

## CHAPTER

# 1

# The Origins, Meaning, and Significance of Qualitative Inquiry in Evaluation

**Thomas A. Schwandt and Timothy J. Cash**

### Key Ideas

- The definition of qualitative evaluation has been constantly contested. Several narrative accounts exist of the genesis of qualitative inquiry; each sets a particular framework for the introduction of qualitative inquiry and establishes its role in the development of program evaluation.
- The move to incorporate qualitative evaluation into the lexicon was, in some ways, a response to educational researchers' failures to demonstrate program effects using experimental designs and came on the heels of Cronbach's (1963) call for evaluators to "reconceptualize evaluation."
- Influenced by anthropology and sociology, qualitative evaluation brought new approaches to understanding human actions and meaning making to evaluation.
- Qualitative evaluation relies on methods used to generate qualitative data (e.g., interviewing, observation, focus groups, document

review) and makes use of such reporting conventions as narratives, stories, and case studies.

- Qualitative evaluation prioritizes value pluralism and considers both the stakeholders' values and the evaluator's values.
- Within qualitative evaluation, an evaluation is more about social communication than about technical reporting.

This chapter sets the stage for subsequent chapters that discuss how qualitative inquiry is related to prominent evaluation approaches and how those who conduct and are committed to the practice of qualitative inquiry view several critical issues in evaluation practice. The chapter is meant as preliminary in the sense of a beginning or general orientation to key issues involved with the origins, meaning, and significance of qualitative evaluation; it is not an exhaustive examination of these issues. The chapter begins with two brief sections—the first presents a perspective on ways in which qualitative evaluation originated; the second discusses the contested definition of the term *qualitative evaluation*. The third section identifies what, in our view, are important contributions of the extensive literature in qualitative inquiry in evaluation to shaping the practice of evaluation.

### THE GENESIS OF QUALITATIVE EVALUATION

Some notable narratives over the past three decades describe the advent and development of qualitative inquiry in the field of evaluation in several unique ways (e.g., Campbell, 1984; Conner, Altman, & Jackson, 1984; Guba & Lincoln, 1987; Madaus & Stufflebeam, 2000; Patton, 1975; Rossi & Wright, 1987; Scriven, 1984). Madaus, Scriven, and Stufflebeam (1983) argued that qualitative evaluation was one of many new conceptualizations of evaluation that arose in the late 1960s and early 1970s, primarily in the field of education. These developments followed on the heels of Cronbach's (1963) call for educational evaluators to "turn away from their penchant for post hoc evaluations based on comparisons of norm-referenced test scores of experimental and control groups" and to "reconceptualize evaluation—not in terms of a horse race between competing programs but as a process of gathering and reporting information that could help guide curriculum development" (cited in Madaus, Scriven, &

Stufflebeam, 1983, p. 12). Similarly, Guba (1969) pointed to what he called the failure of educational evaluation, arguing “the application of conventional experimental design to evaluation situations . . . conflicts with the principle that evaluation should facilitate the continuous improvement of a program” (p. 8).

For some scholars, the advent of qualitative inquiry in evaluation was inspired in large part by the failure of attempts to demonstrate the effects of Title I projects (funded by the Education and Secondary Education Act of 1964) using experimental designs. For Rossi and Wright (1987, p. 59), the introduction of qualitative evaluation was an “intellectual consequence . . . of [the] close-to-zero effects” of the social programs of the 1960s and 1970s. A strong critic of the national evaluation of Follow Through (an extension of the Head Start Program), begun in 1967 as a planned variation experiment, wrote:

We will not use the antiseptic assumptions of the research laboratory to compare children receiving new program assistance with those not receiving such aid. We recognize that the comparisons have never been productive, nor have they facilitated corrective action. The overwhelming number of evaluations conducted in this way have shown no significant differences between “experimental” and “control” groups. (Provus, 1971, p. 12)

Greene (2000) claimed that constructivist, qualitative approaches to evaluation emerged against the backdrop of several intellectual and social developments in the 1970s in the United States, including the “dethroning of experimental science as the paradigm for social program evaluation” (p. 992); a decline in the authority accorded social science theory; a decline in the authority of political figures in view of the Vietnam War, Watergate, and so on; and increased interest in value pluralism.

