American Journal of Evaluation

Toward Distinguishing Empowerment Evaluation and Placing It In A Larger Context

Michael Quinn Patton

The online version of this article can be found at:

Published by:

$SAGE

On behalf of:

American Evaluation Association

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

Email Alerts: http://aje.sagepub.com/cgi/alerts

Subscriptions: http://aje.sagepub.com/subscriptions

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

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

**Citations** http://aje.sagepub.com/cgi/content/refs/18/1/147

Toward Distinguishing Empowerment Evaluation and Placing It In A Larger Context

MICHAEL QUINN PATTON

ABSTRACT

Fostering self-determination is the defining focus of empowerment evaluation and the heart of its explicit political and social change agenda. However, empowerment evaluation overlaps participatory, collaborative, stakeholder-involving, and utilization-focused approaches to evaluation in its concern for such issues as ownership, relevance, understandability, access, involvement, improvement, and capacity-building. A critical question becomes how to distinguish empowerment evaluation from these other approaches. Making such distinctions has become critical as the field debates the boundaries and implications of empowerment evaluation.
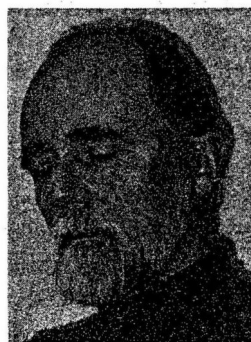
Michael Quinn Patton

Review essay about Empowerment Evaluation: Knowledge and Tools for Self-Assessment & Accountability, edited by David M. Fetterman, Shakeh J. Kaftarian, and Abraham Wandersman. Thousand Oaks, CA: Sage Publications, 1996.

WORDS AS SOCIOLOGICAL CONTEXT

How long a time lies in one little word!

Shakespeare
King Richard II

Words mark decades.

Rock 'n roll in the fifites. *Altered consciousness* in the sixties. *Paradigm shift* in the seventies. *Proactive* in the eighties. *Empowerment* in the nineties.

It's what's happening now. Or what's not happening. Or what shouldn't happen, depending on your perspective and praxis. The bridge to the twenty-first century will be built with words. Some worry that that's all it will be built of.

Vanderplaat (1995), writing from a European perspective, locates empowerment evaluation in the larger context of emancipatory research that grew out of Freire's liberation peda-

Michael Quinn Patton • Union Institute, 3228, 46th Avenue South, Minneapolis, MN 55406.

147

gogy (1970), feminist inquiry (e.g., Harding, 1987; Maguire, 1987), critical theory (ref. Forester, 1985) and communicative action (Habermas, 1984, 1987). More directly related to evaluation, empowerment evaluation draws on and raises the stakes of participatory action research (Whyte, 1991; Wadsworth, 1993; King, 1995) and collaborative evaluation (Cousins & Earl, 1992, 1995). Fetterman, in the volume here under review, traces the roots of empowerment evaluation to community psychology, action anthropology, the school reform movement, and grassroots community development — influences reflecting his own professional and intellectual journey.

The phrase "empowerment evaluation" gained prominence in the lexicon of evaluation when Fetterman, as President of the American Evaluation Association (AEA), made it the theme of the Association's 1993 National Conference. The volume here under review grew out of that conference. Karen Kirkhart, President of AEA in 1994, provided an additional platform for discussing empowerment by choosing "Evaluation and Social Justice" as the theme for the national conference over which she presided. The importance of the idea of empowerment evaluation as a frontier of evaluation practice was further recognized by the profession when David Fetterman, in 1995, and Shakeh Kaftarian, in 1996, won the Alva and Gunnar Mydral Award for Evaluation Practice, and Margret Dugan, in 1995, won the Guttentag Award as a promising evaluator — all authors in the volume under review and all recognized in part for their work and writings on empowerment evaluation.

THEORY AND PRACTICE: DEFINING EMPOWERMENT EVALUATION

Fetterman's introductory and concluding chapters are the bookends that support 14 other chapters, organized in four parts: Breadth and Scope; Context; Theoretical and Philosophical Frameworks; and Workshops, Technical Assistance, and Practice. Fetterman defines empowerment evaluation as "the use of evaluation concepts, techniques, and findings to foster improvement and self-determination" (p.5). Of course, using evaluation processes for improvement is nothing new; it's the emphasis on fostering self-determination that is the defining focus of empowerment evaluation, and the heart of its explicit political and social change agenda. Fetterman elaborates five "facets" of empowerment evaluation: (1) training participants to conduct their own evaluations, i.e., capacity-building; (2) evaluators as facilitators and coaches rather than judges; (3) evaluators advocating on behalf of groups that are disempowered and/or supporting disempowered groups in advocating for themselves; (4) illumination and (5) liberation for those involved. One of the weaknesses of the book, at the level of conceptualization and theory, is that it does not provide a conceptual continuum depicting varying degrees of empowerment emphasis appropriate for differing situations and contingencies. Moreover, because empowerment evaluation overlaps participatory, collaborative, stakeholder-involving, and utilization-focused approaches to evaluation in its emphasis on attending to such issues as ownership, relevance, understandability, access, and involvement, a critical question becomes how to distinguish empowerment evaluation from these other approaches. Dugan, in her chapter on "Participatory and Empowerment Evaluation," addresses this issue most directly in the book.

> Some projects begin as participatory and evolve to a more empowering approach, as our work did. Empowerment evaluation assumes that (a) the evaluation is used explicitly to contribute to the process, and (b) participants not only are involved but also control the

process (Fetterman, 1995). In general, participatory work follows a continuum from limited participation to an ideal of full control. Empowerment evaluation begins closer to the end of participatory work (p. 283).

Dugan's chapter, a reflective sharing of her own experiences and learnings from working closely and collaboratively with program participants, depicts a developmental process (though, she notes, not always a linear one) that deepened participation over time. She's forthright about difficulties: "This process takes time and patience. For example, the evaluation teams we mentored and coached have been slow to internalize evaluation as a natural part of their strategic planning. But slowly this is changing…" (p.297). In the end, however, she muddies the distinction between participatory evaluation and empowerment evaluation by summarizing their advantages and disadvantages together, as if they were really the same. This confusion permeates the book.

