

Pre-publication Draft. Please do not quote without permission.

The Language of Evaluation

Stanley L Witkin, Ph.D.

Professor

Department of Social Work

University of Vermont

Burlington, Vermont 05405

USA

switkin@uvm.edu

Plenary address presented at the 4th International Conference on Evaluation for Practice
Tampere, Finland, July 2002

The Language of Evaluation

Introduction

Evaluation is about language. Whatever we might call our approach to evaluation - realist, interpretivist, constructionist - we all rely on written and spoken language to comprehend, understand, analyze, and communicate our activities. However, despite the inevitability and salience of language in evaluations, relatively little attention has been given to its use and influence, or to its potential for changing the way we understand, conduct and assess evaluations. This presentation is a modest attempt to address this situation.

One of my favorite aphorisms goes something like, “How do I know what I think until I see what I write?” The wisdom of this question became apparent to me as I was preparing this presentation. When I originally proposed the title, “The Language of Evaluation,” to the conference organizers several months ago, I had a general idea of what I hoped to write about. However, it was not until I began writing that I began to “see” what I was thinking about this topic. Now that the paper’s complete, at least complete enough to present, I think a more accurate title would be “Evaluation as Discourse,” which, of course, encompasses language but goes a little further. The aphorism also reflects an important theme of my paper: that the process of evaluation is reflexively related to the world that we evaluate. That is, evaluations enact the environments they evaluate. The objects of evaluation are rendered visible and inscribed through the actions of the evaluator.

The broader intellectual context for my paper is the changes in conceptions of knowledge and knowledge development that have occurred over the past 40 or so years. In particular, I am referring to the sustained social, ideological, and literary critiques of Western, mainstream philosophical thought and its expression in science and research that have led to a more

historically, culturally, and socially contingent view of knowledge. Interestingly, these changes have not had much impact on professional journals or academic conferences. Most academic conferences, for example, have been using the same format for decades: a speaker lectures, an audience listens (or pretends to listen), maybe asks a few questions, and then, everyone goes home. This format generates a kind of disembodied knowledge that minimizes the speaker's personhood and its importance to the topic and how it is interpreted. If you learn anything at all about the speaker, it is about his or her affiliations and accomplishments - information that lends credibility to whatever is spoken or written. This format also ignores the extent to which knowledge is socially produced, giving the impression that it is the product of an individual mind.

In contrast to this disembodied knowledge is the kind of knowledge that emerges from people in relationship. I talk differently to (and hear differently with) people I know, people with whom I have a relationship, than with people I do not know. For example, I am more likely to be humorous, take risks, reflect "out loud," and engage in sustained dialogueⁱ. Information has more of a give and take, co-constructed quality inseparable from its dialogic, interpersonal context. There is an "active negotiation of meaning" (Mehan, 1997, p. 270) between equals. I would suggest that evaluations also differ in the degree to which they generate these different types of knowledges.

Although the canons of conventional academic discourse suggest that who I am has little to do with what I believe, that is not my experience nor that of anyone I know. What I am interested in and how I think about those interests are very connected with my location and experience in the world, including my identification with and vision of the profession of social

work. Knowing something about me provides an interpretive context for my ideas - whether or not you agree with them.

In many ways the evolution of my philosophical thinking and methodological allegiances during my academic career have paralleled the changes in some areas of philosophy and the social sciences that I alluded to earlier. These changes, to some extent, also have occurred in U.S. social work. Thus I went from a being “card carrying” behaviorist early in my career to a social constructionist, a change that some in the profession found difficult to understand (see Thyer, Harrison, & Hudson, 1992). These changes were connected, as best as I can tell, to multiple factors. Some were personal experiences: the birth of my third child with multiple disabilities had a profound effect on my life. It led me to ask basic questions about what it meant to be a human being and it revealed for me the poverty of the behavioral model of understanding. This, in turn, inspired me to search for new models that would better enable me to comprehend this situation and respond in ways that were consistent with my personal and professional values. Other experiences involved encounters with people or writings that raised provocative questions or proposed new conceptual frames for making sense of the world. For example, Michael Mahoney’s work on the psychology of scientists in the 1970s (Mahoney, 1976) and Ken Gergen’s work on social construction in the 1980s (Gergen, 1985) were influential in this regard. Through these and other experiences, I began to realize that hidden within the claims of neutrality and discovery of conventional social science were authority, values, and agency. I began to forge a new understanding of social life and inquiry that took these realizations into account and that was informed by the values and mission of social work. This took me into the realm of the postmodern.

Talking Postmodern

Presenting a “postmodern” perspective is challenging. Because of the emphasis on language and discourse, postmodernists face the dilemma of using a lexicon that reflects a modernist worldview to express postmodern ideas. Brown (1990) articulated this challenge in the following way: One must use a known language, with its inherent vantage point and presuppositions, to say anything intelligible. *But it is difficult to convey a new vision in an established discourse.* If the new perspective is too closely wedded to a new mode of representation, it will appear incomprehensible to users of the old. But if the new vision is encoded in the old language, this very language, although easily comprehended, may contradict the new message that the author is struggling to express (p. 60, emphasis added)ⁱⁱ. For some postmodernists, the challenge not only is contradiction, but how to avoid reproducing the very realities that one is trying to “deconstruct” or replace. Surber (1998), for instance, noted that “the impenetrable and idiosyncratic style of much poststructuralist critique” is due to their strong desire to avoid “. . . the construction of yet another ‘true’ or ‘conceptual’ discourse, which would lead them into the very trap they sought to critique . . .” (p. 188).