For Guba and Lincoln (1981, 1987, 1989; Lincoln & Guba, 1985; Guba, 1978) and Patton (1975), qualitative evaluation is about the birth (or perhaps discovery) in the 1970s of a new paradigm for evaluation derived from fieldwork methods in anthropology and qualitative sociology and from a strong interest in appropriating insights of the *Verstehen* tradition in German sociology. Scholars in that tradition held that understanding the actions of human beings as uniquely meaning-making creatures required methods different from those used to study the behavior

of nonhuman objects. Lincoln and Guba (1985) initially called this new paradigm “naturalistic inquiry” and later refined it as responsive constructivist evaluation or “fourth-generation evaluation” (Guba & Lincoln, 1989). They claimed that three prior generations of evaluation (characterized as measurement, description, and judgment) were beset by several serious problems—a tendency toward managerialism, a failure to accommodate value pluralism, and an over-commitment to a scientific paradigm of inquiry. They argued that the fourth generation addresses these problems and offers a salutary alternative.

Each of these narratives presents a different understanding of the development of program evaluation and the role qualitative inquiry has played in that story. Each employs a particular framework for shaping its account of the introduction and development of qualitative evaluation. Each provides a partial perspective on how it is that qualitative approaches arose in the broad field of evaluation. The strongest reaction to the dominance of experimental and psychometric traditions in evaluation in the 1960s and 1970s came from scholars in education who were initially trained in those traditions, including Robert Stake, Egon Guba, Lee Cronbach, and others.

The history of qualitative evaluation has often been portrayed as a struggle between different methodologies and methods or of fundamental epistemological disagreements between, for example, strong empiricists and interpretivists or post-positivists and social constructionists. These accounts are accurate to the extent that they reflect the dominance of experimental methods and the hypothetico-deductive paradigm found in texts discussing evaluation research in the late 1960s and 1970s, (e.g., Bernstein & Freeman, 1975; Reicken & Boruch, 1974; Rossi & Williams, 1972; Suchman, 1967).

Guba and Lincoln are unique in interpreting the appearance and development of qualitative evaluation as a narrative of progression or generations (although a similar idea has been advanced by Denzin and Lincoln [1994] regarding the development of qualitative research in the social sciences more generally). We are more skeptical of this way of viewing the genesis of qualitative evaluation, for there is a modernist narrative of progress implicit in the movement from one “generation” of evaluation to the next. Our thinking about evaluation may

indeed develop over time—for example, an enlargement on, improvement in, rejection of, or expansion on concepts and ideas—but earlier generations of evaluation thinking are still very active and still very much in dialogue with one another. One generation has not ceased to exist or completely given way to another.

An engaging, intellectual history of the advent and development of qualitative inquiry in the field of evaluation in the United States has yet to be written. Such a history would have to account for more than the methods wars or paradigm wars characteristic of several explanations. It would trace the influence that debates both within and outside the social sciences had on how the field of evaluation took shape, developed its multiple perspectives on what constitutes legitimate approaches to evaluation, wrestled with the politics of knowledge construction, and defined the role of professional evaluation expertise in contemporary society. Developing such an account is not our purpose here. After a brief discussion of definitions of qualitative evaluation, we offer a modest version of three sets of ideas that have their origins in the work of early proponents of the importance of qualitative inquiry in evaluation. These ideas are enduring contributions affecting the way many evaluators aim to practice what is often called qualitative evaluation.

## WHAT IS QUALITATIVE EVALUATION?

The term *qualitative evaluation*, as with most terms, from its advent in the language of evaluation (e.g., Fetterman, 1988) has been constantly contested terrain and not readily definable, because “the meaning one adopts for terms is heavily dependent on the social theory or theories that guide the use of those terms, even [between] those who use the same term” (Lincoln & Guba, 2004, p. 226). To illustrate this point, Lincoln and Guba (2004) offered the term *accountability* as an example of the many meanings a single word can take. Two people using the word may have in mind different standards for judging what accountability means, vastly different audiences, as well as vastly different social and political contexts (see also the two uses of the term in Biesta, 2004). Given that the evaluation field is filled with scholars from diverse disciplinary backgrounds

who bring with them different social theories (as well as social locations and perspectives), no wonder the terms *qualitative evaluation* and *qualitative research* are difficult to define. Schwandt's (2007) *Dictionary of Qualitative Inquiry* is evidence of this view. He offered commentary on some 380 words and phrases that are part of the vocabulary of qualitative research generally and argued that the language constituting the aim, methods, and significance of the multiple practices of qualitative research is constantly being reinterpreted.

Testimony to the fact that qualitative evaluation is not easily definable is apparent when one considers that a popular textbook on evaluation (Fitzpatrick, Sanders, & Worthen, 2011) and a comprehensive review of twenty-six approaches to conducting evaluation (Stufflebeam & Shinkfield, 2007) do not contain an entry for the term in their respective indexes. The former book indexes the term *qualitative research* in reference to a discussion of methods for gathering qualitative data. Both books do, however, discuss two ways of thinking about evaluation that are often identified as qualitative—Stake's responsive evaluation approach (2004) and Guba and Lincoln's naturalistic, constructivist, fourth-generation approach. The former book treats both as instances of participant-oriented approaches to evaluation; the latter treats each as examples of social agenda and advocacy approaches.

