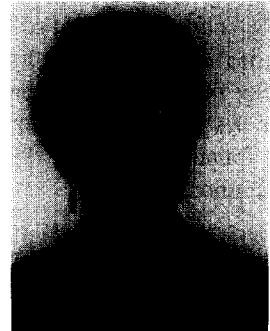


Inclusive Evaluation: Implications of Transformative Theory for Evaluation

DONNA M. MERTENS

ABSTRACT

Evaluators face a challenge in responding to a call for greater inclusiveness of marginalized groups. In this presentation, I examine the contribution that transformative theory can make toward meeting this challenge. Transformative scholars assume that knowledge is not neutral, but is influenced by human interests, and that all knowledge reflects the power and social relationships within society, and that an important purpose of knowledge construction is to help people improve society. I propose the use of an inclusive model of evaluation that can address the tension between what is needed to accurately represent the experiences of marginalized groups and the traditional canons of research.



Donna M. Mertens

INTRODUCTION

Evaluators work in a world of multiple demands—sometimes seemingly conflicting demands that emanate both from the complexity of the programs that we evaluate as well as from the standards, principles, and canons that were generated from within our professional culture to guide us in our work. We have standards (*The Program Evaluation Standards: How to Assess Evaluations of Education Programs*, Joint Committee on Standards for Educational Evaluation, 1994) that direct us to identify the stakeholders, defined as those persons involved in or affected by the evaluation, so that their needs can be addressed. We have guiding principles (*Guiding Principles for Evaluators*, Shadish, Newman, Scheirer, & Wye, 1995) that call upon us to recognize and be sensitive to cultural differences (see Mertens, 1995 for a further explanation of this point) in order to fulfill the evaluator's obligation to consider the public good (see Schwandt, 1997 for a further explanation of this point). At the same time, we have the canons of research that emphasize the search for truth, objectivity, credibility, and validity, all the

Donna M. Mertens • 4600 Marie Street, Beltsville, MD 20705; (301) 937-8282; (202) 651-5545; E-mail: donna.mertens@gallaudet.edu.

American Journal of Evaluation, Vol. 20, No. 1, 1999, pp. 1–14.
ISSN: 1098-2140

All rights of reproduction in any form reserved.
Copyright © 1999 by American Evaluation Association.

while warning against the dangers associated with the appearance of advocacy, political motivations, and ontological nihilism.

In this presentation, I examine the tensions created by these forces that challenge evaluators, especially as we address the need to reach out to marginalized groups. I explore the ways that transformative theory can both inform the theory that guides our work, as well as enable us to design inclusive evaluations directed toward building a stronger bridge between evaluation findings and the transformation of society. Throughout this paper, I argue that inclusive evaluations have the potential for stronger objectivity (Harding, 1993) and increased rigor because they are inclusive. First, I address the need to include marginalized groups, and the dangers associated with the appearance of advocacy. Second, I provide an explanation of transformative theory and inclusive evaluation. Third, I discuss the tensions between inclusion and the canons of research. And finally, I provide specific examples of the application of an inclusive evaluation approach that can support the efforts of evaluators who seek to challenge the status quo, a contribution that Chelimsky described as "our most important task and the best justification for our work" (1998, p. 37). More extensive examples of applications can be found in Mertens (1998), Truman, Mertens, and Humphries (in press), and Cousins and Whitmore (1998).

Need to Include Marginalized Groups and the Dangers Associated with Advocacy

Calls for greater inclusiveness of marginalized groups can be found in a large body of scholarly work from those outside of the professional evaluation community (Stanfield, 1998; Gill, 1998), as well as from evaluation-community insiders (Chelimsky, 1998; Kirkhart, 1995; Greene, Lincoln, Mathison, Mertens, & Ryan, 1998; Mertens, 1998; Schwandt, 1997; Weiss, 1998). In the printed versions of their plenary addresses at the 1997 American Evaluation Association meeting, both Weiss and Chelimsky wrote of the need for greater inclusiveness of marginalized groups (Weiss, 1998; Chelimsky, 1998).