Cousins and Earl (1995) have distinguished collaborative and participatory approaches by goals (increase use of findings, generate social theory, participant emancipation) and degree of researcher-participant collaboration. Cousins, Donohue and Bloom (1995) have identified three dimensions along which evaluations can vary: (1) degree of researcher versus practitioner control of the process; (2) depth of participation; and (3) breadth of stakeholder participation (a continuum from a limited number of primary users to all legitimate groups). Following these efforts at classification and distinction, Fetterman's overview would suggest that, in addition to these dimensions of participation and collaboration, empowerment evaluation adds attention to and varies along the following continua: (1) the degree to which participants' power of self-determination is enhanced, that is, the extent to which "liberation" occurs; (2) the extent to which evaluators are advocates for disempowered groups or enable groups to advocate for themselves, and (3) the degree to which training participants in evaluation skills is an explicit, primary, and attained outcome of the evaluation process. The first — the liberation dimension — would seem, by definition, to be the defining characteristic of empowerment evaluation. The second and third dimensions are enabling processes in support of liberation. Fetterman's other two dimensions — facilitation and illumination — are not at all unique to empowerment evaluation. All participatory, collaborative, and utilization-focused approaches emphasize a facilitative role for evaluators and include illuminative outcomes for participants.

These distinctions become critical as the field debates the boundaries and implications of empowerment evaluation. With regard to the present volume, these criteria help us determine the extent to which the case examples included actually constitute empowerment examples. For there is another dimension that emerges here, the degree to which the language of empowerment is inserted into an evaluation, quite apart from whether any actual empowerment occurs or whether the evaluation process meets the criteria that distinguish empowerment evaluation. In my judgment, several of the cases presented are exemplars of participatory, collaborative, stakeholder-involving, and even utilization-focused evaluations, but do not meet the criteria for empowerment. In these cases, the language of empowerment appears to have been added as an overlay for political and personal value reasons more than because the case data fit the empowerment model. In any event, an interesting and provocative use of the book with students would be to have them analyze each case according to the degree to which Fetterman's criteria are met, which is precisely what I now propose to do.

The Clearest Empowerment Evaluation Cases

The book's first case example is Henry Levin's description of the Accelerated Schools Project. Evaluation processes are integral to school transformation processes in this example. This case also reveals how the language of evaluation can be affected by an empowerment agenda. "Taking-stock committees" gather baseline data. Adherence to values like equity and trust are assessed. "*Cadres*" analyze priorities and learn to conduct research. "Critical reflection" and "self-evaluation" support "transformation." Comparing the "taking-stock baseline" to the "school-as-a-whole vision" to examine overall accomplishments is called "revisiting the process." Evaluation is viewed as "embedded" in the empowerment process. Because the Accelerated Schools Project has an explicit empowerment agenda, the language and processes of empowerment evaluation appear to fit well in the program.

The sixth chapter (Grills, Bass, Brown, and Akers) describes the congruence between empowerment evaluation and "a tradition of activism in the African American community." The authors argue that "evaluation is part of a survival issue. Given that they exist within a reality in which their institutions are constantly under attack…, people who have been historically oppressed recognize the need to be open to self-critique" (p.129). They trace the roots of self-evaluation in African American activism to Marcus Garvey, Malcolm X, the Black Panther Party, and the civil rights movement, including parallel Hispanic community organizing activist efforts. In this case example from South Central Los Angeles, evaluation tools and techniques blend with and become integral to community organizing strategies. The evaluator openly supports the organizers' goals: "The staff of the Community Coalition bring an agenda and philosophy to the table and combine it with scientific techniques brought by the evaluator, *who is equally invested in the cause*" (p. 126, emphasis added). Needs assessment surveys not only document problems, but serve as a recruitment mechanism: an "'excuse' used to talk to neighbors" and get them to feedback meetings on the survey at which time "their involvement in crafting the campaign is then elicited" (pp.127–128). Having determined community needs, community members are asked what they want to do next. *This* is empowerment. It is affirmation and validation of the community and acts as an impetus for action with direction" (p. 133; emphasis in the original). Survey results are also used to support the Coalition's positions in political and media campaigns, and in negotiating with funders and policy makers.

As in the case of the Accelerated Schools Project, the language and processes of empowerment evaluation are congruent with the political activism of the Los Angeles Community Coalition's organizing efforts and goals. These cases demonstrate that the rhetoric of community activism and empowerment evaluation can be mutually reinforcing.

This is also the case with Andrews' chapter on "Realizing Participant Empowerment in the Evaluation of Nonprofit Women's Services Organizations." Subtitled "News From the Front Line," Andrews reports on the early stages of a long term process still unfolding. After providing a philosophical context (feminist principles, e.g., reciprocity) and program description (small community-based nonprofits serving women and families), she reports issues that have emerged in attempting to implement an interorganizational empowerment evaluation study group. Finding resources for evaluation, both money and time, has been a challenge. Fears about exploiting service recipients led to obtaining a small grant to compensate participants in a pilot study. Having good communications without spending lots of time in meetings is an ongoing challenge. Moving from informal, continuous feedback (which can seem evaluative but is easily subject to systematic bias and selective perception as staff hear what they want to hear) to more formal, systematic data collection is a large transition for small agencies

that value informality and fluidity. Issues of confidentiality and informed consent can be complicated when service recipients are involved in data collection, analysis, and reporting. Andrews makes it clear that procedures to address these and other challenges are still being developed; she wrote her chapter before the group's evaluation plan had been finalized. Solutions were being guided by both empowerment evaluation principles and practical wisdom. The sense of struggle is palpable. Ultimately, it is the grounding of these programs in feminist principles that makes the rhetoric of empowerment evaluation feel real and meaningful.

More Ambiguous Cases: Participation, Collaboration, and/or Empowerment?

The tenth chapter on "Evaluation and Self-Direction in Community Prevention Coalitions" by Stevenson, Mitchell, and Florin also conveys a sense of struggle, but here the struggle is to make the philosophy and rhetoric of empowerment evaluation fit the challenges, realities, and uncertainties of community coalitions committed to preventing alcohol and drug abuse. The authors offer a list of questions, none of which have definitive answers, that will arise, implicitly if not explicitly, in any empowerment evaluation. How is empowerment to be conceptualized and measured? Which of the many players in any given situation are to be empowered? Can one really "give evaluation away"? (p. 210) They conclude that the meaning of empowerment, and therefore, empowerment evaluation, is highly situation and context specific. What they end up describing are highly collaborative and participatory evaluation processes aimed at making evaluation useful, meaningful, developmental, and accessible. Whether those evaluation processes are also empowering remains, for me and, I think, for the authors, problematic.