In his *Evaluation Thesaurus*, Scriven (1991) construed the term *qualitative evaluation* as follows:

A substantial part of good evaluation (of personnel and products as well as programs) is wholly or chiefly qualitative, meaning that description and interpretation make up all or most of it. . . . Qualitative evaluation is not a "thing in itself" but rather a complement to quantitative methods, to be combined with them when and to the extent appropriate. (p. 293)

Taking a cue from Scriven's definition, we might consider the term *qualitative evaluation* from the point of view of method—the procedures used in a particular inquiry. In this sense, an evaluation is qualitative to the extent that it relies principally on methods used to generate qualitative data, such as unstructured interviewing, field observation, focus groups, document analysis, and so on, and analyzes and interprets those data by nonstatistical

means such as analytic induction, comparative analysis, thematic analysis, taxonomies, typologies, discourse analysis, narrative analysis, qualitative content analysis, and so on. Qualitative evaluation also, generally, makes use of reporting formats particularly suited to presenting qualitative data, including narratives, stories, case studies, and in some cases, performance texts (texts that are dialogic, multivocal) and performance art (e.g., visual displays, sociodrama). For example, in his “Qualitative Evaluation Checklist,” Patton (2003) noted that qualitative evaluations often derive their data from fieldwork observations to describe activities, behaviors, actions, conversations, interpersonal interactions, and organizational or community processes, as well as open-ended interviewing to generate in-depth responses on the experiences, perceptions, feelings, and knowledge of individuals involved in an evaluation. He pointed out that these kinds of methods are often used because they yield data that facilitate capturing and telling both the “program’s story” and “participants’ stories.”

However, the use of qualitative data per se in an evaluation does not provide a definitive clue as to the type of evaluation approach being employed. Many evaluations make use of qualitative data. That is fairly commonplace. Even evaluators who strongly defend the importance of using experimental designs to evaluate social interventions argue that qualitative data are necessary (Gorard & Cook, 2007; Smith & Smith, 2009). Making use of methods to gather and analyze qualitative data may be necessary for defining an evaluation as “qualitative,” but it is not sufficient. This is so because many qualitative evaluations do not rely exclusively on qualitative data. Consider, for example, this argument from Lincoln and Guba (2004):

Please note that while some individuals think that constructivist evaluation or research is about qualitative methods or utilizes only qualitative methods, this is quite simply not true, at least as we have ‘constructed’ the paradigm. . . Naturalistic and constructivist evaluators utilize whatever methods best collect the data that answer one or another specific question. . . It is our contention that constructivist evaluation utilizes whatever data have authentic meaning for the question at hand, whether qualitative or quantitative. (p. 233)

Thus, method choice is clearly not the sole determinant of what ever one identifies as qualitative evaluation.

When considered from the point of view of methodology—a framework of assumptions and principles for how a particular approach to inquiry should proceed—defining qualitative evaluation becomes even more complicated. A number of approaches to evaluation are possible, including educational connoisseurship and criticism, case study evaluation, responsive evaluation, illuminative evaluation, naturalistic inquiry or fourth-generation evaluation, appreciative inquiry, feminist evaluation, culturally responsive evaluation, and variants of participatory and collaborative evaluation. Each draws heavily on the methods noted above but does not necessarily share the same methodological characteristics and assumptions about what constitutes significant evaluation knowledge and how it is to be established and warranted. The approaches differ in important ways in their assumptions about the nature of knowledge, the politics of knowledge production, and the responsibility and role of the inquirer/evaluator. This suggests that qualitative evaluation is perhaps, more or less, a family of approaches comprising relatives that do not always agree but generally are happy to share a common name.

Some of these approaches may endorse (either explicitly or tacitly) a form of social and cognitive constructionism (Greene, 2000; Guba & Lincoln, 1989). Social constructionism is the view that “the actions, artifacts, and events of everyday life depend (to some degree) for their identity, intelligibility, stability, and consequences on collective practices, conventions, and representations” (Heap, 1995, p. 52). Cognitive constructionism is the view that the identity, stability, and so on of these aspects of everyday life depend (to some degree) “on cognitive representations (plans, schemata, goals, processing rules) and processes as enabled and constrained by cognitive architecture” (p. 52). Constructionists, in brief, hold that we do not discover the meaning of actions, events, or concepts such as deviance, disability, identity, gender, race, selfhood, and so on, but rather we construct or make meaning against a backdrop of shared understandings, language, and practices. This kind of constructionism is opposed to a version of realism that assumes that our knowledge is a reflection or mirror of what is out there in the world. However, neither a necessary nor an irrevocable connection exists between the use of qualitative methods as



discussed and a social constructionist way of thinking. In other words, evaluators can quite readily and reasonably make observations and conduct interviews—look, see, and record—even if they are realists and believe that they are recording the way things “really are.”