Weiss commented on the tendency of evaluators to limit participation in evaluations to program staff. She acknowledged that it makes sense to include program staff, but argued that program clients also have a big stake in the program. She wrote:

I wish their inclusion were more widespread, especially clients from marginal groups in society. Difficult as it is to engage them in evaluation activities, they are apt to have different interests and concerns from those of staff and addressing their questions would broaden the scope of the study. It would also bring people with different orientations and perspectives together to hear and come to understand each other's points of view. Including clients in the evaluative process would also help redress the inequalities in access to influence and the imbalance of power that beset marginalized groups in their social interactions with the larger society. (Weiss, 1998, p. 30)

In Chelimsky's discussion of the role of experience in formulating theory for evaluation, she noted one

...implication of assimilating history, values and past research into the evaluation process is the need for a certain inclusiveness. This is to ensure that the new evaluation will be complete and balanced enough both to achieve credibility in the current political environment, and also to pass muster in the next. One aspect of such inclusiveness involves awareness not only of highly publicized stakeholder positions with regard to a policy or program, but also of stakeholder positions that are less well-known because they are not represented

by an organized lobby. This is often the case for program beneficiaries such as children, patients, poor people, immigrants, or the severely handicapped. Their views are important for an understanding of the program, but they are often overlooked by evaluators.

Paying attention to what the beneficiaries of a program think about it is a hallmark of a credible study, and has nothing to do with advocating for those beneficiaries. Instead, it does three things for an evaluation. First, it aids impartiality and even-handedness by counter-balancing management biases with beneficiary biases. Both sets of views need to be understood and validated, not championed. Second, because so much research has failed to deal with beneficiary views in the past, an evaluation that does so may well produce critical new information. Finally, since beneficiaries know more, from personal experience, about the qualities and inadequacies of a program than anyone else, learning what they know greatly improves the sensitivity of an evaluation. (Chelimsky, 1998, p. 46-47)

I want to extend the list of benefits expected from an inclusive approach to evaluation to strengthen my argument that such an approach is needed. In addition to aiding in impartiality, gaining new information, and improving the sensitivity of evaluation, I would add three other reasons for the adoption of an inclusive approach. First, it allows us to fulfill our moral and ethical obligations as explicated in the Guiding Principles (Shadish et al., 1995) under the category: "Responsibilities for General and Public Welfare" that states: "Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare" (p. 25). Second, we live in a symbiotic relationship with each other on this planet. When serious social problems are allowed to fester and worsen, there is a spill-over effect for all of us. What happens to one group affects others, possibly in terms of higher taxes, increased crime, decreased quality of life or, for some, increased moral angst. Third, without an inclusive approach, a realistic description of the program will not be attained, and without that, the wrong conclusion may be reached.

It was clear to me in Chelimsky's message that advocacy in the form of having an ax to grind, or showing partiality in the sense of advancing the interests of specific groups, has no place in evaluation. Weiss addressed this issue with a similar message:

Use is about change...But it is not change any which-way the evaluator wants. It is not supposed to be change based on the evaluator's ideological or political predilections. That is not why people invited us onto the scene. They assume that the direction for change that comes out of the evaluation will be based on what the evaluator finds out. Listen to that statement: what the evaluator finds out. (Weiss, 1998, p. 31)

I am in complete agreement that the evaluator should avoid advocacy when it is defined as "...taking the views or interests of one group and always championing them over others, regardless of the findings of the evaluation" (House & Howe, 1998, p. 235). However, in order to be inclusive of marginalized groups, evaluators need to struggle with issues of oppression, discrimination, and power differences, and still walk the tight rope of avoiding the bias associated with a position of advocacy (Gitlin, 1994). Greene (1997) presents eloquent arguments for advocacy of values associated with democratic pluralism, and House and Howe (1998) support evaluators serving as advocates for democracy, the public interest and an egalitarian conception of justice. I do not want to see the evaluation community avoid dealing with these important issues by using the labels of advocacy, political motivation, or ideological or ontological nihilism to dismiss efforts to represent validly the viewpoints of those with the least power. I do not want to see the evaluation community use the terms objectivity and neutrality to mask dismissal of issues related to social injustice, discrimination, and oppression.¹

