainty, how scientific paradigms change) might be appropriate; in others, practical demonstrations of scientific features might help participant understanding (Moffet 1996).

Possible difficulties revolve around the issue of cost (see the criterion of cost-effectiveness). It will always be possible to summon more witnesses or experts to take part in a procedure or to supply more time and materials. A trade-off is clearly required. This is also true with respect to the amount of information presented to participants: information overload is possible, which is liable to lead to stress and confusion. To avoid this, concise summaries of information, free of jargon, would seem apt.

Criterion of task definition: The nature and scope of the participation task should be clearly defined. It is important to ensure that there is as little confusion and dispute as possible regarding the scope of a participation exercise, its expected output, and the mechanisms of the procedure. All of these aspects should be clearly defined at the outset (e.g., Chakraborty and Stratton 1993). The effectiveness of a procedure, as well as its credibility, is liable to be influenced by any dispute caused through misunderstandings.

The main objection to this criterion is that an overly prescriptive set of definitions and rules might reduce flexibility in the face of new information or disputes. This might be overcome if the terms under which an exercise took place allowed for changes in terms of reference in the face of important new information.

Criterion of structured decision making: The participation exercise should use/provide appropriate mechanisms for structuring and displaying the decision-making process. The participation exercise should provide participants with appropriate mechanisms for structuring and displaying the decision-making process (e.g., Crosby, Kelly, and Schaefer 1986; Swallow, Opaluch, and Weaver 1992; Renn et al. 1993). This would enable the underlying reasons behind a decision to be examined, as well as the extent to which a conclusion was well supported, and would help organize the process. Documenting the process of reaching a decision (as well as the outcome) is liable to increase transparency (and hence the perceived credibility of the exercise) as well as the efficiency of the process (see criterion of transparency).
A variety of decision-aiding tools might be incorporated into a participation procedure, such as decision analysis, decision trees, multiattribute utility theory, and the Delphi technique. These are basically procedures that help to structure the decision-making process of individuals and groups. They may be used to clarify the relationship between important variables, ensure that weight is given to social and political concerns in addition to technical assessments, combine judgments in an effective manner, and give a clear and explicit representation of how a decision is arrived at (the individual details of these procedures are too complex and lengthy to discuss here). It would seem particularly important to structure the decision process in groups (which are commonly used in more recent participation procedures, such as citizens’ juries and consensus conferences), as groups are often found to function suboptimally and inefficiently due to the influence of dominant/dogmatic individuals, the premature closure of arguments, social loafing (by individuals who do not contribute to the group discussion), and so on (Rowe 1998). It might be that an independent decision analyst could be usefully involved (see also the criterion of resource availability) or that a group facilitator might employ rules for effective group decision making to keep group discussions on track.

Criterion of cost-effectiveness: The procedure should in some sense be cost-effective. Clearly, cost is a key concern to those involved in organizing a participation exercise, and value for money is a significant motivation (e.g., Crosby, Kelly, and Schaefer 1986; Rahl 1996). For example, a major public hearing might be inappropriate for a relatively minor policy decision. Prior to conducting a participation exercise, it is clearly sensible to take account of the potential costs of the alternative methods, in both time and money, and to consider the extent to which they fulfill the other criteria. Although monetary costs are objectively measurable, most discussions on participation methods in the literature do not discuss costs in any depth. Furthermore, given the sheer variety of ways in which any one method might be implemented, it becomes difficult to establish anything but a very vague order of probable “costliness” of procedures.

Results of Evaluation

Table 2 shows our evaluation of the public participation methods described in Table 1. As a consequence of the variety of ways in which any one method is applied (e.g., Lynn and Busenberg 1995 describe the varied applications of citizen advisory committees [CACs]), plus the mediating
effects of social and environmental factors and the lack of precision in defining and measuring the criteria, it is not possible to definitively state whether a particular method succeeds or fails, and hence relative terms are used in the table. The assessments are largely based on our own opinions, and others might disagree with them. We welcome the process of dialogue and debate to help clarify this evaluation framework.

