A Vision of Evaluation that Strengthens Democracy

By Michael Quinn Patton

Context

Let me explain how this title came about.

At about the same time that European Evaluation Society President Frans Leeuw invited me to keynote this fourth EES conference in Lausanne, my local coffee shop went out of business. It had only been open for two years, having been part of a surge of new coffee shops throughout the United States as Americans fell in love with cappuccino. This particular coffee shop was quite unique because it was a "biker" shop. That is, it catered to aficionados of large Harley Davidson motorcycles affectionately called "Hawgs." The owners, Scott and Connie, a young husband and wife team, were also bikers. Because the coffee shop was only a block from my office, I had become a regular, despite lacking the appropriate leather attire and loud steel machine. One morning Connie mentioned that she had decided to try the new laser surgery to correct her near-sighted vision. I left for a trip wishing her a positive outcome from the surgery. When I returned two weeks later the coffee shop was closed and a barely legible handwritten note on the door said:

Closed indefinitely -- vision problem

I made inquiries at a nearby petrol station but all I learned was that the shop had closed very suddenly without notice. Some three weeks later I happened to see Scott riding his motorcycle on the street and waved him over to ask how Connie was doing. He said she was fine. What about the sign on the coffee shop door? That had nothing to do with Connie's surgery, he explained. "I just couldn't vision myself serving coffee the rest of my life."

That got me thinking about the vision for evaluation encompassed in the conference theme "Taking Evaluation to the People: Between Civil Society, Public Management and the Polity." This seemed to suggest a connection between evaluation and democracy. So I began pondering the question: In what ways might evaluation strengthen and support democracy?

Fieldwork

To help me contemplate this question, I used the months leading up to this conference to do some fieldwork. I started asking those I work with what connection they saw between evaluation and democracy. The results have been interesting and informative. Let me share some responses to provide context for considering what it might mean for evaluation to strengthen democracy.
The Oregon Program Evaluation Network (OPEN) invited me to keynote their annual conference. I mentioned the European Evaluation Society conference theme, told the audience about my fieldwork on the connection between evaluation and democracy, and asked anyone with ideas or reactions to see me at the break. When I concluded, I was immediately surrounded by a number of graduate students who told me they were excited by the idea of a vision of evaluation that was more than methods, techniques, and reports -- a vision that suggested our profession could contribute to strengthening fundamental principles and practices of democratic participation. They were hungry to learn more. Their enthusiasm for a larger vision of evaluation was palpable.

I was consulting on the design of a large-scale, statewide arts initiative supported by a major philanthropic foundation. We were involved in a daylong logic modeling exercise. Just before lunch I told the group of diverse stakeholders about the EES conference theme and asked, rather sheepishly, if they saw any connection between what we were doing and democracy. That led to an animated lunchtime discussion about general public access to the arts and how art contributes to the quality of life in a democracy. That discussion carried over into the logic modeling exercise and evaluation design as arts accessibility became a component of the program and a question for the evaluation.

In working with a group of environmental educators and evaluators I asked the participants if they saw any connection between the issues we were discussing and democracy. The result was conceptualizing something they called "ecological citizenship" as both a desired program outcome and focus for evaluation.

In facilitating an evaluation design with diverse stakeholders for an agricultural extension program, I again asked about the connection to democracy. A farmer jumped to his feet and said something like this: "This is the fourth study I've been asked to participate in over the last 10 years. All of them promised to send us back findings. None did so. I don't know what you mean about strengthening democracy. I'd be happy just to get the findings that we've been promised."

