

Evaluation

<http://evi.sagepub.com>

A Vision of Evaluation that Strengthens Democracy

Michael Quinn Patton

Evaluation 2002; 8; 125

DOI: 10.1177/1358902002008001740

The online version of this article can be found at:

<http://evi.sagepub.com>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:



[The Tavistock Institute](#)

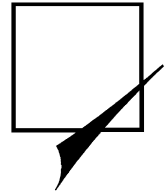
Additional services and information for *Evaluation* can be found at:

Email Alerts: <http://evi.sagepub.com/cgi/alerts>

Subscriptions: <http://evi.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>



A Vision of Evaluation that Strengthens Democracy

MICHAEL QUINN PATTON

The Union Institute, USA

Democratic Evaluation

Over the relatively short history of professional evaluation, those working in the field have directed considerable attention to both a vision of democratic approaches to evaluation and practice wisdom about how to realize that vision. In Europe, the democratic evaluation model of Barry MacDonald (1987) stands out. He argued that 'the democratic evaluator' recognizes and supports value pluralism with the consequence that the evaluator should seek to represent the full range of interests in the course of designing an evaluation. In that way an evaluator can support an informed citizenry, the *sine qua non* of strong democracy, by acting as an information broker between groups who want and need knowledge about each other. The democratic evaluator must make the methods and techniques of evaluation accessible to non-specialists, that is, the general citizenry. MacDonald's democratic evaluator seeks to survey a range of interests by assuring confidentiality to sources, engaging in negotiation between interest groups, and making evaluation findings widely accessible. The guiding ethic is the public's right to know.

Saville Kushner (2000) has carried forward, deepened and updated MacDonald's democratic evaluation model. He sees evaluation as a form of personal expression and political action with a special obligation to be critical of those in power. He places the experiences of people in programs at the center of evaluation. The experiences and perceptions of the people, the supposed beneficiaries, are where, for Kushner, we will find the intersection of Politics (big P – Policy) and politics (small p – people). Much of evaluation these days (i.e. logic models, theories of action, outcomes evaluation) is driven by the need and desire to simplify and bring order to chaos. Kushner, in contrast, embraces chaos and complexity because democracy is complex and chaotic. He challenges the facile perspectives and bureaucratic imperatives that dominate much of current institutionally based evaluation practice. Over and over he returns to the people, to the children and teachers and parents, and the realities of their lives in program settings as they experience those realities. He elevates their judgments over professional and external judgments. He feels a special obligation to focus on,

Evaluation 8(1)

capture, report and therefore honor the views of marginalized peoples. He calls this 'personalizing evaluation', but the larger agenda is strengthening democracy. Consider these reflections on the need for evaluators and evaluations to address questions of social justice and the democratic contract:

Where each social and educational program can be seen as a reaffirmation of the broad social contract (that is, a re-confirmation of the bases of power, authority, social structure, etc.), each program evaluation is an opportunity to review its assumptions and consequences. This is commonly what we do at some level or another. *All programs expose democracy and its failings; each program evaluation is an assessment of the effectiveness of democracy in tackling issues in the distribution of wealth and power and social goods.* Within the terms of the evaluation agreement, taking this level of analysis into some account, that is, renewing part of the social contract, is to act more authentically; to set aside the opportunity is to act more inauthentically, that is, to accept the fictions. (Kushner, 2000: 32-3; emphasis added)

Evaluation as a Democratic Process

On the American side, House and Howe (2000) have been most explicit about linking evaluations to larger sociopolitical and moral structures. They have articulated three requirements for evaluation done in a way that supports democracy: *inclusion*, *dialogue* and *deliberation*. They worry about the power that derives from access to evaluation and the implications for society if only the powerful have such access.

We believe that the background conditions for evaluation should be explicitly democratic so that evaluation is tied to larger society by democratic principles argued, debated, and accepted by the evaluation community. Evaluation is too important to society to be purchased by the highest bidder or appropriated by the most powerful interest. Evaluators should be self-conscious and deliberate about such matters . . .

If we look beyond the conduct of individual studies by individual evaluators, we can see the outlines of evaluation as an influential societal institution, one that can be vital to the realization of democratic societies. Amid the claims and counterclaims of the mass media, amid public relations and advertising, amid the legions of those in our society who represent particular interests for pay, evaluation can be an institution that stands apart, reliable in the accuracy and integrity of its claims. But it needs a set of explicit democratic principles to guide its practices and test its intuitions. (House and Howe, 2000)

While MacDonald, Kushner and House and Howe make explicit linkages between evaluation and democracy, a number of other evaluation approaches imply such linkages by emphasizing various degrees and types of stakeholder participation and involvement and, correspondingly, evaluator responsiveness. For reviews of the variety of such approaches and distinctions among them see Cousins and Earl (1995), Alkin (1997), and Ryan and DeStefano (2000). The work of Mertens (1998, 1999) on 'inclusive evaluation' and the 'empowerment evaluation' model of Fetterman et al. (1996) offer additional examples of evaluation approaches that support democratic principles, social justice and explicitly political foundations of evaluation in support of those whose stakes tend to be

under-represented in policy discussions because they are marginalized economically, socially and politically.