The issue of language seems particularly relevant at a conference where most of the attendees are not native speakers of English and yet, English is the “official” language of the meeting. There are two points I want to make about this. First, I want to express my gratitude for being allowed to speak in my native language of English - the Finnish version of my talk is quite brief - and the deep respect I have for the multilingual talents of so many of you. I am quite aware that we not using English because it is the “best” language, but because it happens to be the language of the world’s most powerful nation. And I am also aware of how difficult English can be, particularly when addressing highly complex topics such as evaluation. For example,

take the word “play.” According to my dictionary there are 94 definitions of this word. It seems to me that it takes a fairly sophisticated language user to be able to discern which of those 94 to use, or are being used, in a given sentenceⁱⁱⁱ.

This leads to my second point. To the extent that evaluations are carried out or reported in English, many of you have a double challenge: to speak the language of evaluation and to speak it in English. This becomes particularly important when language is viewed as constituting evaluations. How adequately can one capture the nuances and complexities of social life when they are translated into evaluation and English languages? What concepts do not exist or exist differently in your native language and English? Perhaps at some point we can discuss this issue.

From Reflection to Constitution

In the correspondence theory of language, the “meaning of a word is taken to be what it references, relates to, or stands for in the real world.” “In this framework, the essential task of language is to convey information, to describe reality” (Gubrium and Holstein, 1997, p. 54). In contrast, from a postmodern perspective, “the distinction between what signifies and what is signified” collapses and “signs no longer claim to depict, mirror, or even disguise an objective reality” (Ashley, 1997, pp. 4-5). What is taken to be real and the language used to represent it cannot be disentangled as any attempt to do so also would require language. Language constitutes rather than reflects reality.

What is the connection between experience and its representation? If there is no independent reality in which to anchor our observations and if language plays a constitutive role in our sense of the real, then our modes or styles of representation take on paramount importance. In some contexts such as ethnography, this has led to a “crisis of representation” [in which] “There can never be a final, accurate representation of what was meant or said—only

different textual representations of different experiences” (Denzin, 1997, p. 5). Whether or not one endorses this view, attention to representational issues has led to increased awareness of the realities that are constituted within texts. Particularly cogent in this regard has been the exclusion or negatively skewed presentations of women, people of color, people with disabilities, and others who carry marginalized statuses.

Questions about the nature of reality are complex and diverse viewpoints abound. Certainly, they are not resolvable here. I concur with Gergen (1994) that “Once we attempt to articulate ‘what is there,’ . . . we enter the world of discourse.” . . . [which is] inextricably woven into processes of social interchange and into history and culture” (p. 72). This position underscores my primary theme: that evaluation be considered as discourse.

Evaluation As Discourse

This “world of discourse” that Gergen identifies is both created by and the context for evaluations. If we take this idea seriously, it opens up new and, I believe, potentially useful ways to think about and practice evaluation. Discourse, however, is not a straightforward concept. Rather, it has various meanings depending on the speaker/author, context, purpose and so on. Below are some examples which capture the senses in which I am using the term:

“Discourses are systems of statements that construct an object produced and reproduced in conversation and written text” (Newman & Holzman, 1997, p. 54).

[Discourses are] “systems of meaning which offer positions of power to some categories of people and disempower others” (Parker, 1992, p. 10).

[Discourses are] “practices that systematically form the objects of which they speak” (Foucault, 1972, p. 49).

“Discourses are structures of knowledge, claims, and practices through which we understand, explain, and decide things. In constituting agents, they also define obligations and determine the distribution of responsibilities and authorities for different categories of people, such as parents, children, social workers, doctors, lawyers, and so on” (Chambon, 1999, p. 57).

Some characteristics of these definitions that I wish to highlight include:

Discourses are ways of understanding the world. As Foucault and others have stressed, discourses are more than “mere words.” Rather they are basic to the thought structures that we use to make sense of things. What we think of as family or science or sex are embedded in and constituted through the dominant discourses of our cultures.

Discourses are expressed through language. Although this point may seem obvious, it is central to my discussion. Further, although discourses may be manifested in ways other than language (for example, in the structure of an organization), their linguistic expression serves to reproduce, disrupt, or regulate social life. Additionally, some properties of language, namely, its historical and cultural embeddedness and ability to be shaped by various social and political forces are particularly important for evaluation.

Discourses are practices. We “do” discourses. They are not passive but are “activated” by speaking, writing, and other forms of action such as administering questionnaires or interviewing. In this sense, evaluation practices enact discourses.