Despite their differences, all members of the family called qualitative evaluation share what Stake (2004, p. 30) has called a “naturalistic disposition” or “persuasion” to value “ordinary activities in their settings.” Patton (2003) similarly speaks of naturalistic inquiry as being about studying a program “as it unfolds naturally and openly, that is, without a predetermined focus or [predetermined] categories of analysis” (p. 3). This disposition can be elaborated to include:

- A concern for programs and projects being evaluated as multifaceted, complex compositions of the experiences of those individuals and groups most strongly influenced or affected by the program or project; experiences are permeated with meaning (perceptions, understandings, judgments)
- A concern with capturing the diachronic (historical or long-term) and synchronic (specific point in time) character of program and project activity
- A strong interest in grasping the views and voices of people associated with that activity via rich description and explanation of processes occurring in local contexts (Schwandt & Burgon, 2006).

## ENDURING CONTRIBUTIONS

We have already identified a shared disposition among members of the family of qualitative evaluation. Here we single out prominent commitments that further shape the way in which those who subscribe to the importance of qualitative inquiry practice evaluation. The emphasis here is primarily on decisions made in practice, where practice is understood not as a site or location where theory or abstract concepts are applied but as engaged, “embodied, materially mediated arrays of human activity centrally organized around shared practical understanding” (Schatzki, 2001, p. 2). Although we do not subscribe to a

sharp theory–practice divide (Schwandt, 2002, 2006), we do not concern ourselves here with an examination of philosophical or, more specifically, epistemological commitments as much as with decisions that orient the way an evaluation is crafted and conducted.

The “shared practical understanding” of many advocates of qualitative inquiry in evaluation is reflected in ways an evaluation is focused, values are understood, and an evaluation is communicated. Given the preliminary nature of this chapter, we do not tackle the considerable literature that involves extensions, commentary, and critical appraisal of these commitments.

### ***Focusing an Evaluation***

Patton (2008, p. 229) noted that the act of focusing an evaluation involves answering the question of “what’s worth knowing,” and he emphasized that many different answers to that question can be legitimately posed in evaluation, including, was the program effective, were objectives achieved, what intended (and unintended) outcomes occurred, how was a program implemented, were resources used appropriately, what is the relation between program costs and outcomes, and so on.

Broadly speaking, three signature responses to “what’s worth knowing” occur for many who practice qualitative forms of evaluation. For those evaluators committed to forms of ethnographic evaluation guided by anthropological theory (Dorr-Bremme, 1985; Fetterman & Pittman, 1986; Hopson, 2002), “what’s worth knowing” relates to understanding a program’s culture. For many evaluators influenced by responsive evaluation, “what’s worth knowing” are issues. Stake (2004) explained:

Issues are regularly taken to be “conceptual organizers” for the investigation, more so than needs, objectives, hypotheses, group comparisons, and cost–benefit ratios. Issues are obstacles, organizational perplexities, and social problems, drawing attention especially to unanticipated responsibilities and side effects. With the term issues, we try to draw thinking toward the interactivity, particularity, and subjective valuing felt by persons associated with the program. (p. 209)

For evaluators more attuned to constructivist, fourth-generation evaluation, issues not only take on a somewhat

different meaning from that found in Stake's responsive evaluation, but also claims and concerns become important aspects of what's worth knowing as well. Guba and Lincoln (1989) explained:

A claim is any assertion that a stakeholder may introduce that is favorable to the evaluand, for example, that a particular mode of reading instruction will result in more than a year's gain in standard test reading scores for every year of classroom use, or that a particular mode of handling domestic disturbance calls by police will materially reduce recidivism in offenders. A concern is any assertion that a stakeholder may introduce that is unfavorable to the evaluand, for example, that instruction in the use of a computer materially reduces pupils' ability to do computations by hand . . . [and] an issue is any state of affairs about which reasonable persons may disagree, for example, the introduction of education about AIDS into elementary schools . . . Different stakeholders will harbor different claims, concerns, and issues; it is the task of the evaluator to ferret these out and to address them in the evaluation. (p. 40)

Both responsive evaluation and fourth-generation evaluation clearly welcome a plurality of stakeholder views. However, as noted later, a difference is seen in the way that plurality is addressed. Responsive evaluation regards multiple perspectives, more or less, as the conceptual structure that the evaluator uses to portray the program in question. Fourth-generation evaluation puts these perspectives into dialogue with one another in the process of conducting the evaluation.

