

## A FRAMEWORK FOR CHARACTERIZING THE PRACTICE OF EVALUATION, WITH APPLICATION TO EMPOWERMENT EVALUATION

Nick L. Smith  
Syracuse University  
Syracuse, New York

**Abstract:** As the evaluation profession experiences a widening diversity of theory, method, and practice, there is a growing need for some means of characterizing significant similarities and differences in evaluation approaches. To that end, a framework is presented which consists of six aspects: context, purpose and social role, phenomena of interest, procedural rules, methods of justification, and sanctions. The utility of the framework is illustrated through its application to a current debate concerning the theory and practice of Empowerment Evaluation.

**Résumé:** À une époque où l'évaluation professionnelle élargit la diversité de ses théories, méthodes et pratiques, il est de plus en plus nécessaire de se doter de moyens de caractériser d'importantes similitudes et différences des approches d'évaluation. À cette fin, on présente un cadre comportant six aspects: contexte, objet et rôle social, phénomènes d'intérêt, règles de procédure, méthodes de justification, et sanctions. L'utilité du cadre est illustrée en l'appliquant à un débat sur la théorie et la pratique de l'évaluation de l'habilitation.

When Empowerment Evaluation was dramatically introduced to the evaluation profession through Fetterman's (1994) presidential address to the American Evaluation Association, the reaction was immediate and heated. Stufflebeam (1994) published a critique of this address, to which Fetterman (1995) responded. Fetterman, Kaftarian and Wandersman (1996) then published *Empowerment Evaluation: Knowledge and Tools for Self-assessment and Accountability*. A chapter on empowerment evaluation by Fetterman (1997a) also appeared in a book by Chelimsky and Shadish. The Fetterman et al. (1996) book has been reviewed by Brown (1997), Patton (1997a), Scriven (1997a), Sechrest (1997), and Wild (1997). Fetterman has published responses to the reviews by Sechrest

(Fetterman, 1997c), and Patton and Scriven (Fetterman, 1997b), who have each issued rejoinders (Patton, 1997b; Scriven, 1997b). “Empowerment Evaluation” is capitalized throughout this article, then, to reflect this more or less discrete approach to evaluation. This is done to clarify the difference between design intent and study result: an evaluation that does not follow the Empowerment Evaluation approach may nevertheless empower someone (an empowering evaluation), just as an Empowerment Evaluation may fail to empower.

It is difficult to construct a clear picture of the varied mix of claims, criticisms, and rebuttals generated by the proposal of Empowerment Evaluation. In order to understand the significance of this debate, it is necessary to first appreciate its emergence within an increasingly diverse evaluation scene, and second to see the criticisms as interrelated issues within a broad framework of evaluation theory and practice. To characterize this Empowerment Evaluation debate, I will first situate it within the growing diversity of evaluation practice, next suggest a framework for examining the various issues raised, and then examine the debate within that framework.

## DIVERSITY OF EVALUATION PRACTICE

Darwin’s Finches of the Gala’pagos Islands —

gradually the descendants of those first finches evolved into different forms. Some came to eat seeds of one sort, some of another; still others came to eat insects. For each way of life a particular species would develop a particular beak, a particular size, a particular scheme of organization. (Asimov, 1972, p. 461)

At the time of Darwin’s visit on the H.M.S. Beagle, there were over 14 species of finches on the Gala’pagos Islands. If we could agree on how to count them, we would probably find more than 14 species of evaluation currently being practised. Darwin’s observation of the finches played a key role in his development of the theory of evolution. On a more modest scale, perhaps our observations on the variety of evaluation practices can give rise to an improved theory of evaluation practice.

Evaluators have observed the proliferation of evaluation approaches for several decades: from the early development of numerous models of evaluation in the 1960s and 1970s, through various attempts

at classificatory schemes in the 1970s and 1980s, to the integrative “multiples” strategies of Cook, Scriven, and others in the late 1980s, to the meta-theories of Scriven and Shadish, Cook, and Leviton in the early 1990s (see Smith, 1994, for a review of this history).

Diversity of viewpoint can result in serious problems, such as when the very definition of the field of evaluation itself is at issue; for example, contrast Rossi and Wright’s early definition of evaluation as a form of experimental social science: “There is almost universal agreement among evaluation researchers that the randomized controlled experiment is the ideal model for evaluating the effectiveness of a public policy” (Rossi & Wright, 1977, p. 13) with Eisner’s definition of evaluation as a form of a-scientific qualitative inquiry: “... qualitative inquiry has no obligation — moral, epistemological, or otherwise — to be scientific in character” (Eisner, 1991, p. 180). These are quite different kinds of birds!

Although such diversity could be seen as contributing to the robustness and scope of the evaluation profession, in related disciplines it is often seen as a problematic fragmentation that could eventually result in the dissolution of the professional enterprise. For example, Yanchar and Slife (1997) summarize why fragmentation in psychology is considered such a problem:

a fragmented discipline does not have common standards of evaluation ... rather, each discourse community within the discipline adopts indigenous rules for the adjudication of knowledge claims. This, in turn, creates confusion as psychologists within a particular discourse community are unable to evaluate research produced in other discourse communities ... As a result, communities make no contribution to the discipline as a whole. There is little or no accumulation in knowledge, and little occurs in the way of genuine scientific progress ... (p. 237)

That this fragmentation represents more than just differences of methodological opinion among academics is evident in Rice’s (1997) recent statement that:

American psychology is viewed as an academic profession that began to develop applied or practitioner interests and that, driven by economic and social forces, has undergone differentiation to the point where a complete

division into 2 professions is a definite possibility.  
(p. 1173)

Rice argues that forces promoting a split within a profession include extreme differentiation of practice due to rapid growth of the profession, increased specialization, and shifting patterns of work settings — conditions currently present in professional evaluation.

Although the symptoms of fragmentation in psychology have their counterparts in evaluation, it may, of course, be alarmist and premature to predict the dissolution of the evaluation profession due to internal dissension. Diverse intellectual contributions have improved evaluation thought and enriched its practice. Indeed, Cook has stressed the benefits of debates about evaluation theory, even expressing the concern that “without active theoretical debates about core issues in evaluation that transcend method, the field may not grow, even if it does not atrophy” (1997, p. 47). At times, however, disagreements among writers have become divisive and counterproductive. Cook warns of the dangers of polarized positions, as in past arguments of qualitative versus quantitative methods, instrumental versus enlightenment evaluation use, and randomized experiments versus quasi-experiments. “But when one examines all these cases, it is immediately apparent that many more options could have been examined, and that the use of disparaging contrast to gain rhetorical advantage was not needed” (1997, p. 43).

How can the profession grasp the diversity of these evaluation positions, systematically clarify the nature of the claims being made, and identify other options that might be examined? In the present case, how can one understand the claims and criticisms of Empowerment Evaluation which have emerged from the epistemological and methodological arguments surrounding the so-called “qualitative/quantitative debate” as one of several stakeholder-oriented alternatives alongside Participatory and Collaborative Evaluation?