Discourses construct objects. Discourses generate their own realities. “Events in the world do not exist for people independently of the representations people use to make sense of them. Instead, objects are defined through elaborate enactments of cultural conventions which lead to the establishment of such well documented ‘institutional facts’ (Searle, 1969) as ‘touchdowns,’ ‘marriages,’ ‘insults,’ ‘banishments,’ ‘property rights,’ (D’Andrade, 1984), and, . . . ‘learning

disabilities' . . . " (Mehan, 1996, p. 273). Despite this constitutive quality, discourses do not exist apart from the objects they construct nor are those objects external to discourse. Instead, discourses and their objects exist in a reflexive relationship, constituting and being constituted by each other (cf. Gubrium & Holstein, 1997).

Discourses define categories of people, their responsibilities and authorities. This characteristic is a slight revision of Chambon's definition since, I would argue, categories do not exist independent of discourses but are created and sustained by them. For example, whether some collection of people is categorized as a family will likely vary depending on whether one is operating from the discourse of fundamentalist Christianity in the U.S., certain Asian religious traditions, or radical feminism. Once categorized, their "responsibilities and authorities" can be explicated as a function of the discourse. This has obvious implications for evaluation as it creates the agents and objects of evaluation (such as evaluators, stakeholders, and outcomes) and guides their interaction.

Discourses involve power relations. The existence of multiple discourses and the lack of uniform criteria for applying them (which is inevitable since such criteria also are part of discourses) means that multiple discourses can exist simultaneously and compete with one another for dominance. Which understandings will prevail in a particular setting or which objects will be constituted depend on relations of power involving factors such as authority, resource control, and sanctions. These relations are associated not only with the people in a setting, but to institutional and organizational structures - themselves the products of discourses (e.g., the hierarchical organization of a hospital).

Foucault, perhaps more than anyone, highlighted the relationship between discourse and power and its institutional dimensions. He used the concept of "discursive fields" to identify

“competing ways of giving meanings to the world and of organizing social institutions and processes”(Weedon, 1997, p. 34). Competing discourses will not be equal in power nor in their political stances; that is, their support of or challenge to the status quo. For example, Weedon (1997) discusses the conservative discourse in which “family” is considered a natural unit of the social order with the primary responsibility to rear children. In this family, there is a gendered division of labor with the male in a position of authority. The dominance of this discourse is reflected in “the organization of society in family units [which] guarantees the reproduction of social values and skills in class and gender terms” (p. 37), and in its institutions such as the legal system and the welfare system. Weedon contrasts this dominant discourse with radical and socialist family discourses in which the family is an instrument of oppression of women. However, because the dominant discourse is inscribed in societal institutions, giving it enormous material advantages, these alternative discourses remains marginalized.

These characteristics of discourse pack a lot of meaning into one word, maybe too much. However, if we accept even some of them, the implications for evaluation are substantial. The remainder of my presentation addresses some of these.

Questions

When evaluation is considered a type of discourse, the questions asked both in and about evaluations change. For instance, instead of the typical focus on methodological rigor or the proper application of technique, questions turn to characteristics and implications of the discourse.

Questions invite others to participate in a discourse. Since questions themselves arise within particular discourses, their invitations are similarly located^{iv}. In this sense all questions are leading questions. For example, asking a question about the characteristics of a particular

psychiatric disorder is to enact a particular discourse in which the disorder exists, where psychiatric disease exists, where psychiatrists have particular authority, where health and illness have particular meanings and so on. Similarly, to ask whether a particular drug is effective in reducing smoking is to invite participants into the discourse of science in which efficient causality, determinism, subjectivity, physicians and patients are constituted. Thus, unless the question or questioner is rejected, the discursive field will be somewhat constricted by the question.

Because of their discourse enacting qualities, questions orient and guide evaluations and can have a considerable influence on their eventual findings. For example, Gubrium and Holstein (1997) discuss how the “what” questions of traditional ethnographers and the “how” questions of ethnomethodologists lead them to constitute different realities. Additionally, evaluations can be assessed by asking questions of their questions, not only their content, but how they function discursively. Below are some examples of questions that might be asked when preparing or analyzing evaluations:

- 👉 How do questions function in this evaluation (for example, as hypotheses, as regulators of what can be said, as incentives to consider new perspectives)?
- 👉 What is the range of permissible (or intelligible) responses to the questions asked? What responses are invited or discouraged?
- 👉 How were evaluation questions identified and formulated; that is whose questions are these?
- 👉 Who asks the questions and who answers them?
- 👉 What is the relationship among questioners, respondents, and organization? Given the nature of that relationship, can the question or questioner be challenged?

- 👉 In whose language are questions formulated and expressed?
- 👉 What values, interests, and commitments do the questions express? (See Witkin, 1999 for further discussion of questions).

Asking these questions about an evaluation - or using them to guide the development of an evaluation - orients us toward relational and value issues that become infused with (and therefore often invisible within) evaluations. They can help reveal the discourse that is operating and how the language of that discourse structures and generates its “findings.”