Transformative Theory and Inclusive Evaluation

I now turn to my second topic, what I mean by transformative theory and inclusive evaluation. The good news is that we do not have to start from ground-zero in our struggle to meaningfully involve marginalized groups. There actually is a large body of scholarship that is only rarely represented in research or evaluation methods or theory books, that has struggled directly with the accurate and credible representation of marginalized groups in and through the research process. Some of these writings fall into the postmodern camp which may indeed border on, or whole heartedly embrace, absolute relativism and ontological nihilism. I reject such a position, and will focus on the contributions from the literature that have traveled under a number of different names, one of these being transformative theory.

Transformative scholars assume that knowledge is not neutral, but is influenced by human interests, that all knowledge reflects the power and social relationships within society, and that an important purpose of knowledge construction is to help people improve society (Banks, 1993). Transformative theory is used as an umbrella term that encompasses paradigmatic perspectives such as emancipatory (Lather, 1992; Mertens 1998), anti-discriminatory (Humphries & Truman, 1994; Truman, Mertens, & Humphries, in press); participatory (Reason, 1994; DeKoning & Martin, 1996; Cousins & Whitmore, 1998); and Freirian approaches (McLaren & Lankshear, 1994), and is exemplified in the writings of feminists (Alcoff & Potter, 1993; Fine, 1992; Hill-Collins, 1990; Reinharz, 1992), racial/ethnic minorities (Stanfield & Dennis, 1993; Madison, 1992), people with disabilities (Oliver, 1992; Mertens & McLaughlin, 1995), and people who work on behalf of marginalized groups.

The *W. K. Kellogg Foundation Evaluation Handbook* (1998) contains a description of the evaluation landscape that is historically dominated by the scientific method paradigm (elsewhere known as postpositivist, Mertens, 1998). From the Foundation's viewpoint, over-reliance on the scientific method has resulted in evaluations that have not addressed issues of process, implementation, and improvement as well as they should. The real concern is that "conventional evaluations...may very well be negatively impacting the more complex, comprehensive community initiatives...because these initiatives are often ignored as unevaluable, or evaluated in traditional ways that do not come close to capturing the complex and often messy ways in which these initiatives effect change" (Connell, Kubisch, Schorr, & Weiss, 1995; Schorr & Kubisch, 1995, cited in the *W.K. Kellogg Foundation Evaluation Handbook*, 1998, p. 6). Their recommendations for a better balance is for evaluators to learn about and reflect on alternative paradigms and methods that are appropriate to our work. Their list of alternative paradigms and methods includes the interpretive/constructivist paradigm, as well as feminist methods and participatory evaluation. I use the Kellogg Foundation's work as an example of the emergence of a call within the evaluation community to be inclusive and aware of the influence of different paradigms on the type of work that we do.

The transformative paradigm is characterized as placing central importance on the lives and experiences of marginalized groups, such as women, ethnic/racial minorities, people with disabilities, and those who are poor. The evaluator who works within this paradigm consciously analyzes asymmetric power relationships, seeks ways to link the results of social inquiry to action, and links the results of the inquiry to wider questions of social inequity and social justice (Mertens, 1998; Mertens, Farley, Madison, & Singleton, 1994; Truman et al., in press).

The transformative paradigm is based on key ontological, epistemological and methodological assumptions that are different from those underlying those of the postpositivist and interpretive/constructivist world views. The ontological question, "What is the nature of real-

ity and by extension, truth?" is answered differently by scholars who align themselves with the three paradigms. The postpositivist paradigm's ontological view holds that there is one reality, one truth that can be known within a certain level of probability. The interpretive/constructivists have argued for the recognition of multiple socially constructed realities, and are thus subjected to the criticism of being mired in absolute relativism such that no one perspective is any "truer" than any other perspective. The transformative ontological assumption holds that there are diversities of viewpoints with regard to many social realities, but we need to place those viewpoints within a political, cultural, and economic value system to understand the basis for the differences. And, then, we need to struggle with revealing those multiple constructions, as well as with the decisions about privileging one perspective over another.