Taken together, these writings on evaluation's role in supporting democratic processes reflect a significant shift in the nature of evaluation's real and potential contribution to strengthening democracy. A decade ago the emphasis was all on increasing use of findings for enhanced decision making and program improvement and, therefore, making sure that findings reflected the diverse perspectives of multiple stakeholders, including the less powerful and participants in programs (instead of just staff, administrators and funders). While this thrust remains important, a parallel and reinforcing use of evaluation focuses on helping people learn to think and reason evaluatively, and how rendering such help can contribute to strengthening democracy over the long term. I turn now to elaborate that contribution.

Supporting Democracy Through Process Use: Helping the Citizenry Weigh Evidence and Think Evaluatively

Start with the premise that a healthy and strong democracy depends on an informed citizenry. Evaluation's contribution, then, is to help ensure an informed electorate. This vision was articulated by House and Howe in the passage cited earlier, worth repeating because it so powerfully makes the case for this role.

Amid the claims and counterclaims of the mass media, amid public relations and advertising, amid the legions of those in our society who represent particular interests for pay, evaluation can be an institution that stands apart, reliable in the accuracy and integrity of its claims. (House and Howe, 2000)

In addition, however, evaluation has a role to play in helping the citizenry weigh evidence and *think evaluatively*. This involves thinking processes that must be learned. It is not enough to have trustworthy and accurate information (the informed part of the informed citizenry). People must also know how to use information, that is, to weigh evidence, consider inevitable contradictions and inconsistencies, articulate values, interpret findings and examine assumptions, to note but a few of the things meant by 'thinking evaluatively'.

Philosopher Hannah Arendt was especially attuned to this foundation of democracy. Having experienced totalitarianism, then having fled it, she devoted much of her life to studying it and its opposite, democracy. She believed that thinking thoughtfully in public deliberations and acting democratically were intertwined. Totalitarianism is built on and sustained by deceit and thought control. In order to resist efforts by the powerful to deceive and control thinking, Arendt believed that people needed to practice thinking. Toward that end she developed 'eight exercises in political thought' (Arendt, 1968). She wrote that 'experience in thinking . . . can be won, like all experience in doing something, only through practice, through exercises' (p. 4). From this point of view, might we consider every evaluation an opportunity for those involved to practice thinking? This would mean that every evaluation is an opportunity to strengthen democracy by teaching people how to think evaluatively. In this regard we might

Evaluation 8(1)

aspire to have evaluation do what Arendt hoped her exercises in political thought would do, namely, 'to gain experience in *how* to think'. Her exercises 'do not contain prescriptions on what to think or which truths to hold', but rather on the act and process of thinking. For example, she thought it important to help people think conceptually, to:

discover the real origins of original concepts in order to distil from them anew their original spirit which has so sadly evaporated from the very keywords of political language – such as freedom and justice, authority and reason, responsibility and virtue, power and glory – leaving behind empty shells . . . (Arendt, 1968: 14–15)

Might we add to her conceptual agenda for examination and public dialogue such terms as outcomes and performance indicators, interpretation and judgment, and beneficiary and stakeholder, among many evaluative possibilities?

Helping people learn to think evaluatively by participating in real evaluation exercises is what I've come to call 'process use' (Patton, 1997, 1998). I have defined process use as relating to and being indicated by individual changes in thinking and behaving that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process. (Changes in program or organizational procedures and culture may also be manifestations of process impacts, but that is not our focus here.) This means an evaluation can have a dual impact in terms of strengthening democracy:

1. a more informed electorate through use of findings; and
2. a more thoughtful and deliberative citizenry through helping people learn to think and engage each other evaluatively.

One way of thinking about process use is to recognize that evaluation constitutes a culture, of sorts. When we engage other people in the evaluation process, we are providing them with a cross-cultural experience. This culture of evaluation, that we as evaluators take for granted in our own way of thinking, is quite alien to many of the people with whom we work at program levels. Examples of the values of evaluation include: clarity, specificity and focusing; being systematic and making assumptions explicit; operationalizing program concepts, ideas and goals; distinguishing inputs and processes from outcomes; valuing empirical evidence; and separating statements of fact from interpretations and judgments. These ways of thinking are far from natural for many people; indeed, they can seem quite alien. When people are involved in a process of evaluation, at least in any kind of stakeholder involvement or participatory process, they are in fact learning things about evaluation culture and often learning how to think in these ways.