As can be seen from Table 2, each method has its strengths and weaknesses. The public hearing, however—which is perhaps the most widespread method for engaging the public—scores relatively low on both acceptance and process criteria. In the past, public hearings have been perceived as being quick, cheap, and simply administered means of satisfying any legal requirement for public participation (Smith 1983) and seen as giving the appearance of community involvement (Fiorino 1990). Their disadvantages, however, seem numerous. For example, they are commonly held during weekday working hours in locations that are “formidable” to the public (e.g., government buildings), which may disadvantage low-income and minority citizens and have a negative impact on the representativeness of those attending (e.g., Checkoway 1981). Communication at hearings is primarily one-way—consisting of presentations and testimony—involving little debate with the various stakeholders (scoring low on the influence criterion). Indeed, public hearings often seem designed to contain and control participation (Middendorf and Busch 1997) by allowing only limited choices on narrow, short-term questions at a late stage of the policy process (e.g., Folk 1991), and so they also score low on the early involvement criterion. It has been suggested that their main aim is often to co-opt public support and to change decisions rather than to seek informed consent and expand democratic choice (e.g., Nelkin and Pollak 1979). Some empirical evidence suggests that they have little influence on citizen behavior or policy choices (Cole and Caputo 1984).

Referenda, public opinion surveys, and focus groups do reasonably well on acceptance criteria but not on process criteria. From this, we suggest that although these methods might gain a fair amount of credibility with the public, the quality of the decisions that arise from their implementation may not be high—and this would clearly be of concern to a sponsor. Thus, their participants are generally representative of the population and independent of the sponsors, and their results and processes are generally simple and transparent. In the case of referenda, a binding commitment by the government to implement their results means that “influence” is high (although conceivably too high in that they delegate authority to an “uninformed/unqualified electorate”; Fiorino 1990). Because opinion surveys and focus groups may serve as the basis for subsequent policy formation, they may be implemented at a
Table 2. An Assessment of the Most Formalized Public Participation Techniques According to a Variety of Evaluation Criteria

<table>
<thead>
<tr>
<th>Acceptance criteria</th>
<th>Referenda</th>
<th>Public Hearings</th>
<th>Public Opinion Survey</th>
<th>Negotiated Rule Making</th>
<th>Consensus Conference</th>
<th>Citizens’ Advisory Committee</th>
<th>Citizen Jury/Panel</th>
<th>Focus Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representativeness of participants</td>
<td>High (assuming full turnout at poll)</td>
<td>Low (generally high)</td>
<td>Low</td>
<td>Moderate (limited by small sample)</td>
<td>Moderate to low</td>
<td>Moderate (limited by small sample)</td>
<td>Moderate (limited by small sample)</td>
<td>Moderate (limited by small sample)</td>
</tr>
<tr>
<td>Independence of true participants</td>
<td>High</td>
<td>Generally low</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>Moderate (often relation to sponsor)</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Early involvement?</td>
<td>Variable</td>
<td>Variable</td>
<td>Potentially high</td>
<td>Variable</td>
<td>Potentially high</td>
<td>Variable but may be high</td>
<td>Potentially high</td>
<td>Potentially high</td>
</tr>
<tr>
<td>Influence on final policy</td>
<td>High</td>
<td>Moderate</td>
<td>Indirect and difficult to determine</td>
<td>High</td>
<td>Variable but not guaranteed</td>
<td>Variable but not guaranteed</td>
<td>Variable but not guaranteed</td>
<td>Liable to be indirect</td>
</tr>
<tr>
<td>Transparency of process to the public</td>
<td>High</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Low</td>
<td>High</td>
<td>Moderate</td>
<td>Variable but often low</td>
<td>Low</td>
</tr>
</tbody>
</table>