## AN ANALYTIC FRAMEWORK

“The practice of evaluation research can be characterized as the making of justifiable claims concerning the object of study” (Smith, 1987, p. 309).

The defining aspect of any given variety of evaluation is the nature of the claims it supports, including how those claims are developed

and justified. The expanding diversity and hybridization of evaluation practice, however, makes it increasingly difficult to clearly discern the type of claims being made in particular instances. This frequently gives rise to confusion and disagreement about the proper conduct of evaluation.

A useful first step in clarifying the diversity of evaluation practice might be the development of a comprehensive framework with which to compare and contrast fundamental attributes of any evaluation approach. One such framework is proposed here; it is reminiscent of the classificatory schemes of the 1970s and 1980s, but cast at a more meta-theoretic level, reflecting the evaluation theory work of the early 1990s. The primary use of this framework is in creating a common frame of reference for examining diverse evaluation positions and approaches; that is, the purpose of this framework is to clarify the discourse of evaluation practice. After illustrating its use in three prior applications, this analytic framework is then applied to the current debate about Empowerment Evaluation (Fetterman, 1994).

Individual theorists change their evaluation positions as a result of many influences: review of discrepant positions, confrontation with others' interpretations of their work, incorporation of related work, accumulation of personal experience with their own approaches, interaction with colleagues, and accumulating research (Alkin, 1991). At the level of the profession, however, review of evaluation approaches requires information that is empirical (i.e., experientially-based evidence, not just value-based arguments), public (i.e., information available to all members of the profession as contrasted with personal knowledge from an individual's own evaluation practice), and nomothetic (i.e., information about the approach in a variety of uses and conditions, as contrasted with idiographic knowledge) (Smith, 1979). There have been repeated calls for increased empirical study of evaluation practice (cf. Scriven, 1991; Shadish, Cook, & Leviton, 1991; Smith, 1993; Worthen, 1990), including statements of different types of empirical tests, criteria for reviewing evaluation methods, and specification of ideal conditions for testing new evaluation methods (Smith, 1981b, 1982). Much of this prior work, however, has presupposed that evaluation approaches share a common purpose and social role so that empirical comparisons are not only possible, but meaningful. Such is not always the case.

Developing a new orientation for evaluation, such as shifting from an experimental to naturalistic paradigm, involves changes in conceptual categories and value

structures as well as changes in operational strategies and techniques. Disciplinary experience, not field trial, tests a paradigm ... *A point of view* is the constellation of conceptual, social, value, and philosophical elements which define a discipline. A new point of view would thus redefine the nature of the phenomena of interest and change the purpose of the disciplinary activity; it would constitute a major upheaval in disciplinary activities and thought. For example, to reclassify evaluation as a branch of aesthetics instead of a branch of inquiry would constitute such a change in point of view. Evaluators would no longer be scientists but artists, theologians, or philosophers. (Smith, 1981a, p. 43)

Thus, to encompass the full range of discourse in evaluation, a conceptual tool is needed which includes consideration of not only the procedural rules of an evaluation approach, but also its purpose, social role, and phenomena of interest.

As evaluation experiences a widening diversity of theory, method, and practice, there is a growing need for some means of characterizing significant similarities and differences in approach. A framework consisting of six comparative aspects: context, purpose and social role, phenomena of interest, procedural rules, methods of justification, and sanctions, is offered here as a possible analytic tool.

## Context

Evaluation as a professional activity is delimited in part by the practical conditions of the settings of its practice. To achieve its societal purpose (whatever that may be), evaluation must be contextually relevant and useful. Practitioners need to recognize the contexts of practice that the various evaluation approaches presuppose.

## Purpose and Social Role

Although the technical purpose of evaluation is to assess merit or worth, the social ends to which this activity is put vary dramatically. The societal purpose of some forms of evaluation is to produce knowledge while, for other forms, its purpose is to promote social reform. The modern/post-modern debate in evaluation, for example, is as much about the proper societal role of evaluation as it is about a proper epistemology.

## Phenomena of Interest

Evaluation approaches differ in the phenomena they consider in two ways: the types of phenomena of interest (e.g., physical, biological, psychological, social, economic, etc.) and the assumptions they make about the reality of those phenomena. Evaluation decisions are facilitated when practitioners recognize that one evaluation approach is based on constructivist assumptions about social reality, while another approach makes realistic assumptions about physical or biological phenomena.

## Procedural Rules

The most obvious way in which evaluation approaches differ is in the procedural rules each espouses. The rules employed are closely related to the prior elements of context, social role, and phenomena of interest. It is this aspect that concerns methodology, not only in a narrow sense, but also in terms of how the evaluator interacts with others with respect to status, power, and communication.

## Methods of Justification

All variants of evaluation practice result, more or less, in various types of claims (knowledge claims, value claims, action claims, etc.). The types of claims considered appropriate within a given approach and how those claims are derived are influenced by the societal role, the phenomena of interest, and the procedural rules of the approach. These elements also shape how claims are justified. Since a primary purpose of any evaluation is inquiry, that is, the development and justification of claims, this aspect is a critical component of the reasoning of any evaluation theory and of each evaluation study.

## Sanctions

Sanctions are the consequences of inappropriate conduct under the various approaches. Depending on the approach, sanctions are seen as issues of epistemology (e.g., being wrong), morality (e.g., being unjust), finance (e.g., losing clients), and so on. While these are philosophical issues for theorists, they are practical issues for practitioners: "What are the consequences of making a mistake?"

Examples of the application of this framework illustrate its utility in clarifying the important dimensions along which evaluation practice currently varies.

## PRIOR APPLICATIONS OF THE FRAMEWORK

The use of this framework is simple: when reviewing an evaluation study, approach, meta-evaluation, and so on, review each element in turn, examine their interrelationships, contrast existing alternatives, and consider other possible options. The framework is not a complex analytic device, it simply directs attention to fundamental aspects of evaluation practice.

The framework is merely an item checklist (Scriven, 1991) to aid description in meta-evaluative reviews. The present version of this framework is a developmental “work in progress” to be applied and modified as needed to improve its analytic and heuristic utility.

Three applications of prior versions of this framework have been made. Smith and Chircop (1989) considered the elements of this framework in their analysis of the Weiss-Patton debate about the nature and level of evaluation impact on decision making. To summarize, using the elements of the framework described above — Weiss and Patton had both discounted that their disagreement was due to differences in Context (e.g., Weiss working with national broad-scale evaluations, and Patton working with local, smaller scale studies). Similarly, they agreed that their disagreement did not reflect differences in the Phenomena of Interest (i.e., Weiss focusing on policy evaluation; Patton dealing with program evaluation). Further, Weiss rejected Patton’s suggestions that their disagreement reflected differences in the Purpose and Social Role of evaluation (i.e., Weiss conducted “academic evaluation” emphasizing research, while Patton conducted “service evaluation” emphasizing stakeholder needs.) Smith and Chircop argued that Weiss and Patton had each developed a different set of effective Procedural Rules because the Methods of Justification and Sanctions differed for each of the decision-making communities with which Weiss and Patton worked.