Helping people learn to think evaluatively can be a more enduring impact of an evaluation than the use of specific findings generated by that same evaluation. Findings have a very short 'half life' – to use a physical science metaphor. The relevance of findings can deteriorate very quickly as the world changes rapidly. Specific findings typically have a small window of applicability. In contrast, learning to think and act evaluatively can have an ongoing impact. The experience of being involved in an evaluation, then, for those stakeholders actually involved,

can have a lasting impact on how they think, on their openness to reality testing, on how they view the things they do, and on their capacity to engage in democratic processes.

Training People to Think Evaluatively

Beyond what people learn about thinking from participating in evaluations, recognizing the value to democracy of people thinking evaluatively opens up new training opportunities for our profession. Most training discussions are focused on the need to train evaluators, that is, on the supply side of our profession. But we also need to be training evaluation users: to build up the demand side as well as to broaden the general public capacity to think evaluatively. Let me share an example of such an effort.

A major initiative underway in my home state of Minnesota involves the ambitious goal of infusing evaluative thinking into the government, not-for-profit, and philanthropic sectors. Developed by a collaboration of leaders from these sectors and with support from a major philanthropic foundation, we are conducting Outcomes Learning Labs for non-profit executives, government managers, and funders. The Outcomes Learning Labs involve a six-month hands-on commitment from participants.

After an initial one-day workshop on outcomes, evaluation, organizational learning, and systems change, small lab groups are formed to meet for 10 sessions, three hours per session, over a six-month period. Each lab group includes a mix of people from non-profit, government and philanthropic sectors. All lab participants must undertake an outcomes project of some kind in their own organization. These change projects are the focus of lab interactions, analyses and discussions. All lab groups are facilitated by an experienced professional. Participants are taught to explore issues in depth through questioning deeply, being sure of context and undertaking systems analyses. In a given six-month period, about 100 participants engage in the lab process in 10 small group labs. At the end of the six months we have a closing workshop. Some groups continue to meet on their own. Advanced groups combining participants from previous labs have also formed.

The project has now operated for three years through six sequences, so roughly 600 executives, managers, senior staff and funders have gone through the Outcomes Learning Labs. The labs are specifically designed to de-mystify and deepen thinking about outcomes. Cautions about outcomes measurement and performance monitoring are part of the workshops and labs. In addition, the labs get influential people actually thinking about and doing outcomes work. Our strategy has been to put outcomes evaluation on the community agenda at a deeper level than mere political rhetoric by engaging leaders in real projects and helping them feel comfortable with and knowledgeable about outcomes management as a leadership function. A heavy emphasis is placed on involving program participants in meaningful ways in their organizational evaluative processes. The scope of the project is aimed at creating a critical mass of leaders who can thoughtfully incorporate evaluative thinking into their deliberations

both internally and externally. The process is now branching out through a train-the-trainers initiative that will expand further the influence of the initiative.

We are also working with grassroots community-development efforts to build training in evaluative thinking into their efforts.

These initiatives illustrate the potential impact of thinking about the role of evaluation in strengthening democracy and thinking creatively about how to create new opportunities for people to not only participate in evaluations but to get training in how to think evaluatively as a way of deepening public dialogues and deliberations. For example, imagine that as part of the annual evaluation conferences of the major evaluation associations (e.g. European, Canadian, American, Australasian, African), a public session was sponsored in which the community was invited to learn about current issues in the field and, at the same time, discover that there actually is a vibrant and growing evaluation profession. This is but one example of how we, as a profession, might become more expansive and less insular within a vision of evaluation strengthening democracy.

In essence, I'm suggesting in this section that when we think about the conference theme 'Taking Evaluation to the People', we include not only taking evaluation findings but also offering the people opportunities to learn how to think and dialogue about evaluation.

Special Issues in the Linkages Between Evaluation and Democracy

The presentation on which this article is based came on the final day of the 2000 EES conference. By agreement with the conference organizers, that gave me an opportunity to comment on some of the issues that arose during the conference.

Transparency

From the opening conference presentations of Katia Horber-Papazian of Switzerland and EES President Frans Leeuw of The Netherlands [see Leeuw, this issue], through a number of session presentations, a prominent theme was the need to ensure that the public has full and unfettered access to evaluation findings. Stories were shared about suppressed evaluation reports and bureaucratic delays in finalizing publication that amounted to suppression. New dissemination initiatives, like posting evaluation reports on websites for speedier and greater public access, were discussed and endorsed.