While facilitating an inner city community development evaluation I continued my fieldwork. Asking whether what we were doing had anything to do with democracy turned a deadly session on evaluation methods and survey design into a lively discussion about community-level activism and evaluation as a way of sharing information and perceptions within their own community.
At a graduate seminar on research methods that included a large number of African-American students, I asked for their reaction to the EES conference theme and my draft keynote speech title. This led to an animated discussion about whether -- and possibly how -- evaluation could be used to capture and communicate diverse perspectives, especially providing a mechanism for the voices of those outside the inner sanctum of power to be heard by those in power. It gave me an opportunity, which had not been part of the original agenda, to share work on the use of evaluation to "give voice" to the disenfranchised, underprivileged, poor, and others outside the mainstream (Weiss and Greene, 1992:145). I told them about an evaluation of a diversity project we had done in the Saint Paul, Minnesota public schools in which a major part of the design involved capturing and reporting the experiences of African-American, Native American, Chicano-Latino, and Southeast Asian (Hmong) parents. Offering a way for them to tell their stories to mostly white, corporate funders was an intentional part of the design, one approved by those same white corporate funders. The final report was written as a multi-vocal, multicultural presentation that presented different experiences with and perceptions of the program's impacts rather than reaching singular conclusions or making recommendations. The medium of the report carried the message that multiple voices needed to be heard and valued as a manifestation of diversity (Stockdill et al, 1992). The findings were used for both formative and summative purposes, but the parents and many of the staff were most interested in using the evaluation processes to make themselves heard by those in power. Being heard was an end in itself quite separate from use of findings. At the seminar where I described this evaluation, the liveliest discussion during the five days we were together occurred around this unplanned topic – how evaluation could be a mechanism to give voice to those whose voices are seldom heard by those in power and thereby, possibly, strengthen democracy by supporting inclusion and dialogue.

Inquiry Into Evaluation and Democracy

The preceding vignettes illustrate the range of reactions I got from my "fieldwork" into how people react to being asked about the connection between evaluation and democracy. Typically, the question first elicited a furrowed brow. It is, after all, a rather pompous query. Evaluation and democracy? Evaluation strengthening democracy? Bombastic. Grandiose. Pretentious. Platitudinous. One person replied sarcastically: "I think you need to make a bit more progress in getting evaluators to do something minimally useful -- like writing relevant and understandable reports -- before you take on democracy."

But most people I engaged in this inquiry, after the initial furrowed brow, took the topic seriously and responded in a way that suggested they cared about the connection. Moreover, as I believe the vignettes show, the very asking of the question, as is so
often the case, constituted a kind of intervention. In several cases, introducing the idea of a link between evaluation and democracy changed the course of subsequent discussion. Public access to the arts. Ecological citizenship. Giving voice to those whose voices typically go unheard in corridors of power. Participatory evaluation.

The reactions I received are, I think, sufficiently intriguing that I would invite collaborators in carrying on this inquiry. The next time you are involved with some stakeholders in designing an evaluation or presenting findings for their reflection, look for an opportunity to ask something like this: "Just curious but, do you suppose that the evaluation work we’re engaged in here might, in anyway, strengthen democracy? Now I know that may seem a strange question, but at a conference of the European Evaluation Society, it was suggested that the question might be worth pondering. What do you think?"

Imagine being involved in an audit process, or a performance indicators exercise, or any of a host of evaluation activities and, at the odd moment, interjecting: "Does anything we’re doing here strengthen democracy?" In that regard, perhaps the "vision thing," as American political pundits now refer to public discussions of vision, is not so much at this point a vision of evaluation strengthening democracy as it is a vision of evaluators engaging in serious inquiry into the possible connections between evaluation and strengthened democracy.

Additional Global Context:
Evaluating Democracy as an Intervention

The 4th EES conference, and therefore the presentation upon which this article is based, took place at the time that people in Yugoslavia were assembled in the streets to validate the election results that removed Slobodan Milosevic from the presidency. At such a time democracy is not an abstraction, but a globally galvanizing street action (carried live on CNN) in which ordinary people risk their lives and fortunes for their children’s and country’s future in support of what they perceive as the greater good.