In epistemological terms, the issue of objectivity is salient, along with implications for the nature of the relationship between the evaluator and the stakeholders in the program. In the postpositivist paradigm, objectivity is considered to be paramount and is thought to be achieved by observing from a somewhat distant and dispassionate standpoint. In the interpretive/constructivist paradigm, interaction between the evaluator and participants is felt to be essential as they struggle together to make their values explicit and create the knowledge that will be the results of the study. In transformative terms, objectivity is valued in the sense of providing a balanced and complete view of the program processes and effects such that bias is not introjected because of a lack of understanding of key viewpoints. However, to obtain this depth of understanding, it is necessary for the evaluator to be involved in the communities impacted by the program to a significant degree.

Finally, in methodological terms, the postpositivist paradigm is characterized as using primarily quantitative methods that are interventionist and decontextualized. The interpretive/constructivist paradigm is characterized as using primarily qualitative methods in a hermeneutical and dialectical manner. The transformative paradigm might involve quantitative, qualitative, or mixed methods, but the community impacted by the evaluation would be involved to some degree in the methodological and programmatic decisions. (Mixed methods designs that use both quantitative and qualitative methods can be used in any paradigm; however, the underlying assumptions determine which paradigm is operationalized.)

These philosophical assumptions form the basis for the definition of inclusive evaluation. Inclusive evaluation involves a systematic investigation of the merit or worth of a program or system, for the purpose of reducing uncertainty in decision making, and to facilitate positive social change for the least advantaged. Thus, inclusive evaluation is data-based, but the data are generated from an inclusive list of stakeholders, with special efforts to include those who have been traditionally under-represented. It does not exclude those who have been traditionally included in evaluations, i.e., the decision makers, intended users, program administrators or staff, and funding agency representatives. It does explicitly recognize that certain voices have been absent, mis-represented, or marginalized, and that inclusion of these voices is necessary for a rigorous evaluation. Inclusive evaluation does borrow ideas from scholars who have struggled with the representation of marginalized voices and applies those ideas in the evaluation context. Conclusions are based on the collection, analysis, and interpretation of inclusive data and are not fore-gone conclusions.

Tensions Between Inclusive Evaluation and Canons of Research

Truth, objectivity, credibility, validity, and rigor are fundamental issues for evaluators. Inclusive evaluation has the potential to contribute to an enhanced ability to assert truth,

objectivity, credibility, validity, and rigor in the sense that ignored or misrepresented views are included. How does adherence to the principle of objectivity find itself on the opposite side of the fence from addressing the needs of marginalized and less empowered groups? If the heart of objectivity is to avoid bias, then it seems to necessitate inclusion of perspectives of all relevant groups. If we have an awareness of historical conditions in our country and around the world, and understand current social conditions, then we know that certain groups of people have been under- or misrepresented in the world of social inquiry, whether it be evaluation or research. Which is the better role for the evaluator? Dispassionate and distanced, or passionately concerned with ascertaining the truth? And truth is defined as being inclusive of the perspectives of those with the lived experience with the problem, whatever it might be – spousal abuse, sexual abuse, poor educational service, or lack of equal access to the justice system.

A good evaluator would want to provide as accurate a picture as possible. When significant voices are missing, the picture is not complete and may actually be a distorted representation of reality. Based on a review of world history, certain groups have been systematically excluded from having meaningful participation in the design, implementation, and use of evaluations that impact them. Can a report be balanced when the voices of important constituencies are missing or inaccurately represented, or lost in the aggregation of data across groups?

Applications of Transformative Theory to Evaluation Work

Transformative theory holds implications for our evaluation theories and practice. As I previously mentioned, a tension exists in the evaluation community as to the proper role of evaluation in regard to recognition of marginalized voices and the accompanying questions of power and privilege, and the objective and neutral stance that is valued as a hallmark of evaluation practice. I do not see these as mutually exclusive roles, but as challenges for evaluators who want their work to be a useful part of bridging the gap between evaluation findings and social change.