Certainly increased transparency is an important thrust in taking evaluation findings to the people. It is worth remembering, however, the long-standing and fundamental distinction between dissemination and utilization of evaluations. Transparency will mean the most when evaluation reports are presented in ways that are understandable, relevant and usable, characteristics that are often lacking in published evaluation reports. Innovative and creative forms of reporting that include actively engaging primary intended users in thinking about findings and their implications will extend and deepen the impacts of increased transparency (see Torres et al., 1996, and Patton, 1997 for examples).

Another way of deepening and extending transparency is to work diligently

with the media to improve the reporting of evaluation findings. Policy makers and the general public become much more interested in evaluations when they find their way into news reports and editorial-opinion columns; but these do not happen by chance. Evaluators interested in use have learned to cultivate reporters and become resources to them in interpreting evaluation findings fairly and fully. The journalistic ethic of balanced reporting provides a point of entree for evaluators to work with journalists, not to tell them what to report, but to help them understand more fully the implications of evaluation findings.

Do the People Want Evaluation?

The conference theme, 'Taking Evaluation to the People' begs the question of whether the people want evaluation. This was the focus of several conference discussions. In this regard, evaluators need to acknowledge that our history of taking evaluation to the people is far from unblemished.

Evaluations have wasted precious program resources, been used to obfuscate rather than illuminate, delayed urgently needed decision making and been used to oppress rather than improve. Evaluation has often been done *to* people rather than *with* people. Researchers have pursued their own research agendas under the guise of evaluation, serving their own publishing interests more than the information needs of intended users. It is not known how widespread these practices are, but as a result, it cannot be assumed that we will be greeted with open arms when we 'take evaluation to the people'.

I find that I have to begin every evaluation exercise by finding out what people's previous experiences have been with evaluation, and I find many of those experiences have been negative. Thus, when we take evaluation to the people, we're not entering virgin territory, a phrase I use advisedly to imply its opposite as the experience many associate with evaluation. Part of the context for each evaluation is the historical experience of evaluation that conditions and affects new evaluation design and implementation efforts. Moreover, for our ongoing professional learning, we need to evaluate our evaluations to find out how they are actually used and become more sophisticated about and adept at doing useful evaluations. The result of this will be better delivery on the positive promise inherent in the idea of taking evaluation to the people.

This conference also featured examples of evaluation processes and findings that were welcomed by the people. What characterized those evaluations, consistent with the findings on evaluation use (e.g. Patton, 1997), was that they focused on issues deemed relevant by the people, treated them with respect, provided them with findings they could understand and use, and engaged them in meaningful ways at critical times during the process. Particularly good examples of this were case studies presented by Mark Bitel (2000) of the United Kingdom in which he reported that the involvement of primary intended users in conducting the evaluation not only built capacity inside participating organizations for ongoing development and accountability, but also built 'social capital' in the larger communities served by those organizations.

At the same time, it is important to recognize that some individuals and stakeholder groups will resist more democratic and inclusive approaches to evaluation.

Evaluation 8(1)

Bitel (2000) reported that such efforts can be threatening to professionals and academics whose superior status and expertise may be threatened by actively involving non-specialists in evaluation deliberations. 'Powerful stakeholders may actively prevent other stakeholders from gaining access to the knowledge and tools of evaluation in order to protect their own professional interests.'

Greene (2000) has reported with unusual openness and forthrightness her frustrations in attempting to implement a deliberative democratic evaluation in a highly contentious environment at the local school district level. As her case study illustrates, democratic evaluation approaches are far from easy to carry out, can generate significant opposition and may be perceived as adversarial, thereby undermining the credibility of evaluation more generally. Moreover, evaluators need a variety of skills beyond their usual methodological training to successfully facilitate participatory evaluations, such as group facilitation, conflict resolution, negotiation and communications.

Communicating with the People

One of the barriers to taking evaluation to the people is the academic jargon we use. The language we use matters (Hopson, 2000; Patton, 2000). Making evaluation more accessible will likely require using language and concepts that are understandable to non-researchers. I find metaphors are especially helpful in this regard. In my home state of Minnesota, 'The Land of 10,000 Lakes', fishing is a primary leisure activity. When working with non-researcher groups, I like to begin by talking with them about connections between fishing and evaluation. What are the desired outcomes of fishing? What are indicators of a good day's fishing? What comparisons are involved in determining whether a fishing experience was successful? What are different kinds of fishing that involve different desired outcomes and are therefore evaluated by different criteria? In the same vein, if I were engaged in this work in Lausanne, the host city for the EES conference and the home of the Olympic movement, I would begin with examples from the Olympics, e.g. discussing different events with different evaluative criteria and varying measurement and judgment challenges.