The larger, longer-term context for the conference deliberations included renewed attention in development efforts to the role of democratic processes and civil society infrastructures in sustainable economic, social, and political development. For example, democracy and civil society have emerged as hypothesized interventions in the reduction of poverty. Some of us, at least, may find ourselves invited to evaluate this hypothesized connection and the programs that derive from and provide a test of this intervention model. How we attend to and evaluate the contributions of civil society to poverty reduction may be one of the avenues through which, in our knowledge generating function, we contribute to strengthening democracy itself. Let me offer examples of what I am referring to.
In September 1999 the World Bank expanded its poverty alleviation mandate by promoting the formulation of new country strategies to reduce poverty in the developing world by using participatory approaches and collaborating with nongovernmental organizations in ways that promote and strengthen civil society. In a broad sense, a strong democratic infrastructure is coming to be viewed as a necessary condition for sustainable economic development that reduces widespread poverty. (For details see the World Bank Website, "Poverty Reduction Strategy Paper: Special Link for Civil Society.")

The European Commission has approved a policy document that sets out the parameters for a new European Community Development Policy aimed at poverty eradication. Based on lessons learned from past efforts, the Commission proposes that future efforts should direct more attention to and be more serious about the ownership of the development process. This implies more bottoms-up, participatory strategies in designing and implementing poverty eradication projects. Such efforts clearly invite evaluation of both participatory processes and resulting outcomes, including effects on civil society and sustainable democracy. (For details see the European Commission Website.)

Harvard social scientist Robert D. Putnam has been one of the major popularizers of the idea that the health of civil society is a leading indicator of the health of democracy more generally. His thesis that U.S. citizens are increasingly retreating from civic life, which he defines roughly as engaging in ongoing activities with non-intimates in ways that build social community, or "social capital," has received a great deal of attention. His data show that fewer Americans are involved with bowling leagues, reading groups, charitable volunteer projects, local political party work, neighborhood associations, and other such voluntary associations (Putnam, 1995). But Putnam's thesis originated from work in Europe, specifically Italy. In a book entitled Making Democracy Work: Civic Traditions in Modern Italy, he investigated government effectiveness and civil society in various regions of Italy. He concluded:

Some regions of Italy... are blessed with vibrant networks and norms of civic engagement, while others are cursed with vertically structured politics, a social life of fragmentation and isolation, and a culture of distrust. These differences in civic life turn out to play a key role in explaining institutional success (Putnam, 1993: 15; emphasis added).

Putnam's work is but one example of research and theory that posits connections between civil society, healthy democracy, and institutional (including but not limited to government) effectiveness. (See, for example, The Good Society Symposium on Commonwealth, Civil Society, and Democratic Renewal, 1999.)

In different ways and from different perspectives, each of these three examples affirms the importance of the connections called to our attention by the sub-theme of this fourth
EES conference, the connections between civil society, public management and the polity. The new civil society initiatives of the World Bank and the European Commission, and the social science research and theory on civil society, offer a challenging agenda for evaluation researchers on democracy as a means to institutional effectiveness and poverty reduction.

But when we call forth a vision of evaluation strengthening democracy, we are not thinking so much about democracy as an instrument or means for accomplishing other desired ends, but rather democracy as an end in itself. Let me turn, then, to ways in which evaluation may contribute to strengthening democracy in accordance with the EES conference theme “Taking Evaluation to the People.”

Democratic Evaluation

Over its relatively short history, the field of professional evaluation has 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 information broker between groups who want and need knowledge of 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 critics of those in power. He places at the center of evaluation the experiences of people in programs. The experiences and perceptions of the people in programs, the supposed beneficiaries, is where, for Kushner, we will find the intersection of Politics (big P -- Policy) and politics (small p - people). Much of evaluation these days (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, 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, do 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: 4)

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 to evaluation in support of those whose stakes tend to be underrepresented 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: 4).

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 thinking evaluatively. In this regard we might 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 distill 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 dual tracks of impact in strengthening democracy: (1) a more informed electorate through use of findings and (2) a more thoughtful and deliberative citizenry though 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 to many people; indeed, they can seem quite alien. When we take people through 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 from an evaluation than use of specific findings generated in 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 having 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, that is, 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 ten 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 ten 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 "Outcome 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 key influencers 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.