They set for themselves the goals of demystifying evaluation, improving rigor, increasing local use of findings, and enhancing the learning capacity of the system being evaluated. They did not use the term "empowerment evaluation" to describe their work, though, they comment, "we do see a close connection between some aspects of our approach and the description of empowerment evaluation offered by Fetterman" (p.210). The areas of greatest congruence are building local capacity to sustain evaluation, supporting improvement, and facilitating illumination — the very areas I have suggested all participatory and collaborative approaches have in common and that fail to distinguish empowerment evaluation. Indeed, they eschew the language of liberation and advocacy. They report that when they presented their approach at a national conference on substance abuse prevention, they called it a "utilization-focused, stage-based model for improving the utility of evaluation by channeling it into a set of relevant tools for planning and self-monitoring by community prevention coalitions" (pp. 214–215). As they reflect on the language and methods of empowerment, they find a host of ambiguities and conclude: "Failure to resolve the ambiguity about what one means by *empowerment* is likely to lead to confusion in the evaluation arena" (p. 212).

Part of the confusion comes from failing to distinguish the potential empowering outcomes of any participatory, collaborative, stakeholder-involving, and/or utilization-focused evaluation from the unique political advocacy and liberation agendas of empowerment evaluation. Many kinds of participatory evaluation processes will feel empowering to those involved because their understanding and sense of mastery has been increased, they feel ownership of the evaluation, and their capacity to further engage in evaluation processes has been enhanced. However, empowerment evaluation as a distinct approach worthy of its own designation, i.e., a distinction with a difference, adds the agendas of liberation, advocacy, self-determination, and self-assessment as primary. In other words, empowering outcomes

may be felt by those involved in any evaluation, even non-participatory ones, where the illumination of findings and corresponding reduced uncertainties about effective interventions feels empowering. Empowerment evaluation, then, cannot be distinguished by empowering outcomes alone, but rather by being a participatory and capacity-building process targeted at groups lacking the power of self-determination whose actual power of self-determination is subsequently increased through the tools and processes of evaluation, including the evaluator's explicit advocacy of and working in support of the goal of self-determination.

A Central Shortcoming of this Book

As is clear from the preceding examples, this volume fails to consistently distinguish between participatory and collaborative evaluation processes that may lead to some feelings of empowerment among those involved versus empowerment evaluation as a distinct political process aimed explicitly at and therefore to be judged by its effectiveness in altering power relationships. Neither conceptually nor in the cases presented is this distinction consistently made and that, I believe, leaves the reader confused about what constitutes empowerment evaluation rather than clarifying and focusing its important niche characteristics. In other words, Fetterman's list of empowerment attributes (listed at the beginning of this review) ought not be treated as a laundry list from which one may pick or choose; rather, I would argue, for an evaluation to be genuinely an *empowerment evaluation*, all five qualities must be present to some extent, including liberation and advocacy. Taking such a position goes a long way toward clarifying the unique contribution of empowerment evaluation and separates it from other participatory, collaborative, stakeholder-involving, and utilization-focused approaches. Let me, then, show how I think the failure to make this distinction muddies the waters in the other cases in the book.

Millett's chapter on evaluation at the W.K.Kellogg Foundation, where he serves as Director of Evaluation, shows how some of the principles of empowerment evaluation can be incorporated into an organization's evaluation philosophy and guidelines. Kellogg's approach builds on commitments to use accumulated wisdom ("We know better than we do"), help people help themselves, work collaboratively, and support learning by doing. Millett is particularly insightful and honest about the tensions that emerge in applying these principles in a mainstream organization, that is, in trying to align "the need to 'improve' with the need to 'prove.' To date, we have not been as successful with this as we would like" (p.70). In seeking balance between supporting self-reflective learning at the grantee level (improvement-oriented evaluation) while seeking outcomes data for Foundation decision makers (proving overall effectiveness for accountability), one central lesson has emerged: "Useful judgments — that is, evaluation outcomes — are more often than not determined by who is making the judgments" (p. 71). The Kellogg Foundation, under Millett's leadership, is trying to balance the needs and interests of divergent stakeholder groups from the field level where grant monies are used, among program staff at the Foundation, and all the way up to the Board of Trustees. True empowerment evaluation would make the self-identified needs of grassroots participants the driving force of evaluation, but those who provide the funds also have a stake. Therein resides the tension, more easily named than resolved. On a continuum of empowerment priority (from the high end where all five of Fetterman's defining characteristics are manifest in meaningful ways to the low end where emphasis on defining characteristics is weak or absent), I believe that Kellogg's approach would be more accurately described as participatory and collaborative than empowering. Kellogg's evaluation effort is driven primarily by the

"illumination" agenda (lessons learned, wisdom accumulated and applied) and, as I argued earlier, that agenda is neither unique to nor a defining characteristic of empowerment evaluation. It is true that the Evaluation Handbook of the W.K.Kellogg Foundation includes among its principles that evaluation ought to "empower people," but this is defined more in terms of capacity-building (skills, knowledge, and perspectives acquired during the process) than with direct liberation. For me, Millett's chapter illuminates the tensions involved in trying to make empowerment evaluation a meaningful part of the agenda of a huge institution.