The message in such an approach is that people already have a great deal of experience with evaluation. They routinely make comparisons and render judgments. What we offer, as a profession, is a more systematic and rigorous approach to evaluation. But in keeping with good principles of adult education, we will be most effective in taking evaluation to the people, I would hypothesize, if we build respectfully on what they already know and have experience with, rather than beginning by asking them to learn our language and jargon.

Decentralization and Evaluation

One of the major threads of the conference was consideration of the implications for evaluation of the trend towards decentralization across most European Community countries and the Community itself. Decentralization creates tensions in the interface between bottom-up, participative approaches to evaluation on the one hand and more managerial modes of evaluation – compliance, control, and auditing – on the other hand. This raises the question of how linkages

can be facilitated in the European context between efforts at participatory, people-centered evaluation processes and the desire for standardized performance indicators and uniform technical accountability.

It may be worth noting that, sociologically, tension between central authority and local actors is inherent in complex, multi-layered systems. We are surrounded by examples of such tensions.

- Overall university administration versus individual departments.
- National voluntary organizations versus local chapters.
- National programs versus local implementation projects.
- Federal mandates versus state and local control.

Such tensions, being natural and inevitable, can be thought of as challenges to manage rather than problems to eliminate. With such a reframing, one potential role for evaluation is to create dialogues between center-based stakeholders and regionally or locally based stakeholders, ensuring that their multiple and contrasting perspectives are represented in comprehensive evaluation designs as well as in interpretation of findings. A particularly fruitful empirical basis for such interchanges is to engage in dialogue about what can be learned from system-wide quantitative indicators in relation to local-level case studies and qualitative data (Patton, 1990). Local-level qualitative data and case studies can illuminate, provide context for and add meaning to system-wide indicators, while those same indicators can help local-level program staff and community participants understand how their experiences and observations compare with larger patterns. In concrete terms, for example, evaluators would facilitate a dialogue between gatherers and reporters of performance indicators and gatherers and reporters of community-based case stories. This would not take the form of conflict resolution where one perspective must 'win' out – be deemed right or true – but rather would aim to enhance understanding of multiple perspectives and system complexities, and thereby inform (rather than presume to direct or determine) policy formulation and program improvement. Such dialogue and deliberation would necessarily include not only varying empirical findings but also explicit attention to diverse values and therefore the intersection of knowledge and values as a foundation of democratic decision making. An important implication of this approach is that budgets for decentralization initiatives need to include resources to support this kind of center–periphery evaluative interaction.

Another dimension of what is sometimes framed as top-down versus bottom-up approaches to evaluation is to realize that taking evaluation to the people needs to occur throughout systems from top to bottom. Much of the work in participatory and democratic evaluation approaches discussed earlier has focused on inclusion and greater involvement of traditionally marginalized people at the community level. But, as Elliot Stern (1999) has persuasively articulated, evaluation also needs to be taken to parliaments. In so doing, he would urge parliaments to hold programs and departments accountable not only for results but also for what they learn.

In the United States evaluators working through legislative audit commissions have been finding it important to involve legislators of different political parties

Evaluation 8(1)

together in a dialogue aimed at setting annual evaluation priorities for state legislatures as well as interpreting subsequent reports. These face-to-face meetings often take place as half-day or full-day retreats that include a subplot of training legislators in evaluative thinking and analysis.

Finally, synthesis evaluations and meta-analyses that combine evaluation findings from a number of local-level studies to generate more system-wide patterns and lessons are an increasingly important way of integrating decentralized evaluations for use by policy makers at the central system level. At the same time, local-level implementers and operators can benefit from knowing about and understanding the results of synthesis evaluations. One way of supporting dialogue between center and periphery is to bring together cross-national, national and local decision makers and evaluators to design synthesis studies and then, later, come together again to discuss and deliberate on the implications of the findings from synthesis evaluations and meta-analyses. Such occasions would also be an especially opportune time to consider how theory can inform findings and practice, a challenge highlighted by EES President Frans Leeuw in his opening presentation at the conference [see Leeuw, this issue]. Again, an important implication of this approach is that budgets for decentralization initiatives need to include resources to support this kind of center-periphery evaluative interaction and collaboration.

Democratic Evaluation and Methodological Quality

More than once at the conference I heard concerns expressed that democratic and participatory approaches to evaluation reduce methodological quality. This is a crucial issue, for if evaluation is to play a supporting role in strengthening democracy, its validity, quality, credibility and independence are the pillars upon which that contribution will be built.

A beginning point to consider this issue, perhaps, is recognition that standards of technical quality vary for different users and varying situations. The issue is not meeting some absolute research standards of technical quality but, rather, making sure that methods and measures are *appropriate* to the validity and credibility needs of a particular evaluation purpose and specific intended users.