We are suggesting that purposive-rational action is generally sufficient for effective evaluation practice with decision-making communities such as Patton describes, but both purposive-rational action and communicative action are required for effective evaluation practice within a Weiss type of decision-making community. The difference in these two types of statements (technical-rational and communicative) can further be highlighted if we consider what types of truth claims they generate

and what happens if an untrue proposition is generated. A truth claim under purposive-rational action is supported by an empirically true or analytically correct proposition. Punishment is built into the system. Failure to observe the proposition will lead to failure. Here failure is similar to incompetence or lack of efficiency. Patton's (1988) adamant accountability position that evaluators are ultimately responsible for any failures of evaluation to improve programs is thus consistent with our claim that his approach is based on a model of purposive-rational action. If we do not follow Patton's advice, evaluation is "nonmarketable." Truth claims under communicative action are supported by consensually achieved social norms (intersubjectivity). Failure to observe these claims is sanctioned by the conventions of the community. The punishment for being "deviant" is external to the proposition and grounded in society. Thus we see Weiss (1988) judging the failures of evaluation in terms of both relative contribution to the decision-making process and community-defined norms of acceptable practice. (p. 10)

Thus, an early use of this framework demonstrated how the Weiss-Patton debate was fundamentally a discussion of differences in the methods by which their respective client groups justified claims and dealt with violations of those methods (see Smith & Chircop, 1989, for further details).

In a second application of this framework, Smith (1995) illustrated how broad societal enterprises, which he characterized as "games," shape the nature of evaluative inquiry by creating a context for the inquiry, defining a particular purpose and social role for the inquiry, specifying the kinds of phenomena that are inquired into as well as the procedural rules for conducting the inquiry, determining methods for justifying claims, and enforcing sanctions for violating the social enterprise. He provides a detailed example of how competing analyses by the American Psychological Association and the Federal Bureau of Investigation of the 1989 explosion aboard the Naval vessel *Iowa*, resulted in differences in the development and justification of claims even when the same phenomena and evidence were examined. The different contexts, purposes, and social roles of the social science game versus the criminal justice game gave rise to fundamental differences in what each considered to be appropriate types of claims and methods of justification.

The social scientists believe that only probabilistic claims are usually warranted and accept the consequence that many cases may therefore go unresolved. They accept clinical judgments, but only when based on extensive experience and warranted by scientific evidence. Successful practice, as evidenced by client satisfaction, is not sufficient to justify a method unless correlated with scientific evidence, because a client may be satisfied with inferior results. They argue that accumulated case evidence does not substitute for scientific inquiry. They value objective, replicable knowledge and invoke both professional scientific and legal sanctions. (p. 9)

The criminal investigators work under a requirement for unequivocal claims, because some legal action must be taken. They accept expert opinion based on extensive experience and warranted by the courts. Successful practice, as evidenced by independent confirmation through apprehensions and convictions, is sufficient to justify a method. They argue that the scientific approach does not substitute for the lack of practical experience. They value expert opinion of proven utility to clients and invoke both professional and severe legal sanctions. (p. 10)

At a broad societal level, then, this framework can illuminate the ways in which societal games shape the various forms of evaluative practice.

Brandon (1997) has conducted a recent third application of the framework to compare and contrast stakeholder-assisted evaluation and participatory-evaluation approaches. Brandon concludes that stakeholder-assisted evaluation (SAE) and participatory evaluation (PE) vary little in terms of context and phenomena of interest.

The PE and SAE approaches differ considerably, however, on two other aspects: the *justification of claims* made in educational program evaluations and the *procedural rules* followed in evaluations (including evaluation methods and issues of status, power, and communication in evaluator-stakeholder interaction). The differences between the two approaches are particularly marked when examining their procedural rules. SAE procedural rules are different from their PE counterparts because the two approaches differ markedly on

two other aspects of the investigative evaluation framework — *purposes* and *social roles*. Purposes and social roles determine the procedural rules; different purposes and social roles call for different procedures. (p. 3)

Brandon's careful analysis not only clarifies important distinctions between the stakeholder-assisted evaluation and participatory evaluation approaches, but he also uses the framework to illustrate how the three primary procedural rules of stakeholder-assisted evaluation were implemented in four stakeholder-assisted evaluation studies.

## A FRAMEWORK ANALYSIS OF THE EMPOWERMENT EVALUATION DEBATE

The following section illustrates the application of this framework to the current debate concerning Empowerment Evaluation. The primary materials considered here include Fetterman's (1994) American Evaluation Association presidential address which made Empowerment Evaluation highly visible, Stufflebeam's (1994) critique, and Fetterman's (1995) response. These sources represent the first shots fired in this theoretical debate. There have been many subsequent volleys, as noted above, but a full analysis of all of them would far exceed space limitations. The results of this analysis are therefore not representative of the expanding literature on Empowerment Evaluation, but are restricted to the initial statements of one of the major advocates of Empowerment Evaluation, Fetterman. The purpose of this illustration is not to provide a definitive critique or defense of Empowerment Evaluation, but to summarize Fetterman's initial position within a framework that can continue to be used to examine the evolving evidence and arguments about Empowerment Evaluation. Note that this illustration applies the framework not to a particular evaluation study such as the "games" example mentioned above (Smith, 1995), nor to a comparative study of theoretical approaches such as Brandon's (1997) work, but to the meta-evaluation of a single theoretical approach.

### Context

Understanding the context of our actions helps us to appreciate their meaning and significance. Chircop (1997) reminds us to be aware of the historical embeddedness of evaluation practice, for this "allows evaluators to look at themselves, as it were, from the outside and

evaluate their own profession in terms of both (a) the profession's emergence as a reaction to a particular historical standpoint that could also have taken other forms, and (b) the social contract they have as a profession with the rest of society" (p. 3).

It is important to acknowledge, then, the social and professional climate within which Fetterman's (1994) call for Empowerment Evaluation has been made. It is a time of high professional interest in pluralist and stakeholder-based approaches to evaluation, with evaluators explicitly urged to adopt positions of stakeholder participation (Cousins & Earl, 1995), utilization focus (Patton, 1997c), and client advocacy (Greene, 1997). The presidential themes for the annual meetings of the American Evaluation Association (AEA) include Empowerment Evaluation in 1993, Social Justice in 1994, and Transforming Society Through Evaluation in 1998. Pluralist reform is currently an active interest of many evaluators and Fetterman's advocacy of Empowerment Evaluation has clearly been seen as a part of that reformist orientation: "Fetterman et al. have nailed their theses to the door of the cathedral. Now the question is, How tolerant is the establishment of dissent?" (Wild, 1997, p. 172). Further, this reformist movement in evaluation is reflective of a broader social context: "the basis of empowerment evaluation rests not in the realm of management science or efficiency but in politics. Especially in the field of public health, empowerment activities are increasingly equated with social change ..." (Brown, 1997, p. 388). Empowerment Evaluation may have received a very different reception had the social and professional zeitgeist been different.