Individuals practicing qualitative forms of evaluation employ these broadly responsive and cultural orientations in different ways. For example, Abma (2006) has incorporated Stake's idea of responsiveness with Guba and Lincoln's emphasis on a strong view of stakeholder participation and the manner in which an evaluation can (should) facilitate dialogue and action among stakeholders. She frames evaluation as "an engagement with and among all stakeholders about the meaning of their practice. Responsive evaluation focuses on stakeholder issues and engages stakeholders in dialogues about the quality of their practice. The aim is to heighten the personal and mutual understanding of stakeholders as a vehicle for practice improvement" (Abma, 2006, p. 31). (See Chapter 7 of this volume to learn how Vivianne

Baur and Tineke Abma facilitated dialogue and action to address power differences.)

Culturally responsive evaluation (Hood & Rosenstein, 2005), another form of evaluation that makes heavy use of qualitative inquiry, also is oriented in this way, although it argues that being responsive to issues, concerns, and experiences of stakeholders in cultures other than one's own is an epistemological, ethical, and political challenge.

In response to claims, concerns, and issues of stakeholders most closely associated with a program, qualitative evaluators are primarily interested in gathering the subjective perceptions of what is going on with respect to a program. In this instance, *subjective* does not mean biased or unreliable (a common use of the word). Rather, subjective is used to indicate that these perceptions come from the subject—they represent the personal view of an individual or the subject's point of view based on his or her (or their) historical, political, cultural, social, material lived experience. Typically this kind of information is gathered via qualitative methods such as interviewing and focus groups, although collecting such information via surveys and questionnaires is also possible.

### ***Valuing in Evaluation***

Grasping the multiple perspectives of key stakeholders in an evaluation inevitably involves the qualitative evaluator in understanding the values associated with those perspectives. As Stake (2004) has expressed it, “responsive evaluation recognizes multiple sources of valuing and multiple grounds” (p. 210). Qualitative evaluations are inherently concerned not simply with the facts of the matter (what happened, when, to whom, how often, and so forth). They are also concerned with values, including making the evaluator's own value commitments recognizable as well as “helping stakeholders articulate their values, considering the values inherent in the evaluation, and working to portray the program through different stakeholders' perspectives” (Fitzpatrick, Sanders, & Worthen, 2011, p. 116).

Generally speaking, qualitative evaluators adopt a stance of value pluralism (Berlin, 1990), which is a belief that genuine values are many and may, and often do, come into conflict with

one another. Endorsing value pluralism means acceptance of the heterogeneity of values and ways of knowing that characterize everyday life and the practices of teaching, learning, managing, leading, providing health and human services, and the like in which we find ourselves embedded. Pluralism stands in opposition to a monist view that only one set of values or one way of knowing is correct, true, and valid; all others are incorrect, false, and invalid. It also opposes a relativist view, which holds that my values or way of knowing are mine, yours are yours, and neither of us can claim to be right.

For some evaluators, pluralism is about finding ways to orchestrate, juxtapose, and place into a coherent, engaging conversation multiple views. Just how that is to be done is the subject of considerable debate among those who endorse a qualitative view of evaluation. In their model of constructivist, fourth-generation evaluation, Guba and Lincoln (1989) argued for a process they labeled the hermeneutic dialectic circle, in which evaluators and stakeholders confront each other's constructions. They explained that

Using what Stake called "portrayals," various aspects of the program are displayed, explained, explored from multiple perspectives, and stakeholders invited to comment or elaborate upon, correct, amend, extend, or otherwise make more accurate or precise the information, data, and interpretations. . . . The constant interaction . . . is what makes this model hermeneutic. Such interaction creates new knowledge, and permits old or taken-for-granted knowledge to be elaborated, refined, and tested. The dialectic of this evaluation model is the focus on carefully bringing to the fore the conflict inherent in value pluralism. (Lincoln & Guba, 2004, p. 235)

Lincoln and Guba (2004, p. 235) claimed that getting at "core values of participants and stakeholders" is necessary "so that when decisions are made, the value commitments that those decisions represent are clear, negotiable, and negotiated between and among stakeholders." Although not fully endorsing all of the assumptions of Lincoln and Guba's constructivist approach to evaluation, House and Howe (1999) argued similarly that value pluralism is best addressed through a process of democratic deliberation in which both facts and values are debated. House (2005, p. 220) stressed that such a process "aspires to arrive

at unbiased conclusions by considering all relevant interests, values, and perspectives; by engaging in extended dialogue with major stakeholders; and by promoting extensive deliberation about the study's conclusion."