Jennifer Greene (1990) examined in depth the debate about 'technical quality versus user responsiveness'. She found general agreement that both are important, but disagreements about the relative priority of each. She concluded that the debate is really about how much to recognize and deal with evaluation's political inherency.

Evaluators should recognize that tension and conflict in evaluation practice are virtually inevitable, that the demands imposed by most if not all definitions of responsiveness and technical quality (not to mention feasibility and propriety) will characteristically reflect the competing politics and values of the setting. (Greene, 1990: 273)

She then recommended that evaluators 'explicate the politics and values' that supply the basis for decisions about purpose, audience, design and methods.

In the utilization literature, the issue of the relationship between technical

quality and user involvement is sometimes described as a truth and utility trade-off. Stakeholders want accurate information; they apply 'truth tests' (Weiss and Bucuvalas, 1980) in deciding how seriously to pay attention to an evaluation. They also want useful and relevant information. The ideal, then, is both truth and utility. In the real world, however, there are often choices to be made and trade-offs to be negotiated.

The simplest example of such a choice is time. The timelines for evaluation are often ridiculously short. A decision maker may need whatever information can be obtained in three months, even though the researcher insists that a year is necessary to get data of reasonable quality and accuracy. This involves a trade-off between truth and utility. Highly accurate data in a year are less useful to this decision maker than data of less precision and validity obtained in three months.

Decision makers regularly face the need to take action with limited and imperfect information. They prefer more accurate information to less, but they also prefer some information to no information. This is why research quality and rigor are 'much less important to utilization than the literature might suggest' (Alkin et al., 1979: 24).

The effects of methodological quality on use must be understood in the full context of a study, its political environment, the degree of uncertainty with which specific decision makers are faced, and thus their relative need for any and all clarifying information. If information is scarce, then new information, even of less-than-ideally-desired quality, may be somewhat helpful.

The scope and importance of an evaluation greatly affect the emphasis that will be placed on technical quality. Eleanor Chelimsky, former Director of the Program Evaluation and Methodology Division of the United States General Accounting Office, has insisted that technical quality is paramount in policy evaluations to Congress. The technical quality of national policy research matters not only in the short term, when findings first come out, but over the long term as policy battles unfold and evaluators are asked to explain and defend important findings (Chelimsky, 1995). But such debates about technical quality are likely to be of more central import in national policy evaluations than in local efforts to improve programs at the street level. At the local level, the concrete challenges of day-to-day program delivery are attended to prior to larger policy-making concerns.

Another factor that can reduce the weight decision makers give to technical quality is scepticism about the return on investment of large-scale, elaborately designed, carefully controlled, multi-year and expensive studies. Cohen and Weiss (1977) reviewed 20 years of policy research on race and schools, and found progressive improvement in research methods (i.e. increasingly rigorous designs and ever more sophisticated analytical techniques). Sample sizes increased, computer technology was introduced, multiple regression and path analytic techniques were employed, and more valid and reliable data-gathering instruments were developed. After reviewing the findings of studies produced with these more rigorous methods, as well as the uses made of their findings, they concluded that 'these changes have led to more studies that disagree, to more qualified conclusions, more arguments, and more arcane reports and unintelligible

Evaluation 8(1)

results' (Cohen and Weiss, 1977: 78). In light of this finding, simple, understandable and focused evaluations have great appeal to practitioners and action-oriented evaluation users.

Technical quality ('truth tests') may get less attention than researchers desire because many stakeholders are not very sophisticated about methods. Increasing their sophistication is one of the impacts of process use, as discussed earlier. Yet, in my experience, many non-researchers know (almost intuitively) that *the methods and measurements used in any study are open to question and attack*. They know that researchers don't agree among themselves about technical quality. As a result, experienced decision makers apply less rigorous standards than academics and, as long as they find the evaluation effort credible and serious, they're more interested in discussing the substance of findings than in debating methods. Credibility involves more than technical quality, though that is an important contributing factor. Credibility, and therefore utility, is affected by:

the steps we take to make and explain our evaluative decisions, [and] also intellectually, in the effort we put forth to look at all sides and all stakeholders of an evaluation. (Chelimsky, 1995: 219)

The perception of impartiality and balance is at least as important as methodological rigor in highly political environments.

As no study is ever methodologically perfect, it is important for primary stakeholders to know firsthand what imperfections exist – and to be included in deciding which imperfections they are willing to live with in making the inevitable leaps from limited data to incremental action.