Further, by strongly advocating Empowerment Evaluation in a presidential address, Fetterman brought the approach into immediate high visibility. This visibility has been continued due, in part, to the considerable journal space devoted to reviews and critiques of Empowerment Evaluation by other eminent theorists and AEA presidents. Fetterman is aware of the context within which his proposals are being considered. He invokes the "culturally embedded nature of our profession" (p.190) to suggest that Stufflebeam's criticism of Empowerment Evaluation reflects a Western privileging of traditional approaches which justifies excluding alternative views. According to Fetterman, attacks on Empowerment Evaluation reflect the rigid context of the current professional evaluation culture:

Charges that empowerment evaluation is pseudo evaluation and threatens "legitimate" evaluation are thus a

familiar refrain. We have heard them before; they are part of an intolerant tradition from our own past.... it is not unusual to observe lifelong resistance (to a new paradigm) particularly from those whose productive careers have committed them to an older tradition of normal science. (Fetterman, 1995, p. 180)

In addition to the professional context within which Empowerment Evaluation is being considered, context arises in two other ways in the Fetterman/Stufflebeam exchanges: context in a particular study, and the context of evaluation practice. In the former case, Fetterman (1995) emphasizes that contextual considerations, such as program life-cycles and environmental conditions, are very important in the conduct of Empowerment Evaluations in order to interpret data more meaningfully and to anticipate aspects of program operation. It is concerning the second case, the nature of the context of evaluation practice, that Fetterman and Stufflebeam again disagree.

Although Fetterman asserts that Empowerment Evaluation is designed to serve a specific need and is not a substitute for other forms of evaluative inquiry, there are apparently no limits on the contexts within which Empowerment Evaluation may be employed: the approach “can be applied to evaluation in any area, including health, education, business, agriculture, microcomputers, non-profits and foundations, government, and technology” (Fetterman, 1994, p. 1). Stufflebeam (1994) argues, however, that U.S. public education must be evaluated within the context of American democracy, the U.S. Constitution, and widely accepted general moral principles, in order to guard against exploitation of relativistic evaluation procedures for personal gain. Fetterman (1995) responds that Stufflebeam’s characterization of the evaluation context is “an atypical managerial stereotype ... layered with untenable and erroneous assumptions...” (p. 185). “Worse, it seriously underestimates program participants’ capabilities, and ignores the context and conduct of empowerment evaluations” (pp. 184-185). Fetterman argues that the objectivist, traditional evaluation approach Stufflebeam advocates is not responsive to the changing evaluation scene in which clients are demanding more than outside experts with little knowledge or vested interest in their programs, and demanding the use of participative, collaborative, and empowerment approaches in community-based evaluations. Stufflebeam and Fetterman clearly disagree on how to characterize the current context within which evaluation is practiced.

## Purpose and Social Role

Although Scriven (1997a) subsequently identified eight different definitions or definitional components of Empowerment Evaluation, Fetterman's initial statements include: "Empowerment evaluation is the use of evaluation concepts and techniques to foster self-determination. The focus is on helping people help themselves" (1994, p. 1); "It has a bias for the disenfranchised, including minorities, disabled individuals, and women. However, empowerment evaluation can be used to help anyone with a desire for self-determination" (1994, p. 12).

Fetterman introduces five forms of Empowerment Evaluation: training, facilitation, advocacy, illumination, and liberation. Each form has its own slightly different purpose and social role, which he presents primarily in terms of the procedural rules by which evaluators are to achieve the overall ends of empowering self-determination.

"In one form of empowerment evaluation, evaluators teach people to conduct their own evaluations and thus become more self-sufficient. This approach desensitizes and demystifies evaluation and ideally helps organizations internalize evaluation principles and practices, making evaluation an integral part of program planning" (Fetterman, 1994, p. 3). Training is not an isolated event, but an ongoing, integral part of the evaluation serving to continuously provide needed skills: "Training a group to conduct a self-evaluation can be considered equivalent to developing an evaluation or research design (since that is the core of the training), which would normally be considered a part of an evaluation" (Fetterman, 1995, p. 182).

"Evaluators can serve as coaches or facilitators to help others conduct their evaluation" (Fetterman, 1994, p. 4). This can be done by providing "general guidance and direction of effort" and being "responsible for helping to clear unnecessary obstacles or identify and clarify miscommunication patterns" (Fetterman, 1994, p. 5); "The empowerment evaluator coach role ensures that the evaluation remains in the hands of program personnel. The empowerment evaluator simply provides useful information, based on training and experience, to provide direction and keep the effort on track" (Fetterman, 1994, p. 6).

Evaluators may also conduct an evaluation for a group, after the goals and evaluation design have been collaboratively established. They may even serve as direct ad-

vocates — helping to empower groups through evaluation. Evaluators often feel compelled to serve as advocates for groups that have no control over their own fates, such as the homeless or dropout populations. Advocate evaluators allow participants to shape the direction of the evaluation, suggest ideal solutions to their problems, and then take an active role in making social change happen. (Fetterman, 1994, p. 6)

“Advocate evaluators write in public forums to change public opinion, embarrass power brokers, and provide relevant information at opportune moments in the policy decision making forum” (Fetterman, 1994, p. 7).

“Empowerment evaluation can also be illuminating” (Fetterman, 1994, p. 7). “This experience of illumination holds the same intellectual intoxication each one of us experienced the first time we came up with a researchable question. The process creates a dynamic community of learners as people engage in the art and science of evaluating themselves” (Fetterman, 1994, p. 8).

Empowerment evaluation can also be liberating. Many of these examples highlight how helping individuals take charge of their lives — and find useful ways to evaluate themselves — liberates them from traditional expectations and roles. They also demonstrate how empowerment evaluation enables them to find new opportunities, see existing resources in a new light, and redefine their identity and their future roles. (Fetterman, 1994, p. 9)

Stufflebeam (1994) agrees with Fetterman that helping people help themselves is important, but he argues that it is not the fundamental goal of evaluation. Stufflebeam concludes that Fetterman has confused the goal of evaluation (to systematically assess merit or worth), with its possible roles (e.g., to achieve self-determination). Stufflebeam says that while training, facilitation, advocacy, illumination, and liberation are important professional services,

such services are not evaluation. Clients who believe or claim that such constructive services from an evaluator constitute evaluation are deceiving themselves or others.... The evaluator must not confuse or substitute helping and advocacy roles with the rendering of assessments

of merit and/or worth of objects that he/she has agreed to evaluate. (Stufflebeam, 1994, p. 323)

Fetterman (1995) appears to have three responses to Stufflebeam's challenges that the purpose and social roles of Empowerment Evaluation do not justify it as a form of evaluation. First, in a footnote, Fetterman simply says he disagrees with the widely accepted definition of evaluation as the assessment of merit or worth. He says that definition is too narrow, and that, at a minimum, evaluation must also include the making of recommendations for program improvements. Further, the definition of evaluation in Empowerment Evaluation must also include "emancipatory facets, including illumination and liberation" (Fetterman, 1995, p. 196).