Evaluation work is often concerned with the need for or quality of services and programs designed to address society's most pressing problems, ranging from welfare reform, drug abuse treatment and prevention, and school system failures. The construction of performance indicators, the determination of costs and the measurement of outcomes all involve the operationalizing of theories as to what is important to the success of a program. Who we talk to, what literature we read, whose opinions are given privilege in the determination of these operationalized tools of evaluation has an impact on our ability to represent the diverse voices of stakeholders. To illustrate the application of transformative theory in evaluation, I have chosen three examples from the process of planning and conducting an evaluation. They include: (1) determination of program theory; (2) sampling decisions, or knowing the stakeholders and how to involve them appropriately; and (3) development of data collection instruments and processes to facilitate social change.

Determining Program Theory

One of the first challenges that an evaluator faces in a new project is to discover the program theory that is operating. The inclusive evaluator would suggest that we need to examine the nature of the theories that the different constituents bring to the table and encourage a crit-

ical examination of those theories. Villegas (1991), Banks (1995), Oakes and Guiton (1995) and Mertens (1996; Mertens et al., 1994) describe theories that program administrators have used to explain such problems as poor academic achievement for ethnic/racial minorities, and sexual abuse of deaf students in a residential setting. According to these transformative researchers, some program theories are based on assumptions of individual or cultural deficits, others on cultural differences, and still others on a transformative theoretical framework that looks at the effects of power inequities.

The cultural deficit theory holds that there are deficits inherent in the cultural group that explain their poor performance or behavior. In a study of high school tracking, Oakes and Guiton (1995) found that a cultural deficit theory was operating amongst the teachers who believed that there was little hope for improvement once a student reached high school because the students either lacked essential basic skills or appropriate motivation. Racial groups were identified with specific tracks in most teachers' minds. The teachers reported believing that Latinos had low representation in higher-level courses because they were "culturally disinclined" to aspire to post-secondary education.

In the study I conducted on sexual abuse in a residential school for the deaf (Mertens, 1996), the cultural deficit theory was manifest in the following comments made by program administrators in their explanations of the conditions that permitted the sexual abuse to occur: "The girls behave in ways that suggest they are easy and loose in terms of their sexual behavior." Or, "People suspected but nobody would do anything. The deaf community knew. That's pretty typical residential school stuff." Or, "The background from their homes is not good. One child was in and out of foster homes; we try to work on it, we can't do miracles." These comments suggest that the problem is with the way the girls dress and act, or with the poor morals associated with members of the deaf culture. This blaming the victim or the culture ignores the responsibilities of the adults in the situation and can be used as an excuse for failing to take action, or taking inappropriate action. It also ignores the power inequities that exist between children and adults, and between students and administrators.

Banks (1995) explained that the cultural difference theorists explain academic failure as a result of serious cultural conflicts between the home culture and the culture of the school. In Villegas's (1991) analysis of minority children's academic achievement, cultural difference theory is manifest in programs that recognize such differences as the nature of language use at home and at school in terms of asking and answering questions or in help seeking behaviors. While the cultural difference theory leaves room for recognizing positive values in the home culture, it does not address the systemic power differences in the dominant culture as compared to minority cultures.

A transformative theory in a program would be based on recognizing the power inequities that are inherent in our society and determining their impact on the program outcomes. In my study of sexual abuse, power issues were raised in a number of different contexts. One power-related theory that surfaced concerned the role of the top level administrator in keeping "dirty laundry" hidden. Illustrative program staff comments included:

- "I think our administrator (the one who left) wanted to hear good things. He discouraged problems coming to him. People under him had learned not to go to him."
- "People wanted to keep a pretty picture; there was a reluctance to have bad news out."
- "The 'sweep it under the rug' attitude was very true. The problem was an administrator who did not want to be actively involved in solving the problems."

- “Possible repercussions if it was reported, fear of loss of job, pressure to not testify from the community.”

This group of comments could be interpreted as blame the ‘bad’ administrator. However, narrowly interpreting the problem as the poor administrative style of an individual ignores the power held by administrators in educational institutions. There needs to be an open acknowledgment of that power and how those without power can gain access to the lines of communication. This acknowledgment could be used as a basis for discussion to change conditions so that the probability of a repeat occurrence of this problem is reduced.