The common perception of decisions about methods among non-researchers is that such decisions are primarily technical in nature. Sample size, for example, is determined by a mathematical formula. The evaluation methodologist enters the values of certain variables, makes calculations and out pops the right sample size to achieve the desired level of statistical robustness, significance, power, validity, reliability, generalizability, etc.; technical terms that dazzle, impress and intimidate practitioners and non-researchers. Evaluation researchers have a vested interest in maintaining this technical image of scientific expertise for it gives us prestige, inspires respect and, not incidentally, it leads non-researchers to defer to us, essentially giving us the power to make crucial methods decisions and then interpret the meaning of the resulting data. It is not in our interest, from the perspective of maintaining prestige and power, to reveal to intended users that methods decisions are far from purely technical. But, contrary to public perception, evaluators know that methods decisions are rarely, if ever, purely technical. Ways of measuring complex phenomena involve simplifications that are inherently somewhat arbitrary, are always constrained by limited resources and time, inevitably involve competing and conflicting priorities, and rest on a foundation of values preferences that are typically resolved by pragmatic considerations, disciplinary biases and measurement traditions.

Democratic evaluations debunk the myth that methods and measurement decisions are purely technical. Non-researchers then develop more understanding about both the technical and non-technical dimensions of evaluation.

Moreover, we know that use is enhanced when practitioners, decision makers and other users fully understand the strengths and weaknesses of evaluation data, and that such understanding is increased by being involved in making methods decisions. We know that use is enhanced when intended users participate in making sure that, as trade-offs are considered, as they inevitably are because of limited resources and time, the path chosen is informed by relevance. We know that use is enhanced when users buy into the design, and find it credible and valid within the scope of its intended purposes as determined by them. And we know that when evaluation findings are presented, the substance is less likely to be undercut by debates about methods if users have been involved in those debates prior to data collection (Patton, 1997).

At their roots, participatory evaluations are informed by a fundamental confidence in the wisdom of an informed citizenry and a willingness to engage ordinary citizens respectfully in all aspects of evaluation, including methodological discussions and decisions. This point is worth emphasizing because some – not all, to be sure, but some – resistance to participatory evaluation derives from the status associated with research expertise and an elitist or patronizing attitude towards non-researchers (they are, after all, ‘subjects’). Egon Guba has described in powerful language this archetype.

It is my experience that evaluators sometimes adopt a very supercilious attitude with respect to their clients; their presumptuousness and arrogance are sometimes overwhelming. We treat the client as a ‘child-like’ person who needs to be taken in hand; as an ignoramus who cannot possibly understand the tactics and strategies that we will bring to bear; as someone who doesn’t appreciate the questions he *ought* to ask until we tell him – and what we tell him often reflects our own biases and interests rather than the problems with which the client is actually beset. The phrase ‘Ugly American’ has emerged in international settings to describe the person who enters into a new culture, immediately knows what is wrong with it, and proceeds to foist his own solutions onto the locals. In some ways I have come to think of evaluators as ‘Ugly Americans’. And if what we are looking for are ways to manipulate clients so that they will fall in with *our* wishes and cease to resist our blandishments, I for one will have none of it. (Guba, 1977: 1)

For others who will have none of it, one way to address the issue of methodological quality in democratic evaluations is to reframe the evaluator’s function from an emphasis on generating expert judgments to an emphasis on supporting informed dialogue, including methodological dialogue. The traditional expert-based status of evaluators has fueled the notion that we provide scientifically based answers and judgments to policy makers while, by our independence, we assure accountability to the general public. Playing such a role depends on a knowledge paradigm in which correct answers and completely independent judgments can be conceived of as existing.

But the real world of research findings is better characterized by probabilities than certainties, and empirical evidence is only one factor in decision making. Having evaluators play more of a facilitative role than expert-judgment role derives, in part, from the influences of postmodernism, deconstruction, critical theory, feminist theory, empowerment evaluation and constructivism, among

other perspectives, all of which share, to some degree, scepticism about the traditional truth-oriented knowledge paradigm. They offer, in contrast, an emphasis on interest-acknowledged perspectives articulated and interpreted within an explicit context (political, social, historical, economic and cultural).

What do they Offer for Evaluation?

They offer informed dialogue and deliberation among various stakeholders rather than expert-based, independent truth and judgment. Properly conceived and facilitated – no small task, I acknowledge – evaluation becomes a process and mechanism for interaction and interface among those with different perspectives and locations in society (top, bottom and middle). Such facilitation and deliberation can occur not only at the local level, but also between central authorities and local actors as described in the previous section on decentralization. This offers a vision of evaluation as a central pillar in support of deliberative democracy in the postmodern knowledge age.