The fourth chapter by Keller opens thus: "This is not your 'normal' case study" (p. 79). What makes it non-normal is that engaging in empowerment evaluation "precipitated a personal odyssey of empowerment" in which the author experienced "illumination and liberation" (p. 79). Keller then describes how, in her job as a state auditor in Texas, she worked with two state agencies to move from compliance-based audit-style evaluations to improvement-oriented agency self-evaluations. What is striking to me about these two cases is that, absent the overlay of empowerment subtitles in the chapter (e.g., "Illumination and Advocacy," "Liberation...: Imparting Ownership"), the actual cases are exemplars of good, old-fashioned (by now) stakeholder-involving, utilization-focused evaluation. These are stories of an evaluator working with agency managers and staff to answer questions they care about. It's a stretch to see these state agency managers as oppressed peoples, disempowered by state regulations, whose lives are dramatically changed by an empowerment process. Rather, this is a story of agency managers figuring out that, given the outcomes mania loose in the land and new legislative requirements for performance reporting, it was in their own self-interest to engage in evaluation to support their budget requests. I don't doubt that, for Keller, the change from an audit mentality of compliance monitoring to the role of improvement-oriented facilitator was dramatic for agency staff and liberating for her, but to call this empowerment evaluation is to diminish the real contribution and niche of empowerment evaluation as an advocacy approach supporting self-determination for the disempowered. The focus of the examples in this case are not of that nature. Keller reports: "We spoke of cost savings and preserving funds to upper management. We spoke of identifying what works and improving processes to lower management and workers" (p. 96). This is classic management-focused, improvement-oriented evaluation in government *a la* Wholey (see Shadish et al, 1991, pp. 225–269). It's effective, important, and useful. What it's not, in any meaningful sense, is empowerment evaluation.

The fifth chapter by Gmez and Goldstein describes an HIV Prevention Evaluation Initiative. The chapter is framed as resolving a problem in bicultural understanding to overcome historical barriers between university researchers and community providers. The MORT syndrome (money, ownership, rigor, time), an acronym describing a "death" in communication, draws attention to central issues in collaborative evaluation efforts (pp. 103-104). This is a splendid case example of collaborative evaluation that greatly enhanced communications between evaluators and community agency staff. A great deal of shared learning is documented, including staff learning evaluation processes and skills. The language of collaboration is appropriate to the case data presented. The language of empowerment, sprinkled in here and there, reads as an overlay and add-on, introduced to fit the book's title, but short on genuine participant liberation. Evaluators interested in collaboration, but leery about the political overtones of empowerment language, can learn much from this chapter by simply ignoring the occasional and largely superfluous insertions of the word "empowerment."

My reaction to the eighth chapter, co-authored by eleven co-researchers with Fawcett as senior author, is similar. It aims to illustrate the steps in an "empowerment evaluation" pro-

cess through two case examples: a community coalition for prevention of adolescent pregnancy and substance abuse in three Kansas communities, and the Jicarilla Apache Tribal Partnership for Prevention of Substance Abuse in north-central New Mexico. The six elements of the process of empowerment evaluation are: "(a) assessing community concerns and resources; (b) setting a mission and objectives; (c) developing strategies and action plans; (d) monitoring process and outcome; (e) communicating information to relevant audiences; and (f) promoting adaptation, renewal, and institutionalization" (p. 169). There's nothing new or particularly empowering about these elements. In the cases presented, most of the work was done by external evaluators with stakeholder input and involvement. Evidence that project leadership made formative use of the evaluations for program development and change are offered, but whether these uses were particularly "empowering" beyond what occurs in other participatory and collaborative evaluation processes remains in doubt. For example, "data helped redirect the energies of coalition staff, leadership, and members away from service provision and toward creating community change. This adaptation was important and empowering because prevention initiatives may be more effective as catalysts for change than as new service agencies" (p. 178). I interpret this as a conclusion about how to be effective, not as evidence of empowerment. I emphasize this distinction because on such interpretations hinge whether empowerment evaluation comes to be viewed as different in kind or only in degree from other stakeholder-involving forms of evaluation. To their credit, the authors note that "the ambiguousness of the construct of `empowerment evaluation' may make it difficult to detect good practice. Empowerment remains a vague concept, referring to both a process and a goal" (p. 179).

The chapter by Yin, Kaftarian, and Jacobs on "Empowerment Evaluation at Federal and Local Levels" exemplifies this conceptual vagueness and ambiguity. The case example is a multisite, multilevel, and multicomponent evaluation of the Center for Substance Abuse Prevention Community Partnership Demonstration Grant Program for 251 local partnerships. The context for this chapter is Stufflebeam's attack on empowerment evaluation (1994) in which he accused it of abrogating the Joint Committee's professional standards for evaluation and undermining the credibility of the entire profession. Fetterman's rejoinder (1995) is also part of the context. Given the Stufflebeam-Fetterman debate, Yin, Kaftarian, and Jacobs undertake a kind of self-reflective meta-evaluation to show how the empowerment evaluation process in this case adhered to the standards. What they demonstrate instead, in my judgment, is that they implemented an exemplary utilization-focused evaluation process rather than an empowerment evaluation. They acknowledge as much when they observe: "The possible outcomes of empowerment evaluation appear initially to resemble the concepts related to evaluation utilization (e.g., Patton, 1986; Weiss, 1978) in which an overall impact — 'utilization' — may be reflected by multiple outcomes (e.g., dissemination, communication, and utilization" (p. 201). That initial appearance of resemblance, far from being diminished, is affirmed by their evidence. They document a collaborative process of evaluation design involving many primary users. They evaluate the evaluation outcomes, what they call "empowerment outcomes," by examining the (a) dissemination of the customized evaluation framework, (b) communications to others regarding the framework, (c) use of the evaluation framework in a variety of activities, and (d) the validity of evaluation results, what they call "empowering ideas" and what the utilization literature calls conceptual or illuminative use. In summarizing "empowerment impact," they conclude: "First, all relevant parties to the original empowerment evaluation process (federal agency staff, local evaluators, and cross-site evaluators) have been involved in some action or validating outcome. Second, these involvements have

occurred in the spirit of 'co-ownership'" (p. 203). In other words, the collaboration between federal and local evaluators led to a framework and findings that were used. The language of empowerment adds nothing in this instance, other than whatever political benefit is derived from using such language in a community-oriented program. No evidence is provided that program participants, as a direct result of the evaluation process, were in any way empowered, emancipated, liberated, enfranchised, or engaged in self-determination. Local evaluators and program staff, usually alienated by top-down and standardized federal evaluation requirements, did apparently feel co-ownership of the evaluation and, as a result, use was greatly enhanced. But, in that regard, this case exemplifies the processes and principles of utilization-focused evaluation more than it does the defining participant liberation and advocacy characteristics that, I have argued here, distinguish empowerment evaluation.