Given the importance of questions, I have wondered why they are given so little attention in research and evaluation texts relative to other topics. A scan of the indexes of several evaluation texts revealed a much greater likelihood of finding “questionnaire” than “question.” My hunch is that this situation might be related to the uneasy fit between developing questions and a method-based approach to inquiry. The creative dimension of questions gives them a disruptive potential. By limiting questions to certain people such as “experts,” or making them conform to certain structural characteristics such as did x cause y, the discourse-invoking nature of question can be restricted. (Witkin, 1999). In contrast, what Raskin (1987) calls “dumb-dumb” questions - “questions which others are too expert or too socialized in a particular discipline, craft, or technology to ask” (p. 280) - can invite alternative discourses. Such questions are often met with incredulity or dismissed as irrelevant or naive. Children often have this experience when they ask “innocent” questions, that is, questions that do not presume the dominant discourse, leaving adults speechless.

An illustration of the “dumb-dumb” question approach to evaluation was demonstrated by one of my former students, Henry Palmer, who was asked by a local community mental health center to conduct an evaluation of a program for people with chronic psychiatric

disabilities. Rather than develop a list of questions that he or the agency was interested in, he went directly to the service users and, after explaining what the agency had requested, asked them what would be the important questions to ask. To his surprise, they responded with a question of their own, “Why does [the agency] want to know this?” Other questions emerged from the ensuing dialogue such as “How can I change case managers?” Such “practical” questions revealed what was important to the recipients of this service, what information they lacked, and provided clues about how the agency was perceived and operated.

Description

In a discussion of how “the events of discourse” are described, Foucault posed a question that provides another example of the disruptive or alternate discourse generating potential of questions. He asked: “how is it that one particular statement appeared rather than another?” The realist or empiricist response to this question might be that the statements chosen most accurately reflect the state of affairs under investigation. When evaluation is seen as discourse, however, the criterion of descriptive accuracy becomes suspect^v. Any description is considered as one way among many ways of construing a situation. The particular description chosen is not demanded by the brute facts of reality, but will vary in relation to social factors^{vi}. Gubrium and Holstein (1997) put it this way, “Descriptions must make sense; they must convince socially defined, culturally competent listeners that the objects, actions, or events in question warrant the attributions and characterizations that are bestowed upon them” (p. 132). Thus, one’s skills with language and rhetoric may be the most critical factors in the plausibility and acceptance of description. The great American writer John Steinbeck illustrated this eloquently in his description of ichthyological research in his book, *Log From The Sea of Cortez* (1941):

The Mexican Sierra (a fish) has 17 plus 15 plus 9 spines in the dorsal fin. These can be easily counted. But if the Sierra strikes hard on the line so that your hands are burned, if the fish sounds and nearly escapes and finally comes in over the rail, his colors pulsing and his tail beating the air, a whole new relational externality has come into being—an entity which is more than the sum of the fish plus the fisherman. The only way to count the spines of the Sierra unaffected by this second relational reality is to sit in a laboratory, open an evil-smelling jar, remove a stiff colorless fish from the formalin solution, count the spines and write the truth There you have recorded a reality which cannot be assailed—probably the least important reality concerning either the fish or yourself.

It is good to know what you are doing. The man with his pickled fish has set down one truth and has recorded in this experience many lies. The fish is not that color, that texture, that dead, nor does he smell that way (pp. 2-3).

Another dramatic (and somewhat amusing) example of this can be found in a little book by Queneau (1981) in which he provides 195 descriptions of the same event.

Describing social interactions, particularly in the contexts within which social work evaluations typically occur, requires complex judgments. For example, calling a parent's response to a child a "rebuke," a "lesson," a "personal attack," a "back-handed compliment," or "praise" may involve consideration of multiple factors such as the context of the interaction, knowledge of and past experience with the family, and favored theories^{vii}. Typically, whatever word or words are used to "describe" what happened will support certain values, that is, there will be an implication that the action was good or bad, healthy or unhealthy, for example, that the rebuke or praise was or was not appropriate. Similarly, using descriptive adjectives such as "domineering," "passive," "dependent," "assertive," "friendly," "uncooperative," "cooperative,"

“distant,” “detached,” “depressed” and so on - almost always imply a value stance in which the adjective used is one side of a dichotomous relation. Therefore, if descriptions involve choice, and if that choice is communicated in language, such choices invariably will imply values.

Sometimes the values contained in descriptions are implicit or indirect as when a statement is made as if it were obviously true or taken-for-granted. For example, in a popular text on human behavior, the authors state, “Children must begin the long process of moving in the direction of independence and separation from parental figures” (Berger, McBreen, & Rifkin, 1996, p. 141). One way to read this sentence is that independence and separation from parental figures are necessary for healthy development. Alternatively, but not necessarily exclusively, independence and separation and their implied relationship to health could be interpreted as value positions of the authors, or suggesting a male-oriented developmental perspective.

Accepting values as an inevitable part of descriptions implies that descriptions may function as prescriptions. This possibility is troublesome to empiricists who use various research designs to try to eliminate values from descriptions (or at least minimize their impact). For the discursive-oriented evaluator, however, the inevitability of values does not pose the same problem. Becoming aware that descriptions involve choice means that they - and the values they imply - can be otherwise. Thus the task becomes not to eliminate values, but to employ them in our inquiries in ways that are consistent with the values and mission of social work (or our visions of the good life)^{viii} and to be explicit about how we do so.