Second, Fetterman asserts that Stufflebeam's distinction between roles and goals is "inappropriate and off-target" (Fetterman, 1995, p. 182) because Empowerment Evaluations are conducted by groups of internal program participants rather than external individual evaluators. This response, however, seems to support Stufflebeam's criticism and to continue to confuse the goal versus role distinction.

Fetterman's third response is the most interesting. "In empowerment evaluation the context has changed: the investigation of worth or merit and plans for program improvement becomes the means by which self-determination is fostered, illumination generated, and liberation actualized" (Fetterman, 1995, p. 181).

Notice that, contrary to Fetterman's statement, it is not the context that has changed, but the purpose and social role of evaluation. With this statement, evaluation becomes not an end in itself, but a means to the ends of self-determination, illumination, and liberation. The primary intended outcome of an Empowerment Evaluation, therefore, is not increased knowledge of the merit or worth of some program, but the increased self-determination of the program participants. Claims about program value are not to be the primary result, but claims about individual and group change. The primary purpose and social role of Empowerment Evaluation is thus to create social change, not to assess merit or worth. This suggests some interesting questions, such as, if one desires to increase an individual's or group's self-determination, are there better ways to do so than by conducting an Empowerment Evaluation? Further, should other professional activities which employ evaluation procedures to achieve non-evaluative ends (such as, medical diagnosis, engineering troubleshooting, and stockmarket investment) also be consid-

ered to be forms of professional evaluation? If not, why should Empowerment Evaluation be considered a form of professional evaluation?

Patton (1997a) has subsequently pointed out that Empowerment Evaluation overlaps with other “process use”-focused evaluation approaches such as participatory, collaborative, stakeholder-involving, and utilization-focused evaluations, especially in its concern for such issues as ownership, relevance, understandability, access, involvement, improvement, and capacity-building (p. 147). Since other forms of evaluation may also empower participants, that outcome does not make Empowerment Evaluation unique. Patton argues that what sets Empowerment Evaluation apart as a distinct approach is its additional emphasis on liberation, advocacy, self-determination, and self-assessment (1997a, p. 151). He concludes that to have a true Empowerment Evaluation, all five facets of training, facilitation, advocacy, illumination, and liberation must be present, although he argues that advocating for participants rather than advocating for use of evaluation results undermines the evaluator’s credibility.

What the actual purpose and social role of Empowerment Evaluation is, and whether it constitutes a legitimate form of evaluation, has continued to be a primary focus of criticism and discussion. Sechrest (1997) has described Empowerment Evaluation as an ideological movement which appears to be “at least disdainful of, if not actively hostile to, traditional evaluation, which is to say, science and the scientific method” (p. 424). Scriven (1997a) concludes that, with Empowerment Evaluation, Fetterman “is primarily teaching and supervising evaluation in new ways and in new areas. These are good things to be doing, and his contributions are commendable, but that does not mean he is doing evaluation in a new, let alone a better, way” (p. 172). Purpose and social role are thus critical dimensions in understanding the various positions taken in the ongoing debate about Empowerment Evaluation.

### Phenomena of Interest

Although previous applications of Empowerment Evaluation have focused on a wide variety of social and educational programs (“conflict resolution, the dropout problem, environmental health and safety, homelessness, educational reform, AIDS, American Indian concerns, and the education of gifted children” (Fetterman, 1994, p.

1), the primary concern of Empowerment Evaluation is not the type of programs being evaluated, but the psychological and social condition of program participants: "Self-determination, defined as the ability to chart one's own course in life, forms the theoretical foundation of empowerment evaluation" (Fetterman, 1994, p. 2). Empowerment Evaluation emphasizes the empowerment of individuals within organizational and community settings, reflecting its earlier beginnings in community psychology and action anthropology. Further, Empowerment Evaluation frequently seems to presuppose that the organization is the dominant source of oppression.

Based on the discussion above of Stufflebeam's critique of Empowerment Evaluation and the conclusion that evaluative claims are not the desired ends of an Empowerment Evaluation but only a means to the ends of self-determination, one might conclude that the phenomena of primary interest in an Empowerment Evaluation is neither a certain type of program nor even evaluative judgments about programs, but the nature and extent of participant self-determination, illumination, and liberation. Fetterman (1994) has said that Empowerment Evaluation had its roots in community psychology — a telling metaphor since the purpose, social role, and phenomena of interest of Empowerment Evaluation more closely resemble social psychology interventions than they do program evaluation.

What does an Empowerment Evaluation look like? Would we recognize one if we saw it? A significant disagreement has arisen concerning what constitutes a clear instance of an Empowerment Evaluation.

In his review of the Fetterman et al. (1996) volume, Patton concludes that only three of the many studies presented are fairly clear cases of Empowerment Evaluation, with Levin's (1996) Accelerated Schools Project being the best example. Brown (1997) is even more positive, stating that the Accelerated Schools Project "embodies the principles of empowerment evaluation (self-determination and ongoing, internal improvement exercises) and is thus an appropriate model for what empowerment evaluation might look like in a large project" (p. 390). Although Scriven (1997a) praises Levin's work, he points out that Levin's definition of Empowerment Evaluation does not exactly match Fetterman's definition. In dissent, Sechrest (1997) argues that the evidence of the Accelerated Schools exemplary performance is program impact data collected by traditional methods. "Fetterman appears to have designated the Accelerated Schools program as a model of empowerment evaluation, even though it was

not in any explicit way guided by the philosophy of empowerment evaluation” (pp. 423–424). Fetterman (1997c) has responded that it was Levin, himself, who concluded that the Accelerated Schools project was an example of Empowerment Evaluation. “By combining systematically the theme of school empowerment with a reflective and continuous process in which assessment and evaluation are firmly embedded, the Accelerated Schools appears to embody rather fully the concept of empowerment evaluation” (Levin, 1996, p. 63).

So is the Accelerated Schools Project an outstanding example of an Empowerment Evaluation? The answer to that question seems to depend on what one takes to be the defining characteristics of the Empowerment Evaluation approach. In reaching their answers, several of these authors appear less concerned with whether participant self-determination was achieved than with evaluator adherence to certain methodological procedures.