(continued)
### Table 2 Continued

<table>
<thead>
<tr>
<th>Process criteria</th>
<th>Referenda</th>
<th>Public Hearings</th>
<th>Public Opinion Surveys</th>
<th>Negotiated Rule Making</th>
<th>Consensus Conference</th>
<th>Citizens’ Advisory Committee</th>
<th>Citizen Jury/Panel</th>
<th>Focus Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource accessibility</td>
<td>Low</td>
<td>Low-</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Variable</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task definition</td>
<td>High</td>
<td>Generally high</td>
<td>Low</td>
<td>High</td>
<td>Generally high</td>
<td>Generally high</td>
<td>Variable but may be high</td>
<td>Variable but may be high</td>
</tr>
<tr>
<td>Structured decision making</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate (influence of facilitator)</td>
<td>Potentially high</td>
<td>Variable (influence of facilitator)</td>
<td>Low</td>
</tr>
<tr>
<td>Cost-effectiveness</td>
<td>Variable/low</td>
<td>Low</td>
<td>Potentially high</td>
<td>Potentially high</td>
<td>Moderate to high</td>
<td>Moderate to high</td>
<td>Variable</td>
<td>Potentially high</td>
</tr>
</tbody>
</table>
fairly early stage of any decision-making process and hence score high on the
criterion of early involvement. These approaches take little citizen time and
fewer resources than many other procedures and are ranked high on the crite-
rion of cost-effectiveness.

On the negative side, participants in referenda, public opinion surveys,
and focus groups have no structured access to resources to enable them to
make good decisions, and as such their output may reflect biases and misun-
derstandings that have no opportunity for resolution (i.e., they score low on
the criteria of resource accessibility and structured decision making). None
of these methods allows a dialogue between risk regulators and the public
(e.g., Middendorf and Busch 1997), and they may even be said to displace
active forms of public debate (e.g., Davison, Barnes, and Schibeci 1997). In
the case of referenda, the output simply reflects the direction and not intensity
of beliefs, with no documentation of the basis for consensus or rationale
underlying decisions (Fiorino 1990), and they reflect opinions at just one
point in time (e.g., Kathlene and Martin 1991). Although the output from sur-
veys and focus groups tends to be more explicit, this generally fails to focus
on a particular question and may have only minimal impact on ultimate pol-
icy (hence the low scores on the criterion of influence). We have also scored
focus groups low on the criterion of transparency since they tend to be con-
ducted behind closed doors, although it is possible that the procedure might
be extended so that group sessions are followed by press releases or even
question-and-answer sessions, which would increase the transparency of the
method.

Generally, referenda would seem inappropriate for most complex, multi-
faceted decisions—particularly risk assessments—while the involvement of
all citizens in all decisions in modern society is clearly impractical. Refer-
enda may also be expensive to implement, which, given the nature of their
output, make them of questionable cost-effectiveness in most situations. For
public opinion surveys and focus groups, their advantages lie in clarifying
bases of agreement and disagreement and identifying values that underlie
opinions, rather than setting a clear direction for policy makers (e.g., Fiorino
1990). As such, they might best be regarded as exploratory methods for com-
plementing other procedures (indeed, Crosby, Kelly, and Schaefer 1986 used
a telephone survey to identify public views for the purpose of selecting repre-
sentative members for a citizens’ panel procedure). For such a purpose, low
“transparency” would seem less of a problem, while the lack of any clear link
between the participation process and the final decision (i.e., a low score on
the criterion of influence) would actually seem preferable.

The negotiated rule-making approach makes a more concerted attempt to
structure the decision making of the participants, with more emphasis on
ensuring that the resources are available to make a good decision (criterion of resource availability) and that the task is precisely defined (criterion of task definition). Since the process may be fairly focused and require relatively few participants, the approach may be relatively cost-effective. The approach scores highly on the criterion of influence, in that recommendations generally influence policy, but scores more highly than approaches such as the referendum in that the sponsor is represented in the committee, and their approval is needed for the required consensus to be reached (i.e., the sponsor does not concede all power to the participants).

The main problem with the approach is that it scores poorly on most of the acceptance criteria (see Table 2), and indeed it might be argued that it barely qualifies as a public participation method at all. It scores poorly on the criterion of representativeness, as only leading members of the appropriate communities are represented, and there is no attempt to include “ordinary” citizens in the participation process (Laird 1993). Furthermore, such committees tend to meet behind closed doors, scoring low on the transparency criterion (although hearings could be held in public). The technique might only work for a limited set of clearly definable technical policy issues and may not work when the issues involve intrinsic and far-reaching value conflicts among contending parties.