References

- Alkin, M. (1997) 'Stakeholder Concepts in Program Evaluation', in A. Reynolds and H. Walberg (eds) *Evaluation for Educational Productivity*. Greenwich, CT: JAI Press.
- Alkin, M. C., R. Daillak and P. White (1979) *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills, CA: Sage.
- Arendt, H. (1968) *Between Past and Future: Eight Exercises in Political Thought*. New York: The Viking Press.
- Bitel, M. (2000) 'Taking Evaluation to the People. Who Wants It?', Paper presented at the European Evaluation Society Conference, Lausanne, 12–14 October.
- Chelimsky, E. (1995) 'The Political Environment of Evaluation and What It Means for the Development of the Field', *Evaluation Practice* 16(3): 215–25.
- Cohen, D. K. and J. A. Weiss (1977) 'Social Science and Social Policy: Schools and Race', in C. H. Weiss (ed.) *Using Social Research in Public Policymaking*, pp. 67–84. Lexington, MA: D. C. Heath.
- Cousins, J. B. and L. M. Earl (eds) (1995) *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning*. London: Falmer Press.
- Fetterman, D. M., A. J. Kaftarian and A. Wandersman (eds) (1996) *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*. Newbury Park, CA: Sage.
- Greene, J. C. (1990) 'Technical Quality Versus User Responsiveness in Evaluation Practice', *Evaluation and Program Planning* 13(3): 267–74.
- Greene, J. C. (2000) 'Challenges in Practicing Deliberative Democratic Evaluation', in K. E. Ryan and L. DeStefano (eds) *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation, New Directions for Evaluation* 85. San Francisco, CA: Jossey-Bass.
- Guba, E. (1977) 'Overcoming Resistance to Evaluation', Paper presented at the Second Annual Conference of the North Dakota Study Group on Evaluation, North Dakota, USA.
- Hopson, R. (ed.) (2000) *How and Why Language Matters in Evaluation, New Directions for Evaluation* 86. San Francisco, CA: Jossey-Bass.
- House, E. R. and K. R. Howe (2000) 'Deliberative Democratic Evaluation', in K. E. Ryan

Patton: A Vision of Evaluation that Strengthens Democracy

- and L. DeStefano (eds) *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation, New Directions for Evaluation* 85, pp.13–26. San Francisco, CA: Jossey-Bass.
- Kushner, S. (2000) *Personalizing Evaluation*. London: Sage.
- MacDonald, B. (1987) 'Evaluation and Control of Education', in R. Murphy and H. Torrance (eds) *Issues and Methods in Evaluation*. London: Paul Chapman.
- Mertens, D. M. (1998) *Research Methods in Education and Psychology: Integrating Diversity with Quantitative and Qualitative Approaches*. Thousand Oaks, CA: Sage.
- Mertens, D. M. (1999) 'Inclusive Evaluation: Implications of Transformative Theory for Evaluation', *American Journal of Evaluation* 20(1): 1–14.
- Patton, M. Q. (1990) *Qualitative Evaluation and Research Methods*. Newbury Park, CA: Sage.
- Patton, M. Q. (1997) *Utilization-Focused Evaluation: The New Century Text*, 3rd edn. London: Sage.
- Patton, M. Q. (1998) 'Discovering Process Use', *Evaluation* 4(2): 225–33.
- Patton, M. Q. (2000) 'Language Matters', in R. Hopson (ed.) *How and Why Language Matters in Evaluation, New Directions for Evaluation* 86. San Francisco, CA: Jossey-Bass.
- Ryan, K. E. and L. DeStefano (eds) (2000) 'Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation', in *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation, New Directions for Evaluation* 85, pp. 13–26. San Francisco, CA: Jossey-Bass.
- Stern, E. (1999) 'Why Parliament Should Take Evaluation Seriously', *Parliamentary Review*.
- Torres, R. T., H. S. Preskill and M. E. Piontek (1996) *Evaluation Strategies for Communicating and Reporting: Enhancing Learning in Organizations*. London: Sage.
- Weiss, C. H. and M. Bucuvalas (1980) 'Truth Tests and Utility Tests: Decision Makers' Frame of Reference for Social Science Research', *American Sociological Review* 45 (April): 302–13.

MICHAEL QUINN PATTON is former President of the American Evaluation Association (AEA) and author of *Utilization-Focused Evaluation* (1997) and *Qualitative Evaluation and Research Methods* (3rd edn, 2002). He received the Myrdal Award for Contributions to Useful and Practical Evaluation Practice from the Evaluation Research Society and the Lazarsfeld Award for Contributions to Evaluation Theory from the AEA. Please address correspondence to: The Union Institute, Utilization-Focused Evaluation, 3228 46th Ave South, Minneapolis MN 55406, USA. [email: MQPatton@Prodigy.net]