Authority, Rhetoric and Representation

If descriptions cannot derive their authority by claiming to be accurate depictions of “what is,” then from where does their authority come? From a discursive perspective, textual authority is attributed on the basis of social factors such as the perceived expertise of

authors/speakers and from institutional jurisdiction - themselves products of discourse - over particular knowledge areas. Authority also can be constructed by how a discourse is structured.

My talk today illustrates one dimension of authorization. Most of you know, or will be told, or will have read that I hold an advanced degree, that I am a full professor, that I have published widely, and that I have held prestigious positions. In addition, you are probably aware that this is a plenary session (given that there are no other presentations to attend at this time) which implies professional distinction (or that I am friends with the organizers). Other contextual factors such as the university sponsorship and prestige of this conference further contribute an air of authority to my talk, even before I have uttered a word.

Notice too the complementarity between context and personal characteristics. You would not be impressed (or even puzzled) if my qualifications for this talk included being a graduate of a culinary institute and owning a dog. Such information is not part of the discourse of authority relevant to my presentation (although, you may have noticed in my introduction that I tried to change the discourse a bit by suggesting that some aspects of my “personal” life *are* relevant). However, if we were dining at a restaurant rather than attending a professional conference, my culinary experience would be relevant (and academic degrees and publications meaningless). Thus, in a way the setting calls forth or makes visible certain personal characteristics that lend authority to my words. In a similar way, evaluators’ credentials and institutional auspices authorize their discourses^{ix}.

Authority also can be produced by a text’s structure and the use of tropes and other literary devices. The format of articles in professional journals and the use of particular writing styles such as that of the American Psychological Association create an authoritative text. The

latter does this by reproducing the discourse of science generating what Billig (1988) calls “depopulated texts.”

Dorothy E. Smith has described how various literary devices may be used to authorize a text. She identified “practices of objectification” that create the impression that a description is factual rather than mere opinion. One way this is done, according to Smith (1997), is by “suspending the presence of the subject” by converting verbs expressing subjects’ actions into nouns (called nominalization). Thus, instead of describing how someone “does” depression or hurts family members, we discuss depression and family violence. A related strategy is to convert subjective states of persons (e.g., attitudes or opinions) into entities “that can interact with other entities” (p. 59). Thus, an evaluation report might discuss the relationship of attitudes to beliefs. In a related literary practice termed “reattributing agency from subject to social phenomena,” Smith notes that “Once nominalized social phenomena are constructed, agency can be attributed to them rather than to people” (p. 59); for example, attitudes, beliefs, and the like may be attributed causal properties.

When people are constructed as categories, it is common to treat those categories as real entities and to assign (via research) characteristics or attributes to them. (Smith calls this “reconstructing subjects as figments of discourse” p. 61.) This is a common practice in social and psychological discourse. Most formal psychological measurement is based on the “existence” of such entities whose characteristics are then measured. In contrast, the discourse-oriented evaluator, rather than seeking referents or characteristics of a particular term, such as “borderline personality,” might (following Foucault) analyze the term as a discursive phenomenon, “as a thing brought into speech by the workings of power” (Shapiro, 1987, p. 369).

In mainstream scientific discourse, authority is related to facticity and objectivity. A common way of constructing objectivity in inquiry is through the use of “distancing devices” which separate the objects of study from the ways they are constituted in our representations of them. For example, the use “exemplary extracts of informants’ comments or conversations” in ethnographies separates the author’s commentary from the lives which are her/his subject matter (Gubrium & Holstein, 1997, p. 91, see also Atkinson, 1990). Statistics in the empiricist tradition may function similarly - turning readers attention away from the possibility that the research text constitutes its subject matter.

In a similar vein, Gergen (1994) discusses “distention devices,” ways of using language that create a separation between subjective experience and its linguistic referents. Distention language can be as simple as using words like “the” and “that” instead of “my,” or by using “distending metaphors” such as using words like “found,” “detected,” and “discovered” to imply that learning about the world is like searching for “buried treasure.” Gergen demonstrates the rhetorical impact of such metaphors by contrasting their use with more personalized language; for example, “Smith discovered the fact” versus “Smith labeled his impression” or “Jones found that . . .” versus “Jones selected new terms for his experience” (pp. 174-175).

These rhetorical devices - and I have only scratched the surface of this topic - not only authorize texts (such as evaluation reports), but help maintain the very characteristics of authority such as objectivity, that the discourse creates. By implication, they also support the metatheory of subject-object dualism. Critical, literary readings can interrupt these self-authorizations and increase a text’s potential interpretations.

Concern with issues of representation has led to exploration of different literary styles and presentation formats. Although a discussion of these efforts is beyond the scope of this

paper, some examples have included writing forms such as personal essay, memoir, autoethnography, dialogue, and poetry. Conjoint or distributed representations in texts in which researchers and participants or other relevant voices co-construct a text have begun to appear (Gergen & Gergen, 2000). In addition, performance texts in which participants literally “act out” their study within a dramaturgical context have been reported (for example, Ellis & Bochner, 1992).