Other broadly qualitative approaches to evaluation, such as empowerment evaluation (Fetterman, Kaftarian, & Wandersman, 1996) and transformative evaluation (Mertens, 2008), argue strongly for privileging the voices of the least advantaged individuals and groups with a stake in a program being evaluated. One distinction among practitioners of qualitative evaluation (and all practitioners of evaluation, for that matter) is whether they adopt a descriptive or prescriptive approach to valuing. In a descriptive approach, an evaluator

Describes values held by stakeholders, determines criteria they use in judging program worth, finds out if stakeholders think the program is good, and sees what they think should be done to improve it. The claim is not that these values are the best but that they are perceptions of program worth that are grist for the mill of decision making" (Shadish, Cook, & Leviton, 1991, p. 49).

A prescriptive theory of valuing argues that some value perspectives are more important than others.

Another dimension of valuing in evaluation is specifically concerned with the evaluator's own value stance. Again, broadly speaking, evaluators committed to qualitative inquiry in evaluation are highly critical of a stance of value neutrality on the part of the evaluator—the view that evaluators (and social scientists, more generally) should concentrate on improving methods for generating descriptions and explanations and leave questions of valuing to others. They freely acknowledge that the evaluator brings to an evaluation his or her own value commitments and perspectives and that evaluation itself, as a social practice, embodies particular political and moral commitments. Transformative evaluation, culturally responsive evaluation, and Greene's (2006) thinking about the relationships among evaluation, democracy, and social justice are all illustrative of this idea.

### ***Communicating an Evaluation***

Early critics of the dominance of experimental and psychometric methods in evaluation often opposed the use of these methods on the grounds that the kinds of data collected and reported were simply not useful to helping those most intimately involved with a program better understand what was happening in the program and how it might be improved. The reporting was often so technically sophisticated that the experience of what the program was actually like for those involved with it was obscured or completely lost. Although a concern with use permeates all forms of evaluation, qualitative inquiry in evaluation strongly emphasizes that an evaluation is less an act of technical reporting and more an act of social communication. This idea is clearly evident in Guba and Lincoln's view of how information is shared with stakeholders, as noted previously.

In a keynote presentation at a conference on "New Trends in Evaluation" in October 1973, at the Institute of Education at Göteborg University, Stake (1974) was highly critical of approaches to educational evaluation that he labeled "pre-ordinate," which relied on a statement of program goals and objectives, the use of objective tests as the primary means of collecting data, on standards held by program personnel, and on research-type reports. He argued that his idea of responsive evaluation is less reliant on formal communication and more reliant on natural communication. He added:

We need a reporting procedure for facilitating vicarious experience. And it is available. Among the better evangelists, anthropologists, and dramatists are those who have developed the art of storytelling. We need to portray complexity. We need to convey holistic impression, the mood, even the mystery of the experience. The program staff or people in the community may be 'uncertain.' The audiences should feel that uncertainty. More ambiguity rather than less may be needed in our reports. (p. 12)

The central idea here is that evaluation reporting should be a means of enhancing understandings of a program across the principal stakeholders in that program. More recently, Stake

(2004) has explained that responsive evaluation reports are “expected to be, in format and language, informative and comprehensible to the various audiences. . . . Thus even at the risk of catering, different reports or presentations may be prepared for different groups. Narrative portrayals, storytelling, and verbatim testimony will be appropriate for some, data banks and regression analyses for others” (p. 213). The preference of many evaluators committed to qualitative inquiry in evaluation for case-study forms of reporting that heavily emphasize careful description and detail is consonant with this way of thinking about evaluation as a form of social communication. Consider, for example, this definition drawn from a US General Accounting Office (1990) report: “A case study is a method for learning about a complex instance, based on a comprehensive understanding of that instance obtained by extensive description and analysis of that instance taken as a whole and in its context.”