Empowerment Evaluation and Capacity-Building

The cases in this book, wherever they fall along continua of participation, collaboration, and empowerment (separate dimensions conceptually), share a commitment to capacity-building. While not the defining characteristic of empowerment evaluation, it can be thought of as critical. In other words, I infer from these cases that commitment to capacity-building might be thought of as a necessary but not sufficient condition for an evaluation to be considered empowering.

Not all references to "participatory" or "collaborative" evaluation make explicit the link to participant learning and capacity-building that empowerment evaluation emphasizes. For example, Levin (1993) distinguished three purposes for collaborative research: (1) the pragmatic purpose of increasing use, (2) the philosophical or methodological purpose of grounding data in practitioner's perspectives, and (3) the political purpose of mobilizing for social action. A fourth purpose, emphasized throughout this book, is teaching evaluation logic and skills as a way of building capacity for ongoing self-assessment. In modeling terms, such skills are seen as enhancing the capacity for self-determination. Indeed, empowerment evaluation makes learning to do self-assessment a primary agenda, sometimes *the* primary outcome of the evaluation.

The chapter by Mayer on "Building Community Capacity With Evaluation Activities That Empower" epitomizes the focus on capacity building. Rainbow Research in Minneapolis, which Mayer directs, uses the language of empowerment sparingly, choosing more often to emphasize "building community capacity."

> We believe that evaluation can assist capacity building, especially when it gives intended beneficiaries of a project the opportunity to get involved in its evaluation. Evaluation that allows the project's intended beneficiaries to get involved in the evaluation process in ways that give them more commitment, resources, and skills could be said to fit the description of "empowerment evaluation" (p.333).

Mayer offers a number of impressive examples of capacity-building evaluations conducted by Rainbow Research, most of which fit the mold of participatory and collaborative evaluation without the added liberation emphasis of empowerment evaluation. These cases and the capacity-building framework emphasize identifying strengths and assets as well as needs or weaknesses, creating partnerships of "codiscovery," listening to and including those not normally included ("the more marginalized of the community"), and connecting together people who can use evaluation information. In addition, Mayer offers the most detailed and

best example in the book of an evaluation where the agenda and language of empowerment evaluation coalesce quite perfectly with the agenda and language of the program.

To support evaluation of a program that worked with mothers trying to get off welfare, they co-created and gathered data through a "Women's Empowerment Group Acts of Empowerment Evaluation Logbook" that is included as a chapter appendix. The label for the instrument is derived from and congruent with the program's name for the women's support groups. The instrument describes "acts of empowerment" thusly:

> These are actions a group member might take — as a result of her participation in the empowerment group — that indicate self-sufficiency: a willingness or capacity to take control of her life, improve her life (p. 340).

Categories for "acts of empowerment" and a framework for recording them make up the actual instrument. Reviewing the instrument, it is clear that the act of logging empowerment acts and then sharing these with other participants would support outcomes of self-determination. Mayer's example illustrates how empowerment evaluation is especially appropriate where the goals of the program include helping participants become more self-sufficient and personally effective. In a similar vein, Weiss and Greene (1992) have shown how "empowerment partnerships" between evaluators and program staff were particularly appropriate in the family support movement because that movement explicitly emphasized participant and community empowerment.

Empowerment Evaluation as Process Use

In its focus on capacity-building, empowerment evaluation is one among many types of evaluation that emphasize "process use" (Patton, 1997b). The effectiveness of participatory strategies for increasing use of findings is well-documented (Patton, 1997c). The new emphasis on *process use* directs attention to the impacts of participating in an evaluation *in addition to* generating findings, impacts that derive from using the logic, employing the reasoning, and being guided by the values that undergird the evaluation profession (e.g., Fournier, 1995; House, 1980). These impacts include enhanced mutual understanding among those participating in the evaluation, support and reinforcement for the program intervention, program, organizational, and community development (e.g., developmental evaluation, Patton, 1994), and increased participant engagement in and ownership of program and evaluation processes. Participation and collaboration can lead to a long term commitment to use evaluation logic and techniques thereby building a culture of learning among those involved.

The cases in *Empowerment Evaluation* document the ways in which participants in an evaluation can come to value both the processes and findings of evaluation. A theme running throughout the book is that learning to see the world as an evaluator sees it often has a lasting impact on those who participate in an evaluation — an impact that can be greater and last longer than the findings that result from that same evaluation, especially where those involved can apply that learning to future planning and evaluation situations. This capacity-building emphasis of empowerment evaluation has helped illuminate the nature and importance of process use for those who study and theorize about utilization of evaluation. With regard to practice implications when negotiating an evaluation, making explicit to stakeholders the possibility of enhancing process use, i.e., building capacity, adds to the menu of potential evaluation benefits. Framing potential process use and capacity-building options in the language of empowerment evaluation appears most appropriate when the program itself has

explicit empowerment goals, as in the cases in the book of the Accelerated Schools Project, the South Central Los Angeles Community Coalition effort, and the Curriculum-Based Support Groups effort. These cases suggest that empowerment evaluation will work best where primary intended users want a process that purposefully and intentionally supports explicit empowerment goals by training and coaching program staff and participants in evaluation logic and techniques so that they can take charge of their own evaluation processes.

However, the link between learning evaluation skills and achieving self-determination is far from direct. Nothing in these cases supports a direct connection between liberation as an outcome and learning evaluation skills as a process. The two may be mutually reinforcing but not determinative. The development and use of "The Plan Quality Index" described by Butterfoss, Goodman, Wandersman, Valois, and Chinman in chapter 14 provides an excellent example of a participatory and collaborative process as well as an instrument for facilitating the learning and application of evaluation logic to improve intervention plans, but evidence is lacking for having achieved liberation outcomes. As an example of how learning the logic of evaluation can be a primary and useful outcome of collaborative evaluation, the case is exemplary — and quite open in presenting the challenges encountered. As an example of empowerment evaluation in its full sense, this case illustrates how much more is needed beyond a mutually developed instrument. Here, then, is another case where the authors embrace the participatory, collaborative, and even capacity-building dimensions of empowerment, but downplay or ignore its more political activist dimension. The focus is on increased effectiveness, not changed power relations.