### Procedural Rules

Fetterman suggests appropriate procedures for conducting an Empowerment Evaluation primarily through his discussion of its various forms: training, facilitation, advocacy, illumination, and liberation (1994). As discussed above, Patton (1997a) has concluded that all five facets have to be present in order to distinguish Empowerment Evaluation from other forms of process-use, stakeholder-involvement approaches. Fetterman (1994) does discuss specific evaluation activities such as participatory and collaborative study design and the use of both qualitative and quantitative data collection methods. He gives special attention to the development of participant self-ratings of both the overall program and major elements of concern to participants. These ratings are then used in the creation of statements of desired ends or program goals and used subsequently to monitor program progress.

Such procedural details are too generic and vague for some critics, however. For example, in his review of the Fetterman et al. (1996) volume, Sechrest (1997) complains that the entire volume “is virtually devoid of discussions of specific research methodology such as might be ‘required to conduct evaluations,’ and the very few references that do occur are disconcerting” (p. 424).

Stufflebeam (1994) is especially critical of the procedures employed in Empowerment Evaluation because he believes they constitute a

type of relativistic evaluation in which the criteria for judging a program “are dependent upon the particular beliefs, goals, or preferences of the client or other stakeholder,” which jeopardizes the ability of the evaluation to serve the interests of all stakeholders (p. 325). In other instances, Stufflebeam objects to the procedures advocated under Empowerment Evaluation, in part, because he disagrees with the role he believes they would have the evaluator play:

evaluators must not usurp the authority of the person(s) duly charged to make and be accountable for organizational decisions. For example, they should not publicly campaign for a particular course of action, unilaterally contact and pressure the decision maker’s policy board or funding agency to implement the evaluator’s recommended course of action, attempt to publicly discredit the decision maker, or seek to take over the decision maker’s job. (p. 328)

In other cases, Stufflebeam seems to criticize Empowerment Evaluation procedures because they lead to what he considers to be inappropriate and unjustified evaluation claims:

The approach advocated by Dr. Fetterman gives over authority to the client/interest group to choose criteria, collect data, and write/edit and disseminate reports, all in the name of self-evaluation for empowerment. The client/interest group seems to be given license to tell some preferred story, obtain the evaluator’s assistance in getting the story across to constituents or others, possibly project the illusion that the evaluation was done or endorsed by an outside expert, and remain immune from a metaevaluation against standards of the evaluation field. (p. 324)

Stufflebeam expresses his own recommendations for the proper procedures for making and justifying evaluation claims as:

It is the responsibility of the evaluator or evaluation team to carefully make judgments about quality and worth, to ground the judgments in explicit societal values and valid information, to report the findings and judgments to the right-to-know audiences, to stand able and ready to defend the evaluation report, and to promote and assist appropriate use of the findings. (p. 328)

Notice that because Fetterman and Stufflebeam disagree on the proper purpose and social role for evaluation, this, in part, leads them to also disagree on proper procedural rules and the type and grounds for justification of evaluation claims.

### Methods of Justification

The evolving contexts and social roles of evaluation have had a dramatic effect not only on the types of claims put forth in evaluation, but also on the nature of the evidence and arguments proffered in their support:

evaluation, like any other practical science, has had to enter the forum of the polis and be a contender with others for the attention of policy and decision-makers. Argument and persuasion become part-and-parcel of evaluation. In the polis the evaluative argument is being communicated to the journalist, the economist, the legislator, ... all of whom have their own professional reasoning processes ... it is the dynamic encounter of evaluators with these different contexts that has, over the past three decades, necessitated the diverse approaches of evaluation methods. (Chircop, 1997, p. 5)

The practice of evaluation as a public, as contrasted with an academic or private, enterprise, has broadened evaluators' search for acceptable methods of justification beyond the earlier reliance on solely instrumental reasoning.

Since the purpose of Empowerment Evaluation is less instrumental analysis of cause and effect and more the increasing of social power to certain constituencies, instrumental rationality is an insufficient means of justifying its evaluation claims. Appeals to commonly held values and political arguments are likely to be given greater weight in Empowerment Evaluations.

Fetterman (1994) says relatively little about how the process, outcomes, or resulting claims from an Empowerment Evaluation are to be justified. He implies, however, that it is to be the program participants themselves who are to make these judgments. They are, after all, the ones to become empowered and self-determined. As they engage in self-evaluation, they are expected to become more systematically analytical and accustomed to justifying and docu-

menting their own self-assessments. They may even become “fiercely independent and self-critical once they become a part of the self-evaluative process, often offering harsher criticism than any outside evaluator may offer” (Fetterman, 1994, p. 10-11). The role of the outside evaluator is “to help keep the effort credible, useful, and on track, providing additional rigor, reality checks, and quality controls throughout the evaluation. Without these elements in place, the exercise may be of limited utility and potentially self-serving. With these elements in place, the exercise is a dynamic community of transformative learning” (p. 11). Fetterman seems to suggest that the information produced in an Empowerment Evaluation should be subjected to traditional concerns of rigor and quality. But the primary claims resulting from an Empowerment Evaluation concern the manner and extent to which participants have achieved greater self-determination. These latter claims are apparently warranted only on the basis of subjective judgments by the participants themselves.

In response to criticisms from Stufflebeam (1994) that Empowerment Evaluation might result in self-serving studies of dubious quality, Fetterman responds that it is the program participants who are the primary evaluators in Empowerment Evaluation. This group of participants is collectively responsible for the evaluation and “serve as a check on its own members, moderating the various biases and agendas of individual members” (Fetterman, 1995, p. 183). Since Empowerment Evaluations are to be conducted in an open, public forum, bias and poor quality data are constantly subject to exposure.

Fetterman (1994) is vague about who, specifically, is to be empowered (purpose), by what means (procedural rules), and is thus vague about how one would know that such empowerment and self-determination occurred (methods of justification). Many do not find his claims and reassurances convincing, calling for follow-up studies of bias, comparative empirical tests, meta-evaluations, use of external evaluators, and so on (see Scriven, 1997a; Stufflebeam, 1994).

In responding to his critics, Fetterman raises the question of what the proper methods are for justifying evaluative claims, not about programs being evaluated, but about Empowerment Evaluation itself. In his response to Stufflebeam (1994), Fetterman (1995) charges that Stufflebeam’s criticisms are not based in logic, or practice of Empowerment Evaluation, or data, which he says exist but does not cite (Fetterman, 1995, p. 184). Fetterman claims that in the cri-

tique of his presidential address, Stufflebeam has rendered a judgment of Empowerment Evaluation without conducting an evaluation of it by collecting data, studying specific case examples, or being responsive to Fetterman's input to an earlier draft of the critique (Fetterman, 1995, pp. 185–186). Fetterman also suggests criteria he considers appropriate in evaluating innovations like Empowerment Evaluation, such as sensitivity to developmental life cycles, and knowledge of Empowerment Evaluation's goals and ethos. Fetterman also invokes *ad personem* arguments, claiming that Stufflebeam was “unresponsive,” showed a “fundamental disregard for data,” evidenced a “minimum of homework,” and used Empowerment Evaluation “as a foil or scapegoat” for many of his own concerns (Fetterman, 1995, pp. 185, 197).