A compelling concern for those for whom representation issues are salient is its political aspect. Social scientists and evaluators inevitably wind up speaking for (and creating) others. What has become increasingly clear over the past several years, is that those representations are often not how those being represented would choose to portray themselves. This issue is particularly relevant for social work evaluators whose inquiries involve people who historically have been least able to represent themselves.

The problem of criteria

The issue of representation also relates to the criteria used to judge knowledge claims. How an evaluation is described and reported will be influenced by the criteria believed most relevant to such claims. Demonstrating how an evaluation meets those criteria will increase its authority. Whether we subscribe to validity or verisimilitude, we will construct our reports in ways in which important evaluative criteria are salient.

Although by undermining the notion of a final, external authority, a discursive standpoint complicates the issue of criteria, we cannot eliminate the need to make judgments. Few of us (I hope) are willing to embrace an “anything goes” or “nothing goes” strategy that would leave “what goes” to those with the most power. However, viewing inquiry and evaluation as discourse raises new challenges for how to adjudicate among knowledge claims or even how to assess such

claims. Undergirding this challenge is the shift in perception of evaluation as a truth-bearing or truth-discovering enterprise to an activity that primarily is moral and political in nature^x (Smith & Deemer, 2000). This shift has led to the identification of criteria that are sensitive to the social and practical aspects of evaluation. For example, Chambers (2000) contends that criteria of utility are at least as important as more traditional ones such as validity. He identifies five such criteria: accessibility, relevance, responsiveness “to different claims on the *significance* of a course of action,” credibility, and the extent to which a study addresses “matters of *prospect* and judgment” (p. 863)^{xi}. It is important to note that all of these criteria are applied in reference to stakeholders and client groups.

In recent years the increasingly vocal demands and “counterstories” by groups who have felt unrepresented or misrepresented by traditional research and evaluation have highlighted the moral dimension of evaluative criteria. One alternative has been the development of “standpoint epistemologies” in which the starting point of inquiry or interpretation is located in the gendered, racial, sexual, or ethnic experience of the researcher, critic, subject, or author. These inquiries have generated new understandings of marginalized groups that stand in contrast to “Eurocentric, masculinist” representations of their lives (see Denzin, 1997).

My own interest in the moral dimension of evaluation was first explored in a paper entitled, “Alternative Criteria for Theory Evaluation” published in 1988 with Shimon Gottschalk (Witkin & Gottschalk, 1988). At the time, we were concerned about what we saw as fairly narrow and rigid criteria for assessing whether a particular theory (using that term loosely) was potentially valuable for social work. In particular, we were concerned with the absence of considerations for the values and ideals that to us made social work both unique and important, issues like social justice, context, agency, and the legitimation of the life experiences of the

people social workers serve. I continued to explore these ideas in two other papers in which I proposed the use of human rights as evaluative criteria (Witkin, 1993, 2000). Within this framework, knowledge claims would be judged, among other ways, with regard to their protection, advancement, weakening, or thwarting of human rights. For example, explanations of, or responses to, social problems which justified restrictions on freedom and well-being (posited as fundamental human rights) would be less favored than those which might expand or protect these rights.

These papers represent one attempt to reframe the criteria issue in a way that addresses the moral dimension of evaluation. Others have spoken eloquently to this issue. For example, Lincoln (1995) asserts the need for a “vision of research that enables and promotes social justice, community, diversity, civic discourse, and caring” (cited in Smith & Deemer, 2000, p. 892). Christians (2000) proposes a feminist communitarian model that requires research to meet the criterion of “interpretive sufficiency” which he describes as a discourse that “represents multiple voices, enhances moral discernment, and promotes social transformation” (p. 145).

Different evaluative criteria generate different interpretations within and across discourse communities. Although the existence of multiple criteria may mean increased competition for recognition, I hope that we can heed Smith & Deemer’s (2000) caution that “If we cannot talk and listen to each other, it is difficult to imagine why anyone else would want to talk with us, listen to us, or attend to our judgments” (p. 893).

Conclusion

In this presentation I have attempted to identify some issues related to the shift from evaluation as method to evaluation as discourse. Although my treatment of these issues was

brief, my hope was to create enough interest in the topic so that some of you will consider how this perspective may enhance the practice and interpretation of evaluation.

Although the shift to discourse can feel like a slide into nihilism, intellectual anarchy, or immobilization, it need not be so. In fact, we can look at this change as a way of enhancing our evaluation activities. Viewing language not “as a neutral carrier of meanings or a mere transparent medium of facts . . . [but as] . . . the constitutive method and material of the world that it projects . . . [means that]. . . the way that we talk about the world [is] as important as the objects of the worlds that, in talking about them, become available as objects of our experience” (Brown, 1990, p. 72). Thus, consideration of linguistic and textual practices and their relationship to evaluation can, in my opinion, enrich the conversation around evaluation, broaden our understanding, increase our sensitivity to topics and people that may have been invisible or silent, and align our practices more closely with our values and commitments. It can also expand the range of practices available to evaluators.