Advocacy: Changed Roles and Evaluator Credibility

Advocacy has two meanings in empowerment evaluation: (1) enhancing the capacity of disempowered groups to advocate on behalf of themselves and (2) evaluators advocating on behalf of disempowered groups. The first is considerably less controversial among evaluators than the second.

All participatory, collaborative, and utilization-focused styles of evaluation change the role of the evaluator from the traditional lone judge of merit or worth to a facilitator of judgments by others involved in the process, sometimes in addition to the evaluator's judgment and sometimes without independent judgment by the evaluator (Patton, 1997a). These forms of evaluation build the capacity of those involved to make their own judgments and use findings. Beyond facilitating and coaching, however, empowerment evaluation, in its purer forms, changes the role of the evaluator to that of direct advocate and social change agent. Certainly, evaluators undertaking such an approach need to be comfortable with and committed to it, and such an activist agenda must be explicitly recognized by, negotiated with, and formally approved by primary stakeholders and intended users, including funders.

Since using evaluation to mobilize for social action and support social justice are options on the menu of empowerment evaluation, what does the book teach us about the implications and consequences when evaluators assume such roles directly? Very little. In introducing the advocacy dimension (pp. 13-14), Fetterman cites examples of evaluators who have championed their findings after an evaluation was completed. He offers examples of Congressional testimony, participating in policy discussions, and writing editorials. In none of these examples is it clear that the evaluators became advocates for a specific program or program participants rather than advocates for use of findings. Even advocat-

ing for a generic type of program, e.g., needle exchange to prevent HIV transmission or the need for a program to serve high school drop-outs is different from advocating support for a specific program. *These are different things.* Those of us interested in evaluation use have a long history of advocating follow-up engagement to champion *findings.* The cursory examples offered in Fetterman's introduction cloud the issue, at least for me, rather than illuminating new territory and new roles. I can't tell enough about what was advocated to whom to know what the evaluator's role was. Moreover, the book's detailed case study chapters provide no good examples of evaluators playing the direct advocacy role or, there-fore, of the consequences and implications of such a role. In the case exemplars, evalua-tors provide technical assistance and play the roles of coaches and facilitators (so participants can advocate for themselves), but do not get involved in advocating directly for the program and its participants beyond advocating for and supporting their involvement in and ownership of the evaluation. Because the proposed role of evaluator as direct pro-gram advocate is the most controversial and problematic dimension of empowerment eval-uation, it's disappointing that no cases are presented that really illuminate and evaluate that role. Doing so in the future will be important in clarifying the nature and consequences of empowerment evaluation.

This point deserves elaboration. The idea and language of advocacy is not new to evaluation (e.g., Sonnichsen, 1988), but in the past the emphasis has been on advocating that evaluations be conducted and that findings be used, that is, championing use. Advocat-ing for evaluation, even self-evaluation and capacity-building approaches, is quite different from advocating for a program or cause. Chelimsky (1995, 1996, 1997), whose advocacy for use and political sophistication about facilitating use are second to none, has asserted consistently that the evaluator's most precious resource in promoting use is credibility grounded in pre-study neutrality about the effectiveness of the program or cause being eval-uated. A balanced presentation, for example, analyzing both strengths and weaknesses, tends to be more credible than one that is primarily negative or positive. The book under review provides a case in point. For me, the chapter on empowerment evaluation in South Central Los Angeles would be more credible and illuminating if it acknowledged difficul-ties and imperfections in the process, as most other cases in the book do. The advocacy tone of the case — working to convince us of the merits and community embrace of empowerment evaluation — undermines the argument that self-reflection facilitates looking at both strengths and weaknesses. In the case data, everyone respects everyone else. The community embraces the evaluator. Everyone learns from each other. "The community accurately appraises the reality of the mitigating conditions levied against it and the multi-tude of levels at which change must occur" (p.138). It's an inspiring case. Indeed, it's an almost *unbelievable* case, one that stands in strong contrast to most other cases that describe challenges and difficulties encountered. Therein lies the problem with advocacy-oriented evaluation reporting. When the data appear to be carefully shaped to support a predeter-mined position, knowledgeable readers will be skeptical. That, to me, makes the case less believable, not in the sense that the authors are not telling the truth, but that *their* version and understanding of the truth is so shaped by their advocacy that the truth they present loses credibility for lack of balance.

I am not raising an ethical issue here, as Stufflebeam (1994) and Scriven (1991, 1993) have, about whether evaluators ought to be program advocates, though their concerns deserve careful attention. Rather, I am raising the instrumental question of whether evaluators can act effectively as program advocates over the long run — and still be credible, and therefore use-

ful, as evaluators. Program advocacy may undercut the evaluator's credibility, not only with external audiences, but even with program participants who may sense that the evaluator has been co-opted and that the very capacity to do meaningful reality-testing that the evaluator was supposed to be teaching is not being practiced by the evaluator.

This does not mean, on the other hand, that the evaluator cannot openly support the program's goals and share the program's values. An evaluator need not be neutral about alcohol and drug abuse, or any other program goal. What an evaluator is commonly expected to be neutral about, short of empirical evidence, is the effectiveness of a particular program in attaining those goals and acting in accordance with shared values. Crossing the line between advocating use of findings to advocating for participants and the program may undercut the evaluator's greatest utility to an effective program, his or her credibility. This may be, in particular instances, a matter of subtlety and nuance, but the issue is central to the debate about the legitimacy of empowerment evaluation. From an instrumental perspective, it's an empirical issue about the long-term effects and effectiveness of program advocacy by evaluators — an issue this book does little to address or illuminate.

THE LANGUAGE OF EMPOWERMENT EVALUATION

Reading this book raises over and over again questions about language. Sorting out the differences between participatory, collaborative, and empowerment evaluation is partly a matter of language appropriateness and congruence between a program's culture and the rhetoric of empowerment evaluation. It's also partly a matter of labeling alternatives in a way that stakeholders can make meaningful choices about what approach to evaluation they prefer. One theme that stands out in these cases is that evaluation facilitators need heightened sensitivity to the language preferences of those involved. For example, Linney and Wandersman, in describing the *Prevention Plus III* materials they developed in support of community-based substance abuse prevention evaluations (chapter 12) abandoned the term "evaluation" in favor of the term "assessment" based on feedback that the latter word was viewed as more appropriate (p. 272). Nor does the language of empowerment appear in the workbook materials they developed, though, for purposes of this book, they use that language occasionally in the case presentation to talk about the materials. They've increased community access to potentially useful evaluation tools to meet funding mandates for evaluation. The workbook excerpts they present are standard fare. Use of those materials to support a political agenda of self-determination beyond the evaluation itself remains undocumented.