Fetterman (1995) dismisses Stufflebeam's (1994) critique largely by claiming that Stufflebeam fails to provide supporting evidence for his claims. Unfortunately, Fetterman commits the same fault. A full fourth of his article is an appendix in which he discusses how the Joint Committee Standards (1994) apply to Empowerment Evaluation. Fetterman makes general statements about what should be done in an Empowerment Evaluation, without providing arguments as to why such practice is appropriate or evidence that it produces the claimed results.

In the Fetterman (1995) versus Stufflebeam (1994) interchange, then, we see five different bases for justifying evaluative claims made for and against Empowerment Evaluation: (1) logic (e.g., the definition of evaluation), (2) criteria (e.g., arguing that Empowerment Evaluation must be judged in terms of its developmental stage), (3) values (e.g., what the proper social role of evaluation should be), (4) *ad personem* arguments (e.g., the critic was ill-prepared and non-responsive), and (5) empirical evidence (e.g., that Empowerment Evaluation does, indeed, result in greater self-determination).

## Sanctions

Fetterman (1994) tends to portray Empowerment Evaluation as an unequivocal good; the only dangers of doing a poor Empowerment Evaluation seem to be that individuals will fail to become more self-determined. According to Fetterman (1994), when employed appropriately with proper support and training given to the participants conducting self-evaluation activities, Empowerment Evaluations should not result in less rigorous studies than those of other evalu-

ation approaches. Further, Fetterman cites no loss to society or the evaluation profession from evaluators “giving evaluation away” or “educating ourselves out of a job,” only benefits from “creating new roles for evaluators to help others help themselves” (p. 10).

Stufflebeam (1994) argues in response that the activities recommended under Empowerment Evaluation are not consistent with the standards of the evaluation profession, namely, the Joint Committee Standards (1994). Throughout his critique he footnotes each standard he feels would be violated by conducting an Empowerment Evaluation. He asserts that conducting Empowerment Evaluations results in “giving away the control of the evaluation’s quality and integrity and turning evaluation into pseudo evaluation exercises, in the quest to foster self-determination” (Stufflebeam, 1994, p. 333). He argues that conducting evaluations such as Empowerment Evaluations “engenders inaction and can help immobilize progress in a program ... permits authority figures to press their advantage and impose their self-interests with relative immunity to external review regarding the logic, philosophical base, and defensibility of their judgments and decisions” (p. 326). Stufflebeam insists that allowing evaluators to practice Empowerment Evaluation could “delude those who engage in such pseudo evaluation practices, deceive those whom they are supposed to serve, and discredit the evaluation field as a legitimate field of professional practice” (p. 325). “For example, it would be a serious breach of professional ethics for the evaluator to help a client project and publish a positive, noncritical image for a particular program and cause or knowingly allow the client to claim that this advocacy service constituted a defensible evaluation of the program” (p. 324). Stufflebeam argues that, at a minimum, independent meta-evaluations should accompany Empowerment Evaluations.

Fetterman (1995) responds to Stufflebeam’s (1994) charge that Empowerment Evaluation procedures repeatedly violate the Joint Committee Standards (1994) by pointing out that no recognized professional association has adopted the standards and that no mechanism for enforcing compliance with them currently exists. Further, he asserts that Stufflebeam’s (1994) analysis is an example of how the Standards can be “inappropriately implemented, resulting in misevaluation” (Fetterman, 1995, p. 185). Fetterman does accept that the Standards are relevant to Empowerment Evaluation, however, and provides a lengthy appendix discussing how he thinks Empowerment Evaluation is responsive to the Standards.

Several critics raise concerns about the potential for serious harm Empowerment Evaluation might bring to society and the evaluation profession. They urge caution and further study (e.g., Patton, 1997a; Scriven, 1997a; Sechrest, 1997; Stufflebeam, 1994). Patton notes that Empowerment Evaluation is “an idea in tune with the politics of our times. Therein lies its strength, weakness, power, and potential danger” (1997a, p. 162).

## CONCLUSION

This analysis of Empowerment Evaluation illustrates how the aspects of the framework (context, purpose and social role, phenomena of interest, procedural rules, methods of justification, and sanctions), are highly interrelated. For example, the nature of the phenomena being evaluated and their context influence what types of procedural rules are likely to be effective in achieving the desired evaluation purpose and social role. The methods for justifying evaluative claims are, of course, related to the evaluation's procedural rules, violations of which may bring on sanctions.

Empowerment Evaluation has arisen at a time of both widespread interest in reformist approaches to evaluation and resistance to radical alternatives to the hard-won methods achieved to date. There is thus disagreement as to the extent to which the current context necessitates more empowering evaluation approaches.

Under Fetterman's construction, the primary purpose and social role of Empowerment Evaluation is to promote self-determination and liberation, and the assessment of merit or worth is one means to this end. Others have responded that Empowerment Evaluation is therefore primarily a means of creating social change and not of evaluating social programs. They argue that, although Empowerment Evaluation may be a new and important use of evaluation, it is not, per se, a new form of evaluation.

The primary phenomena of interest in Empowerment Evaluation are participant self-determination, illumination, and liberation, and not the worth of programs. In identifying exemplary Empowerment Evaluations, however, the level of participant self-determination is seldom mentioned by any of these writers who instead focus on the extent to which a study contains the definitional components of Empowerment Evaluation or collects warranted evaluative data.

Fetterman suggests procedures for implementing the five forms of Empowerment Evaluation: training, facilitation, advocacy, illumination, and liberation. His recommendations have been criticized for being both vague and problematic because they encourage evaluators to act outside conventional evaluator roles. His recommendations are consistent, however, with his purposes of achieving self-determination.

According to Fetterman, traditional concerns of information quality are relevant in Empowerment Evaluation, but it is study participants who make such judgments. More importantly, the central claims of an Empowerment Evaluation, which are those concerning participant self-determination, appear to be justified solely by the perceptions of the participants themselves. When critics have charged that such a position is too subject to bias and abuse, Fetterman has shifted the debate to ask by what kinds of arguments any evaluation approach is justified. In this debate, attempts to justify the various positions use arguments employing a range of logic, criteria, values, empirical evidence, and personal attacks.

Fetterman emphasizes the benefits Empowerment Evaluation could bring to society by increasing self-determination; he mentions few detrimental effects. Others caution that use of Empowerment Evaluation would lead to loss of quality control of evaluation, greater bias and abuse in evaluation practice, and would threaten the integrity of the evaluation profession. There is strong disagreement as to how well Empowerment Evaluation complies with the Joint Committee Standards.