If evaluation is a moral activity, then evaluators have an obligation to be self-conscious about their representational practices given their privileged position as assessors and decision makers. Drawing on Foucault, Gubrium and Holstein (1997) argue that, “Because we speak and write a discourse of scientific ‘truth’ and there are substantial institutional arrangements that lend credence to what we say, self-consciousness obligates us to reflexively deconstruct our own ‘truths’ and consider the power that resides in producing and owning knowledge” (p. 111). Such consciousness can “encourage social scientists to undertake forms of analysis that avoid the uncritical valorization of the realities created by dominant, ‘official’ modes of discourse” (Shapiro, 1987, p. 366). We can address this issue of self-consciousness by being open to and

experimenting with diverse forms of inquiry and representation, and by expanding our evaluative lexicon and modes of expression.

In closing, I want to note some of the important work being done in social work research and evaluation that resonates with the ideas that I have expressed. These include, for example, Christopher Hall's book on *Social Work as Narrative* (1997), Mary Katherine Rodwell's book on *Social Work Constructivist Research* (1999), Bob Broad's edited book, *The Politics of Social Work Research and Evaluation* (1999), Carol Truman's, Donna Mertens', and Beth Humphries' edited book, *Research and Inequality* (2000), and two collections from Finnish authors: *Constructing Social Work Practices* (1999) edited by Arja Jokinen, Kirsi Juhila, & Tarja Pösö, and *Reconstructing Social Work Research* (1999) edited by Synnöve Karvinen, Tarja Pösö, and Mirja Satka. In my view, these works demonstrate the promise of expanding the boundaries (conceptually, methodologically, and politically) of our evaluative practices. Increasing the diversity of intellectual resources available to evaluators can help us to work more sensitively and collaboratively with others. If we can do this, our evaluations, our services, and ultimately our clients, will benefit.

Notes

ⁱ The fact that phenomena like humor are not considered relevant to activities like evaluation can itself be viewed as a characteristic of academic or research discourse and how it influences what it is possible to know. In contrast, the Russian literary critic, Mikhail Bakhtin (1984) has stated that "Laughter has deep philosophical meaning, it is one of the essential forms of truth concerning the world as a whole, concerning history and man; it is a peculiar point of view relative to the world; the world is seen anew, no less (and perhaps more) profoundly than when seen from the serious standpoint" (p. 66).

ⁱⁱ These ideas are expressed in the following joke: "What do you get when you cross a poststructuralist with a mafioso? Answer: An offer you can't understand" (p. 188).

ⁱⁱⁱ For example, one can play a game, play a musical instrument, play a cd, or play Macbeth *in* a play. We can play down, play out, play for, play on, play into, or play it. We can play with fire,

play dead, in which case we would be playing possum, play with others or with oneself, or play with a full deck which, by the way, is not about playing cards (neither the game nor the objects).

^{iv} Of course, one may decline an invitation or respond in an unexpected way thereby challenging the hegemony of the discourse. However, in many social work settings respondents (that is, clients) may not be free to challenge the discourse that is implied by a question.

^v Accuracy is itself seen as the discursive expression of the belief that linguistic representations of reality can be separated from an extralinguistic reality. Of course, this position also is part of a discourse. Thus, it is not the discursive nature of these beliefs that are troublesome, but rather the implications of accepting or denying such a view.

^{vi} Clearly, one is not free to say anything; however, the restrictions on what may be said are socially influenced, for example, language conventions, power relations, and social context.

^{vii} Even the description “response” is a judgment that imposes a temporal sequence on the interaction.

^{viii} This position does not mean that one can say anything. Our utterances still will be judged by various criteria depending on the language community to which one is aiming. Also, it seems reasonable to presume that professionals are guided by a sense of ethics such as honesty which certainly are not suspended in this case.

^{ix} These discourses can be challenged, particularly by others claiming their own personal and contextual authority, for example, a psychiatrist in a mental health clinic.

^x I hasten to add that nothing in this conceptualization compels me to give up having standards by which to judge a program or the evaluation of that program. What changes is the way I do this as was illustrated previously in my discussion of questions.

^{xi} Although Chambers’ concern is applied ethnography, I believe his criteria apply to evaluation studies of various persuasions.