Nuances of language laced with dynamite also exist within the lexicon of empowerment. One doesn't get very far into the political culture of empowerment advocacy before encountering the question of whether one person can empower another, for to so claim is elitist and disempowering of the self-empowering capacity of each individual. The philosophy of empowerment evaluation is that no one can empower another; people empower themselves. But, in practice, the language can be tricky. Consider, for example, the effort at organizing the community in South Central Los Angeles. The chapter opens with the strong assertion that "it is understood that evaluation is being conducted within a community that is already self-determining. In this example, the evaluation process does not create self-determination within the community but facilitates a preexisting sense of self-determination within a community that has its own philosophy, style, strategies, and history" (p. 124). Two pages later, however, the community is described as "disenfranchised." Evaluators who embrace the admonition to

reject "the paternalistic and patronizing quality characteristic of many well-intentioned traditional evaluation approaches" (p. 130), an admonition worthy of contemplation, will need to exercise care in using appropriate language and understanding the nuances of the difference between being self-determining, yet disenfranchised, and how to know when a community is already empowered or merely on the path toward empowerment.

Political Correctness and Empowerment Evaluation

In the closing plenary session of the 1996 national conference of the American Evaluation Association, President Len Bickman presented sample survey data from Association members in which 50% described themselves as "liberal" and 50% as "moderate." Reasons for the absence of self-described "conservatives" would make for interesting speculation, but for my purposes here, the data raise the question of how much empowerment evaluation might be subject to attack as a manifestation of political correctness. Certainly the language of empowerment is not unique to the left; witness Jack Kemp's advocacy of inner city 'empowerment zones' for economic development. But, for conservatives, empowerment means self-sufficiency (off welfare) more than self-determination (political activism).

How, then, might those who worry about the mind-numbing effects of political correctness read this book? What they'd find, overall, is a great deal of attention to methodological rigor and genuine concern about whether self-evaluation involves a loss of rigor. The cases provide ample evidence that self-assessment can be relatively rigorous within the confines of producing context-specific, improvement-oriented learnings, that is, evaluation by participants and/or staff *for* participants and/or staff. Trade-offs between rigor and participation arise as a common tension in design discussions. Evaluations based entirely on self-assessment will likely give rise to credibility concerns with external audiences, even if carefully done.

A related challenge arises when operating in the political correctness environment of empowerment rhetoric that constrains how things get talked about. Causal explanations can sound pre-packaged. Successes get attributed to community engagement and empowerment; failures are due to external oppression. Consider this: "The ever-present political, economic, and racial climate within which the community constantly seeks to assert itself is never underestimated or forgotten. This helps to avoid misplacing attributions of failure on the community for not changing forces that are formidable and require concerted, oftentimes long-term efforts and systematic strategies to bring about change" (pp. 130–131).

This problem of causal attribution comes to mind from a recent personal experience in which feedback to a community-based organization about failure of steering committee members to show up at meetings, falsified reports by staff, and embezzlement by the project director were initially explained away by external oppression and my failure to appreciate the true victim status of perpetrators. Eventually, as the group engaged the findings, participants took ownership and responsibility, as the empowerment approach predicts, but getting there took careful facilitation. Sensitivity to historical sociopolitical oppression does not mean ignoring or explaining away personal and professional responsibility. Evaluators who cannot facilitate real reality-testing about genuine successes as well as culpable failures will not be nurturing empowerment in the long run any more than do educators who graduate students who cannot read.

FINAL COMMENTS

The cases in this book provide sufficient description for the reader to understand pretty much what actually occurred in each setting. Providing adequate descriptive information in an analytic framework, all within the page constraints of a single chapter, is a challenge in case study reporting. The quality of the case presentations, overall, strikes me as strong.

The book does not really deliver on its subtitle: "Knowledge and Tools for Self-Assessment & Accountability." The actual examples of tools are few and not very generalizable. Moreover, the process of developing tools within a local context has to be a part of how capacity is built to do evaluation. In that sense, generalizable tools, even if possible, would scarcely be desirable. The weakest "empowerment" examples in the book are the projects that involved disseminating tools to local partners.

Nor does the book do much to address the accountability theme in the subtitle. Millett sheds the most light on the conflicting goals of improving versus proving. Fetterman devotes one paragraph to the tensions between accountability and empowerment in the closing chapter (and that's the only reference to accountability in the index). His discussion is mostly a call for more study of and attention to the issues involved. No real guidance is offered for negotiating the tension between self-assessment and external accountability. The philosophy comes down to this: the highest form of accountability is self-accountability. This perspective, however, is largely implicit, and is not articulated in a way that its rationale could be communicated convincingly to most evaluation funders. None of this would be a problem were accountability not highlighted in the subtitle.

The cases and discussion, as I have noted throughout this review, confuse distinctions between participatory, collaborative, and empowerment evaluations. Perhaps the new AEA Topical Interest Group (TIG) on Collaborative, Participatory, and Empowerment Evaluation will sort out the commonalities and differences. The eclectic and inclusive nature of the case examples in the book add to rather than reduce the ambiguousness of the empowerment construct.

In my judgment, the book offers three fairly clear examples where empowerment evaluation criteria are reasonably primary and the rhetoric of empowerment seems to fit: the Accelerated Schools Project (chapter 2); the South Central Los Angeles Community Coalition effort (chapter 6); and the Women's Empowerment Group Acts of Empowerment Logbook (chapter 15). Two others come close, but remain problematic because they are still unfolding: the Nonprofit Women's Services study group (chapter 7), and the Curriculum-Based Support Groups effort (chapter 13). The other cases, from my perspective, illustrate participatory and collaborative approaches, but fall well short of meeting Fetterman's full definition of empowerment evaluation.