The preceding section only suggests the scope and complexity of the arguments being made in the debate over Empowerment Evaluation, a full analysis of which would take considerable space. My intent is not to resolve the debate, nor even to summarize all positions to date, but to suggest a framework of six categories within which to view the ongoing discussions. I believe the framework can sharpen the focus of the arguments being made, help compare and contrast the different positions, and suggest new insights and questions worthy of further analysis. The framework is useful not only in telling our current evaluation birds apart, but in recognizing and describing new species as they continue to arrive.

#### ACKNOWLEDGMENTS

An earlier version of this paper was presented at the American Evaluation Association annual meeting, November, 1997, in San

Diego, California. My thanks to Paul Brandon, Sharon May, and Andy Rowe for their helpful suggestions for improving this article; remaining shortcomings are my responsibility

## REFERENCES

- Alkin, M.C. (1991). Evaluation theory development II. In M.W. McLaughlin & D.C. Phillips (Eds.), *Evaluation and education: At quarter century* (pp. 91–112). Chicago: University of Chicago Press.
- Asimov, I. (1972). *Asimov's biographical encyclopedia of science and technology*. New York: Avon.
- Brandon, P.R. (1997, November). *A contrast of the stakeholder-assisted and participatory evaluation approaches, using Smith's meta-model of evaluation practice*. Paper presented at the meeting of the American Evaluation Association, San Diego, CA.
- Brown, J.W. (1997). Review of *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. *Health Education & Behavior*, 24(3), 388–391.
- Chircop, S. (1997, November). *Outlining the broader landscape: Multiple reasoning processes supporting an evaluative claim*. Paper presented at the meeting of the American Evaluation Association, San Diego, CA.
- Cook, T.D. (1997). Lessons learned in evaluation over the past 25 years. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st century* (pp. 30–52). Thousand Oaks, CA: Sage.
- Cousins, J.B., & Earl, L.M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Eisner, E.W. (1991). Taking a second look: Educational connoisseurship revisited. In M.W. McLaughlin & D.C. Phillips (Eds.), *Evaluation and education: At quarter century* (pp. 169–187). Chicago: University of Chicago Press.
- Fetterman, D.M. (1994). Empowerment evaluation. *Evaluation Practice*, 15(1), 1–15.

- Fetterman, D.M. (1995). In response to Dr. Daniel Stufflebeam's "Empowerment evaluation, objectivist evaluation, and evaluation standards: Where the future of evaluation should not go and where it needs to go." *Evaluation Practice*, 16(2), 179–199.
- Fetterman, D.M. (1997a). Empowerment evaluation and accreditation in higher education. In E. Chelimsky & W.R. Shadish (Eds.), *Evaluation for the 21st century: A handbook* (pp. 381–395). Thousand Oaks, CA: Sage.
- Fetterman, D.M. (1997b). Empowerment evaluation: A response to Patton and Scriven. *Evaluation Practice*, 18(3), 253–266.
- Fetterman, D.M. (1997c). A response to Sechrest's review of Empowerment evaluation. *Environment and Behavior*, 29(3), 427–436.
- Fetterman, D.M., Kaftarian, S.J., & Wandersman, A. (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.
- Greene, J.C. (1997). Evaluation as advocacy. *Evaluation Practice*, 18(1), 25–35.
- Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage.
- Levin, H.M. (1996). Empowerment evaluation and Accelerated Schools. In D.M. Fetterman, S.J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment and accountability* (pp. 49–64). Thousand Oaks, CA: Sage.
- Patton, M.Q. (1988). The evaluator's responsibility for utilization. *Evaluation Practice*, 9(2), 5–24.
- Patton, M.Q. (1997a). Toward distinguishing empowerment evaluation and placing it in a larger context. *Evaluation Practice*, 18(2), 147–163.
- Patton, M.Q. (1997b). Of vacuum cleaners and toolboxes: A response to Fetterman's response. *Evaluation Practice*, 18(3), 267–270.
- Patton, M.Q. (1997c). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.

- Rice, C.E. (1997). The scientist-practitioner split and the future of psychology. *American Psychologist*, 52(11), 1173–1181.
- Rossi, P.H., & Wright, S.R. (1977). Evaluation research: An evaluation of theory, practice, and politics. *Evaluation Quarterly*, 1(1), 5–52.
- Shadish, W.R. Jr., Cook, T.D., & Leviton, L.C. (1991). *Foundations of program evaluation*. Newbury Park: Sage.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park: Sage.
- Scriven, M. (1997a). Empowerment evaluation examined. *Evaluation Practice*, 18(2), 165–175.
- Scriven, M. (1997b). Comments on Fetterman's response. *Evaluation Practice*, 18(3), 271–272.
- Sechrest, L.E. (1997). Review of *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. *Environment and Behavior*, 29(3), 422–426.
- Smith, N.L. (1979). Requirements for a discipline of evaluation. *Studies in Educational Evaluation*, 5, 5–12.
- Smith, N.L. (1981a). Developing evaluation methods. In N.L. Smith (Ed.), *Metaphors for evaluation: Sources of new methods* (pp. 17–49). Beverly Hills, CA: Sage.
- Smith, N.L. (1981b). Evaluating evaluation methods. *Studies in Educational Evaluation*, 7, 173–181.
- Smith, N.L. (1982). Assessing innovative methods. In N.L. Smith (Ed.), *Field assessments of innovative evaluation methods* (pp. 3–9). *New Directions for Program Evaluation*, 13. San Francisco: Jossey-Bass.
- Smith, N.L. (1987). Toward the justification of claims in evaluation research. *Evaluation and Program Planning*, 10(4), 309–314.
- Smith, N.L. (1993). Improving evaluation theory through the empirical study of evaluation practice. *Evaluation Practice*, 14(3), 237–242.
- Smith, N.L. (1994). Evaluation models and approaches. In *The international encyclopedia of education* (2nd ed.). (Vol. 4, pp. 2101–2109). Oxford: Pergamon.

- Smith, N.L. (1995). The influence of societal games on the methodology of evaluative inquiry. In D.M. Fournier (Ed.), *Reasoning in evaluation: Inferential links and leaps* (pp. 5–14). San Francisco: Jossey-Bass.
- Smith, N.L., & Chircop, S. (1989). The Weiss-Patton debate: Illumination of the fundamental concerns. *Evaluation Practice*, 10(1), 5–13.
- Stufflebeam, D.L. (1994). 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.
- Weiss, C.H. (1988). If program decisions hinged only on information: A response to Patton. *Evaluation Practice*, 9(3), 15–28.
- Wild, T. (1997). Review of *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. *Canadian Journal of Program Evaluation*, 12, 170–172.
- Worthen, B.R. (1990). Program evaluation. In *The international encyclopedia of educational evaluation* (pp. 42–47). Oxford: Pergamon.
- Yanchar, S.C., & Slife, B.D. (1997). Pursuing unity in a fragmented psychology: Problems and prospects. *Review of General Psychology*, 1(3), 235–255.