On other fronts we are working with grassroots community development efforts to build training in evaluative thinking into their efforts.

What these initiatives illustrate is the potential for impact opened up when we think about the role of evaluation in strengthening democracy and think 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 (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’ve been suggesting in this section that when we think about the conference theme “Taking Evaluation to the People,” we include not only giving them evaluation findings but also offering them opportunities to learn how to think and dialogue evaluatively.

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, through a number of session presentations, a prominent theme was the need to assure 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. Policymakers and the general
public become much more interested in evaluations when they find their way into news
reports and editorial opinion columns. These things do not happen by chance.
Evaluators interested in use have learned to cultivate reporters and become resources to
them in interpreting fairly and fully evaluation findings. 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. We don’t know how widespread these practices are, but we do know that we
cannot assume 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 used advisedly to imply its opposite as the experience many associate with evaluation. Part of the context for each evaluation is the historical experience with 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, thereby better delivering 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. 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 off, 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, skills like 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 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 began 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 toward decentralization across most European Community countries and the Community itself. Decentralization creates tensions in the interface between bottoms 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, assuring 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 to 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, not in a conflict resolution mode where one perspective must “win” out -- that is, be deemed right or true -- but for the purpose of enhancing understanding of multiple perspectives and system complexities, and thereby informing (rather than presuming 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 Elliott 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 together in a dialogue aimed at setting annual evaluation priorities for state legislatures as well as interpreting subsequent reports. These days-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 is an increasingly important way of integrating decentralized evaluations for use by policymakers 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. 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” (p.273). She then recommended that evaluators "explicate the politics and values" that undergird 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 matter of truth and utility trade-offs. 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 accurate information, 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 affects the emphasis that will 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 called on to explain and defend important findings (Chelimsky, 1995). But such debates about technical quality are likely to be much more center stage in national policy evaluations than in local efforts to improve programs at the street level where the policy rubber hits the day-to-day programming road.

Another factor that can reduce the weight decision makers give to technical quality is skepticism 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 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 methods decisions 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 become savvier 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 its 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 toward 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 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, skepticism 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 tasks 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.

When in Rome:

Concluding reflections

Let me can conclude by thanking all those involved with EES for a stimulating conference and for setting me on this eight-month odyssey of thinking about a vision of evaluation that strengthens democracy. I conclude my fieldwork with a new sense of gratitude for living at a time when we can actually imagine engaging in processes and doing work that strengthens democracy, when such a vision, as grandiose, even bombastic, as it may be, can be openly shared, discussed, and debated. Some final context may help explain this sense of appreciation.

Prior to the EES conference I had the opportunity to participate in a seminar sponsored by the Italian Evaluation Society. While in Rome I wandered among the ruins of ancient Rome, visited the ruins of the great Roman Forum, and spent some meditative time in the remains of the ancient Senate seeking inspiration about what I might say at the EES conference regarding evaluation's potential contributions to democracy. Nothing came to me, at least nothing about evaluation. I couldn't get past vivid images of Caesar's death in that place as portrayed by Shakespeare.

After leaving the Forum, I walked across to the Coliseum. There, standing in the central platform area reserved for the Senators, I got a distinct image. I imagined an evaluator presenting important policy findings to the citizens of Rome in the Coliseum. At the end of the report, the Emperor would invite the crowd to render a thumbs-up or thumbs-down on the evaluation. Thumbs-up would mean a laurel wreath, coin of the realm, and an all expenses paid trip to present at the annual meeting of the Imperial Evaluation Society. Thumbs-down would mean the lions got fed. I left the Coliseum quickly.
So I'm delighted to be engaged in evaluation at a time when the stakes, though high, are not quite so high as my Coliseum vision, and when, instead, in the beautiful city of Lausanne, we can, as an international community of professionals, spin visions of strengthening democracy through our evaluation work together.

References