## CONCLUSION

This brief description of what qualitative inquiry means for the practice of evaluation is intended to be more illustrative and generative of further thought than definitive. A full explanation of the role that qualitative inquiry plays in evaluation would require examining the many evaluation approaches that make use of qualitative inquiry and investigating how these approaches address, for example, evaluation purpose and evaluator roles and responsibilities. We have endeavored to illustrate the complexity behind how qualitative inquiry came to be regarded as a significant dimension of evaluation theory and practice. Likewise, we have cautioned that no simple exercise in defining terms can capture what it means to endorse the importance of qualitative inquiry in evaluation. Much more remains to be said about how the theory and practice of qualitative inquiry in evaluation rests on important issues in the epistemology, politics, and ethics of evaluation theory and practice. Those caveats considered, we believe this chapter serves as a preliminary look at what the practice of qualitative evaluation entails. It is reasonable to believe that the family of evaluators committed to the value of qualitative evaluation does in fact share the general orientations we have identified.



## KEY CONCEPTS

Constructivism	Interpretivists
Continuous improvement	Naturalistic inquiry
Culturally responsive evaluation	Post-positivists
Deliberation	Prescriptive
Descriptive	Qualitative evaluation
Empiricists	Qualitative inquiry
Epistemology	Responsive evaluation
Evalund	Social constructionists
Experimental designs	Subjective
Fourth-generation evaluation	Value pluralism

## DISCUSSION QUESTIONS

1. The introduction of qualitative inquiry to evaluation in the late 1960s and early 1970s significantly informed and expanded the field. How have the values associated with qualitative inquiry influenced the field of evaluation?
2. The authors describe the importance of storytelling to portray the complexity of a program. Discuss various ways that an evaluator can write a formal report that “tells the story” of the program in a manner that is meaningful, interesting, and useful to the sponsor of an evaluation.
3. The authors discuss the concept of “what’s worth knowing” for evaluators who practice qualitative inquiry in evaluation. They present the information from the perspective of the evaluator. How might an evaluator navigate a situation in which the evaluator thinks one thing is worth knowing, but the funder thinks something else is worth knowing?

## REFERENCES

- Abma, T. (2006). The practice and politics of responsive evaluation. *American Journal of Evaluation*, 27(1), 31–43.
- Berlin, I. (1990). *Four essays on liberty*. Oxford, UK: Clarendon Press.
- Bernstein, I., & Freeman, H. E. (1975). *Academic and entrepreneurial research: Consequences of diversity in federal evaluation studies*. New York, NY: Russell Sage Foundation.
- Biesta, G. J. J. (2004). Education, accountability, and the ethical demand: Can the democratic potential of accountability be regained? *Educational Theory*, 54(3), 233–250.
- Campbell, D. T. (1984). Can we be scientific in applied social science? In R. F. Conner, D. G. Altman, & C. Jackson (Eds.), *Evaluation studies review annual* (Vol. 9, pp. 26–48). Beverly Hills, CA: Sage.
- Conner, R. F., Altman, D. G., & Jackson, C. (1984). A brave new world for evaluation? In R. F. Conner, D. C. Altman, & C. Jackson (Eds.), *Evaluation studies review annual* (Vol. 9). Beverly Hills, CA: Sage.
- Cronbach, L. J. (1963). Course improvements through evaluation. *Teachers College Record*, 64(8), 672–683.
- Denzin, N. K., & Lincoln, Y. S. (Eds.). (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Dorr-Bremme, D. W. (1985). Ethnographic evaluation: A theory and method. *Educational Evaluation and Policy Analysis*, 7(1), 65–83.
- Fetterman, D. M. (1988). A qualitative shift in allegiance. In D. M. Fetterman (Ed.), *Qualitative approaches to evaluation in education* (pp. 3–19). New York, NY: Praeger.
- Fetterman, D. M., Kaftarain, S., & Wandersman, A. (Eds.). (1996). *Empowerment evaluation: Knowledge and tools for self-assessment and accountability*. Thousand Oaks, CA: Sage.
- Fetterman, D. M., & Pittman, M. A. (Eds.). (1986). *Educational evaluation: Ethnography in theory, practice and politics*. Newbury Park, CA: Sage.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson.
- Gorard, S., & Cook, T. (2007). Where does good evidence come from? *International Journal of Research & Method in Education*, 30(3), 307–323.
- Greene, J. C. (2000). Understanding social programs through evaluation. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (2nd ed., pp. 981–1000). Thousand Oaks, CA: Sage.
- Greene, J. C. (2006). Evaluation, democracy and social change. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation*. London, UK: Sage.