Approaches such as the consensus conference, citizens’ jury/panel, and CAC score reasonably well on both acceptance and process criteria, such as the criteria of early involvement and task definition (largely as a consequence of extended group interactions providing opportunities for defining the problems that need to be debated). Unlike the previous approaches, however, extensive efforts are made in these to provide public participants with the appropriate resources to make good decisions, and hence they score well on the resource accessibility criterion (although severe time constraints might be seen as a resource deficiency). Given these advantages (in addition to the time constraints imposed and the limited number of participants used), these approaches might be seen as relatively cost-efficient when compared to timely and expensive approaches such as public hearings and referenda (although our assessment of the cost-effectiveness of procedures is perhaps the least emphatic of our evaluations and the most open to debate, given the variety of ways of enacting any one participation method).

We have rated the citizens’ jury no more highly than the consensus conference on the cost-effectiveness criterion because, although probably cheaper to implement, the loss of transparency in the process from being held behind closed doors might reduce its effectiveness overall. The transparency of these approaches in general is also at risk from a single, competing high-interest news event reducing media exposure of their conclusions. This, however, is
an environmental variable not related to the approach per se and is liable to affect the effectiveness of most public participation methods. Similarly, environmental variables—such as national political culture—are liable to determine how well any one exercise scores on the criterion of influence (e.g., Davison, Barnes, and Schibeci 1997). For example, in Denmark there is a link between the organizers of consensus conferences and policy-making bodies (the Danish Board of Technology and Danish parliament, respectively), which has resulted in the clear implementation of past decisions, although this link is not evident in other cultures (e.g., in the United Kingdom). Lynn and Busenberg (1995) have also noted how the impact of CACs on policy outcomes can be seen to vary according to the intentions and expectations of the institutions being advised.

Despite generally good ratings, these approaches nevertheless leave room for improvement in a number of areas. Although each attempts (to a degree) to gain representative public samples, their scores on this criterion are only moderate, largely as a consequence of the small sample sizes they employ (necessitated by their group-based modus operandi). It has also been suggested by a number of authors that there may be various imbalances in the selection of panelists, in that volunteers (e.g., for consensus conferences) may be self-selected to a degree (e.g., Middendorf and Busch 1997) or influenced by the intentions of the sponsors (particularly regarding CACs [e.g., Lynn and Busenberg 1995], hence the lower rating on the criterion of representativeness for this method), or that committee members may become elitist and lose touch with their constituencies (e.g., in CACs; Creighton 1993). The significant impact of the sponsors in CACs has led us to give this method a relatively low rating on the criterion of independence.

The group-based mechanism underlying these approaches is also a potential source of difficulty, in that group behavior has often been shown to be suboptimal as a consequence of a number of psychological and social factors (e.g., Lenaghan, New, and Mitchell 1996), as when vociferous individuals monopolize discussions. As such, the quality of any decision reached might be a result of group dynamics and social influence, more so than the public participation approach itself. We have rated the three techniques more highly than other approaches on the criterion of structured decision making, however, because the influence of a facilitator and the definition of rules and guidelines usually provided might help to overcome some of these difficulties and provide a degree of support to the decision-making process, although the use of other procedures to structure group interaction might ultimately improve the performance of these techniques and their score on this criterion (e.g., Rowe 1998). Similarly, techniques from decision theory may be helpful in coping with limitations in the knowledge and decision-making skills of
laypeople, and these are key components in a number of recent participation methods (e.g., Soby, Simpson, and Ives 1994), including a version of the citizens’ jury (e.g., Renn et al. 1993).

Providing support for both decision making and group behavior is an important component of a wider concern for creating an appropriate environment (with all appropriate resources) to enable lay members of the public to contribute effectively to complex and important policy issues. Naturally, the more complex the intervention is, the more expensive it is likely to be for a sponsor. However, the potential damage that might be caused to public trust—and, indeed, to public health and welfare from a poorly made decision—must also be considered by the sponsor.