While the mix of cases and the ambiguous nature of the empowerment construct may disappoint readers who expect clarity, consistency, and generalizable tools, this collection moves the idea of empowerment evaluation from rhetoric to practice, and invites application to empowerment evaluation of the very evaluation skills so lauded in the book as accessible to nonresearchers (e.g., clear criteria, operational concepts, and specified models of causality). As the practice develops, so will the theory, if the practice proves sustainable and valuable beyond the rhetorical moment of this decade.

New approaches are not born fully developed and clearly defined. Thus, the conceptual ambiguities I have identified are reflective of the early stage of development of empowerment

evaluation. When I first published *Utilization-Focused Evaluation* 20 years ago (1978), I failed to include a rather critical conceptual item: a definition of evaluation utilization. Through interactions with AEA colleagues and responding to criticisms of the first edition, I finally offered a definition of utilization in the second edition (1986), one that could serve as an evaluation goal and accountability criterion: intended use by intended users. The latest edition (1997) examines more fully the implications of that definition and deals with conceptual, ethical, and operational issues that I could not have imagined in 1978. Refining the dimensions of empowerment evaluation, examining how they interact, distinguishing them from related approaches, and testing them in practice offer a full development agenda.

As for controversy about empowerment evaluation, Oscar Wilde in "The Critic as Artist," mused: "An idea that is not dangerous is unworthy of being called an idea at all." This book shows empowerment evaluation to be an idea in tune with the politics of our times. Therein lies its strength, weakness, power, and potential danger.

REFERENCES

Chelimsky, E. (1995). *The political environment of evaluation and what it means for the development of the field.* American Evaluation Association Presidential Address, Vancouver, November, 1995. Evaluation Practice. 16(3), 215–225.

Chelimsky, E. (1996). *From incrementalism to ideology and back: Can producers of policy information adjust to the full spectrum of political climates?* Distinguished public policy lecture series (pp. 39–63). Evanston, IL, Northwestern University Center for Urban Affairs and Policy Research.

Chelimsky, E. (1997). The coming transformations in evaluation. In E. Chelimsky & W. Shadish (Eds)., *Evaluation for the 21st century.* Thousand Oaks, CA, Sage Publications.

Cousins, J. B., & Earl, L. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14,* 397–418.

Cousins, J. B., & Earl, L. (1995). The Case for Participatory Evaluation: Theory, Research, Practice. *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 3–18). London, Falmer Press.

Fetterman, D. (1995). In response to Dr. Daniel Stufflebeam. *Evaluation Practice, 16(2),* 177–197.

Forester, J. (Ed)., (1985). *Critical theory and public life.* Cambridge, MA, MIT Press.

Freire, P. (1970). *Pedagogy of the oppressed.* New York, Seabury Press.

Fournier, D. M. (Ed)., (1995). *Reasoning in evaluation: Inferential links and leaps. New Directions for Evaluation.* Volume 68. San Francisco, Jossey-Bass.

Habermas, J. (1984). *The theory of communicative action.* Volume I. Boston, MA, Beacon Press.

Habermas, J. (1987). *The theory of communicative action.* Volume II. Boston, MA, Beacon Press.

Harding, S. (Ed)., (1987). *Feminism and methodology.* Milton Keynes, Open University Press.

House, E. R. (1980). *Evaluating with validity.* Beverly Hills, CA, Sage Publications.

King, J. A. (1995). Involving practitioners in evaluation studies: How viable is collaborative evaluation in schools. In J. B. Cousins & L. E. (Eds)., (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning.* (pp. 86–102). London, Falmer Press.

Levin, B. (1993). Collaborative research in and with organizations. *Qualitative Studies in Education,* 6(4), 331–340.

Maguire, P. (1987). *Doing participatory research: A feminist approach.* Amherst, MA, Center for International Education.

Patton, M. Q. (1986). *Utilization-focused evaluation.* 2nd edition Newbury Park, CA, Sage Publications.

Patton, M. Q. (1994, October). Developmental evaluation. *Evaluation Practice, 15(3),* 311–320.

Patton, M. Q. (1997a). Being active-reactive-adaptive: Evaluator roles, situational responsiveness, and strategic contingency thinking. *Utilization-focused evaluation: The new century text,* 3rd edition (pp. 117–145). Thousand Oaks, CA, Sage Publications.

Patton, M. Q. (1997b). Intended process uses: Impacts of evaluation thinking and experiences. *Utilization-focused evaluation: The new century text,* 3rd edition (pp. 87–113). Thousand Oaks, CA, Sage Publications.

Patton, M. Q. (1997c). Intended uses of findings. *Utilization- focused evaluation: The new century text,* 3rd edition (pp. 63–85). Thousand Oaks, CA, Sage Publications.

Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. McLaughlin and D. C. Phillips (Eds.), *Evaluation and education: At quarter century.* 90th Yearbook of the National Society for the Study of Education (pp. 18–64). Chicago: University of Chicago Press.

Shadish, W. R., Jr., Cook, T. D., & Leviton, L. C. (Eds)., (1991). *Joseph S. Wholey: Evaluation for program improvement. Foundations of program evaluation: Theories of practice* (pp. 225–269). Newbury Park, CA, Sage Publications.

Sonnichsen, R. C. 1988. Advocacy evaluation: A model for internal evaluation offices. *Evaluation and Program Planning. 11*(2), 141–148.

Stufflebeam, D. (1994, October). Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go. *Evaluation Practice, 15*(3), 321–338.

Vanderplaat, M. (1995). Beyond technique: Issues in evaluating for empowerment. *Evaluation, 1*(1), 81–96.

Wadsworth, Y. (1993). *What Is participatory action research?* Melbourne, Australia, Action Research Issues Association, Inc.

Weiss, C. H. (1978). Improving the linkage between social research and public policy. In L. E. Lynn, Jr. (Ed)., *Knowledge and policy: The uncertain Connection* (pp. 23–81). Washington, D. C., National Academy of Sciences.

Weiss, H. B., & Greene, J. C. (1992, February/May). An empowerment partnership for family support and education programs and evaluations. *Family Science Review, 5*(1,2), 145–163.

Whyte, W. F. (Ed)., (1991). *Participatory action research.* Newbury Park, CA, Sage Publications.