References

- Ashley, D. (1997). *History without a subject: The postmodern condition*. Boulder, CO: Westview Press.
- Atkinson, P. (1990). *The ethnographic imagination*. London: Routledge.
- Berger, R. L., McBreen, J. T., & Rifkin, M. J. 1996, *Human behavior: A perspective for the helping professions, 4th ed.* White Plains, NY: Longmans

-
- Billig, M. (1998). Repopulating social psychology: A revised version of events. In Bayer & Shotter (eds.) *Reconstructing the psychological subject*. pp. 126-152.
- Bloom, Fischer, and Orme, 1995
- Brown, R. H. (1990). Social science and the poetics of public truth. *Sociological Forum*, 5 (1), 55-74.
- Chambers, E. (2000). Applied ethnography. In N. K. Denzin & Y. S. Lincoln (eds.) *Handbook of qualitative research*, 2nd ed. (pp. 851-869), Thousand Oaks, CA: Sage.
- Chambon, A. S. (1999). Foucault's approach: Making the familiar visible. In A. S. Chambon, A. Irving, & L. Epstein, (Eds.) *Reading Foucault for social work* (pp. 51-81), NY: Columbia University Press.
- Christians, C. G. (2000). Ethics and politics in qualitative research. In N. K. Denzin & Y. S. Lincoln (eds.) *Handbook of qualitative research*, 2nd ed. (pp. 133-155), Thousand Oaks, CA: Sage.
- Denzin, N. K. (1997). *Interpretive ethnography*. Thousand Oaks, CA: Sage.
- Ellis, C., & Bochner, A. P. (1992). Telling and performing personal stories: The constraints of choice in abortion. In C. Ellis & M. G. Flaherty (eds.), *Investigating subjectivity: Research on lived experience* (pp. 79-101). Newbury Park, CA: Sage.
- Foucault, M. (1977). What is an author? In D. F. Bouchard (ed.), D. F. Bouchard & S. Simon, trans., *Language, countermemory, practice: Selected essays and interviews* (pp. 115-138), Ithaca, NY: Cornell University Press.
- Foucault, M. (1979). *Discipline and punish*. Harmondsworth: Penguin.
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. *American Psychologist*, 40, 266-275.

-
- Gergen, K. J. (1994). *Realities and relationships: Soundings in social constructionism*. Cambridge, MA: Harvard University Press.
- Gergen, M. M., & Gergen, K. J. (2000). Qualitative inquiry: Tensions and transformations. In N. K. Denzin & Y. S. Lincoln (eds.) *Handbook of qualitative research, 2nd ed.* (pp. 1025 - 1046). Thousand Oaks, CA: Sage.
- Gubrium, J.F. & Holstein, J. A. (1997). *The new language of qualitative method*. NY: Oxford University Press.
- Hanson, N. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Humphries, B., Mertens, D. M., & Truman, C. (2000). Arguments for an 'emancipatory' research paradigm. In Truman, C., Mertens, D. M., & Humphries, B. (eds.) *Research and inequality*. London: UCL Press.
- Karvinen, S. (1999). The methodological tensions in Finnish social work research. In S. Karvinen, T. Pösö, & M. Satka (Eds.) *Reconstructing social work research*. Jyväskylä: SoPhi.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Mahoney, M. J. (1976). *Scientist as Subject: The Psychological Imperative*. Cambridge, MA: Ballinger.
- Mehan, H. (1996). The construction of a LD student: A case study in the politics of representation. In Silverstein, M., & Urban, G. (eds.) *Natural histories of discourse*. Chicago: University of Chicago Press.
- Newman, F. & Holzman, L. (1997). *The end of knowing*. NY: Routledge.
- Queneau, R. (1981). *Exercise in style*. NY: New Directions.

-
- Raskin, M. G. (1987). Conclusion: A Manifesto of Reconstructive Knowledge. In M. G. Raskin & H. J. Bernstein (Eds.) *New ways of knowing: The sciences, society, and reconstructive knowledge*. Totowa, NJ: Rowman & Littlefield.
- Shapiro, M. J. (1987). The rhetoric of social science: The political responsibilities of the scholar. In J. S. Nelson, A. Megill, & D. N. McCloskey (Eds.) *The rhetoric of the social sciences*. Madison, WI: The University of Wisconsin Press
- Smith, D. E. (1999). *Writing the social: Critique, theory, and investigations*. Toronto: University of Toronto Press.
- Smith, J. K. & Deemer, J. K. (2000). The problem of criteria in the age of relativism. In N. K. Denzin & Y. S. Lincoln (eds.) *Handbook of qualitative research, 2nd ed.* (pp. 877-898), Thousand Oaks, CA: Sage.
- Steinbeck, J. (1941). *Log from the sea of cortex*.
- Surber, J. P. (1998). *Culture and critique*. Boulder, CO: Westview Press.
- Thyer, B., Harrison, D. F., & Hudson, W. (1992).
- Weedon, C. (1997). *Feminist practice and poststructuralist theory, 2nd ed.* Oxford, UK: Blackwell.
- Weimer, W. (1979). Notes on the methodology of scientific research. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Witkin, S.L. (1993). A human rights approach to social work research and evaluation. ____ *Journal of Teaching in Social Work, 8, 1/2, 239-253*.
- Witkin, S. L. (1996). If empirical practice is the answer, then what is the question? *Social Work Research, 20, 69-75*.
- Witkin, S. L. (1999). [Editorial] Questions. *Social Work, 44, 197-200*.

Witkin, S.L. (2000). An integrative approach to human rights and social research. In C. Truman,

D. M. Mertens, & B. Humphries (Eds.) *Research and inequality*. London: UCL Press.

Witkin, S.L. & Gottschalk, S. (1988). Alternative criteria for theory evaluation. *Social Service*

Review, 62 (2), 211-224.