Discussion

There is increasing contention that public participation in policy making in science and technology is necessary to reflect and acknowledge democratic ideals and enhance trust in regulators and transparency in regulatory systems. While a number of different participation techniques have been developed to allow this, their relative usefulness is difficult to ascertain because systematic comparisons between them are rare. The main problem in the evaluation of participation methods is the absence of any optimal benchmark against which they might be compared and measured, which arises in part because of confusion as to what we mean by “effectiveness.” In this article, we have stipulated a number of evaluation criteria as benchmarks. Using this proposed framework (or normative model), we have conducted a preliminary assessment of some of the most formalized participation methods, although we have used only our opinion as the meter for assessment. As such, our conclusions should not be overinterpreted: more reliable and valid measurement tools are required, and we are currently engaged in developing these. Likewise, the framework we present should be seen as a device for helping to think about the issue of public participation method effectiveness, and the evaluation criteria should not be taken as definitive but rather as the focus for debate and a spur to future experimental research.

From our analysis, it is difficult at this juncture to categorically declare that any one method is the best. Indeed, Smith, Nell, and Prystupa (1997) conclude—and in this we agree—that the most appropriate techniques for public participation are likely to be hybrids of more traditional methods. Similarly, Fiorino (1990) suggests that a potentially effective approach to participation may be to complement one mechanism with another—such as using a survey to clarify the bases of disagreement on issues prior to a series
of public hearings or using a series of citizens’ panels to add balance and depth to what policy makers might learn in open hearings. An effective technique also is liable to access one or more of the decision aids that already exist, and there is no reason why these may not be used to enrich standard participation methods—as done by Renn et al. (1993) in the case of the citizens’ panel.

The intrinsic features of any one participation method will not act alone, however, in determining whether that method will be effective. Rather, a variety of contextual and environmental factors will interact with the characteristics of a method to determine effectiveness. Thus, while one method may be appropriate in a certain situation, another method may be more apt under different circumstances. In this sense, Table 2 is limited, for it does not identify the contextual/environmental factors that will contingently affect effectiveness but instead represents a broad analysis in which “fuzzy” evaluations are made and caveats are expressed. This has been somewhat inevitable, given the gaps that exist in our knowledge. Some of the factors that might affect method effectiveness have been discussed elsewhere, such as national political styles, expectations about the role of government, and local mechanisms for participation (e.g., Nelkin and Pollak 1979). Other authors have attempted to specify situations in which a specific technique might be useful or otherwise (e.g., for consensus conferences, see Grundahl 1995), implying the influence of more mediating factors, such as whether an issue contains conflict, depends on expert contribution for clarification, or involves a particular risk domain. It is likely that an in-depth review of environmental factors (similar to this review) would provide a useful framework for directing future research. As with our evaluation criteria, suitable measurement instruments will need to be developed to measure the identified factors if they are to be useful in directing the choice of one participation method over another.

References

26 Science, Technology, & Human Values


Gene Rowe graduated from the University of Bristol in psychology and gained his Ph.D. from the Bristol Business School at the University of the West of England (UWE). He is currently a senior researcher at the Institute of Food Research, Norwich, United Kingdom. His research interests include judgment and decision making, expert systems, forecasting, the Delphi technique, and the influence of media on public risk perceptions. His main interest at present lies in public participation methods and their evaluation.

Lynn J. Frewer graduated from the University of Bristol in psychology (B.Sc.) and followed this with an M.Sc. in ergonomics from University College London and a Ph.D. in applied psychology from the University of Leeds. She is currently acting head of the consumer science group at the Institute of Food Research. Her current interests include the psychology of risk perceptions and attitudes, the influence of the media on risk perception, public reactions to genetic engineering, the impact of trust on the effectiveness of risk communication, public understanding of risk uncertainty, and developing methodologies for fostering public participation in the strategic development of food technologies.