- Guba, E. G. (1969). The failure of educational evaluation. *Educational Technology*, 9, 29–38.
- Guba, E. G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. *CSE monograph series in evaluation* (Vol. 8). Los Angeles, CA: Center for the Study of Evaluation.
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation*. San Francisco, CA: Jossey-Bass.
- Guba, E. G., & Lincoln, Y. S. (1987). The countenances of fourth generation evaluation: Description, judgment, and negotiation. In D. S. Cordray & M. W. Lipsey (Eds.), *Evaluation studies annual review* (Vol. 11, pp. 70–88). Beverly Hills, CA: Sage.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- Heap, J. L. (1995). Constructionism in the rhetoric and practice of fourth generation evaluation. *Evaluation and Program Planning*, 18(1), 51–61.
- Hood, S., & Rosenstein, B. (2005). Culturally responsive evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation*. Los Angeles, CA: Sage.
- Hopson, R. M. (2002). Making (more) room at the evaluation table for ethnography. In K. E. Ryan & T. A. Schwandt (Eds.), *Exploring evaluator role and identity*. Greenwich, CT: Information Age Press.
- House, E. R. (2005). Deliberative democratic evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (p. 220). Thousand Oaks, CA: Sage.
- House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, CA: Sage.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Lincoln, Y. S., & Guba, E. G. (2004). The roots of fourth generation evaluation: Theoretical and methodological origins. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 225–241). Thousand Oaks, CA: Sage.
- Madaus, G. F., Scriven, M., & Stufflebeam, D. L. (Eds.). (1983). *Evaluation models: Viewpoints on educational and human services evaluation*. Boston, MA: Kluwer-Nijhoff.
- Madaus, G. F., & Stufflebeam, D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 3–18). Norwell, MA: Kluwer Academic Publishers.
- Mertens, D. M. (2008). *Transformative research and evaluation*. New York, NY: Guilford Press.
- Patton, M. Q. (1975). *Alternative evaluation research paradigm*. Grand Forks, ND: University of North Dakota.
- Patton, M. Q. (2003). *Qualitative evaluation checklist*. Available at [www.wmich.edu/evalctr/archive\\_checklists/index.html](http://www.wmich.edu/evalctr/archive_checklists/index.html)

- Patton, M. Q. (2008). *Utilization-focused evaluation*. Thousand Oaks, CA: Sage.
- Provus, M. (1971). *Discrepancy evaluation for educational program improvement and assessment*. Berkeley, CA: McCutchan.
- Reicken, H. W., & Boruch, R. F. (1974). *Social experimentation: A method for planning and evaluating social intervention*. New York, NY: Academic Press.
- Rossi, P. H., & Williams, W. (Eds.). (1972). *Evaluating social programs: Theory, practice and politics*. New York, NY: Seminar Press.
- Rossi, P. H., & Wright, J. D. (1987). Evaluation research: An assessment. In D. S. Cordray & M. W. Lipsey (Eds.), *Evaluation studies annual review* (Vol. 11, pp. 48–69). Beverly Hills, CA: Sage.
- Schatzki, T. R. (2001). Introduction: Practice theory. In T. R. Schatzki, K. K. Cetina, & E. Von Savigny (Eds.), *The practice turn in contemporary theory* (pp. 2–14). London, UK: Routledge.
- Schwandt, T. A. (2002). *Evaluation practice reconsidered*. New York, NY: Peter Lang.
- Schwandt, T. A. (2006, November). Initiation into the theory and practice of evaluation: A commentary on key ideas guiding our thinking. Paper presented at the annual meeting of the American Evaluation Association, Portland, OR.
- Schwandt, T. A. (2007). *Dictionary of qualitative inquiry* (3rd ed.). Thousand Oaks, CA: Sage.
- Schwandt, T. A., & Burgon, H. (2006). Evaluation and the study of lived experience. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation: Policies, programs and practices* (pp. 98–117). London, UK: Sage.
- Scriven, M. (1984). Evaluation ideologies. In R. F. Conner, D. G. Altman, & C. Jackson (Eds.), *Evaluation studies review annual* (Vol. 9, pp. 49–80). Beverly Hills, CA: Sage.
- Scriven, M. (1991). *Evaluation thesaurus*. Thousand Oaks, CA: Sage.
- Shadish Jr., W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Smith, M. S., & Smith, M. L. (2009). Research in the policy process. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of education policy research* (pp. 372–397). Washington, DC: American Educational Research Association.
- Stake, R. E. (1974). Program evaluation, particularly responsive evaluation. Reproduced in W. B. Dockrell & D. Hamilton (Eds.). (1980). *Rethinking educational research*. London, UK: Hodder and Stoughton.
- Stake, R. (2004). Stake and responsive evaluation. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 372–397). Los Angeles, CA: Sage.

- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.
- Suchman, E. A. (1967). *Evaluative research: Principles and practice in public service and social action programs*. New York, NY: Russell Sage Foundation.
- US General Accounting Office. (1990, November). *Case study evaluations. USGAO program evaluation and methodology division*. Washington, DC: Author.