Raising issues of power and discrimination and oppression as possible forces that are operating in a program is not usually a comfortable process. Program administrators may be unwilling to entertain the notion that the program is structured in such a way as to perpetuate the status quo. Entering an evaluation setting with a sensitivity to program theories that reflect discrimination and oppression does not necessarily mean that you will find them. However, excluding such issues from consideration almost guarantees that they won’t be found even when they are present. Raising such issues makes demands on evaluators in terms of having an accurate grasp of the program itself, which requires a certain “closeness” to the community and certainly access to the perspectives of those with the least power in the program. It also requires a sensitivity to protecting those who are most vulnerable in the program.

Knowing the Stakeholders

I want to return to the *Standards* that relate to utilization and more specifically to the Standard that calls upon us to define those persons involved in or affected by the evaluation. The evaluator can choose to take the more conservative approach and limit interactions to the intended users of the evaluation, even if that means allowing the program administrators or funders to say that only funders and program staff need to be involved. I would argue that this does not allow the evaluator to hold to the *Guiding Principles* that direct us to channel our work to the greater public good.

To conduct an inclusive evaluation, it is necessary to both know the diversity within the community that is impacted by the evaluation, and appropriate ways for gathering data from diverse sub-groups. Stanfield (1993) warned against falling into the myth of homogeneity, i.e., that all members of a minority group share the same characteristics. In order to explode that myth, I suggest that we look at a modified version of the questions formulated by the Canadian Research Institute for the Advancement of Women (1996) in their ethical guidelines when we are designing our sampling plans. Questions we might ask ourselves include:

- “Are we including people from both genders and diverse abilities, ages, classes, cultures, ethnicities, families, incomes, languages, locations, races, and sexualities?”
- “What barriers are we erecting to exclude a diversity of people?”
- “Have we chosen the appropriate data collection strategies for diverse groups, including providing for preferred modes of communication?”

Asking ourselves such questions would serve to prod us as to examine the inclusiveness of our sample, and to consider barriers to participation that we might have otherwise overlooked.

In my work with improving court access for deaf and hard of hearing people in the United States, we conducted focus groups to determine problems and successful strategies that deaf

and hard of hearing people had experienced in the court systems (Mertens, in press). Although I have worked in this community for over 16 years, I was appreciative of the insights gained from interaction with our advisory board, which included people who preferred a variety of communication modes and represented different aspects of the judicial system. We prepared a careful sampling design in order to conduct focus groups with people who represented the diversity of communication modes in the deaf and hard of hearing communities. The five groups we included were:

- Group 1: Deaf users of American Sign Language, highly educated, professionals who used American Sign Language and could read English (with one oral deaf person who relied on voicing, lip reading, and reading English).
- Group 2: Deaf adults, with limited education and reading skills, some of whom could use American Sign Language, one who had very low language functioning and communicated through some sign language and gestures and pantomime, and a deaf/blind adult who used an American Sign Language interpreter at close range.
- Group 3: Hard of hearing adults, highly educated, good English reading skills, and used various personal assistive listening devices.
- Group 4: Deaf adults who used Mexican Sign Language, with little to no English reading skills or understanding of American Sign Language.
- Group 5: Deaf adults who rely on oral communication (watching an oral interpreter who carefully enunciates speech instead of signing) and who could read English.

Simply identifying these groups was not enough. We also had to determine the best way of collecting data from each group. Different configurations of technical support were needed to facilitate effective communication. For example, in the group in which the individuals had a low level of language functioning, we had hearing and deaf co-moderators, an American Sign Language (ASL) interpreter to sign for the hearing moderator and voice for the deaf participants, a deaf relay interpreter who translated the ASL signing into a combination of signs, pantomime, and gesture for one individual whose language functioning was too low to understand a pure ASL presentation, and a deaf-blind interpreter who signed into the hands of the visually impaired woman in the group. In addition, we had a court reporter observing the group and entering the comments of everyone into a written transcript that ran as real-time captions across two television screens that were strategically situated in the room. In the focus group in which the participants used Mexican Sign Language (MSL), the communication loop consisted of a hearing focus group moderator who voiced in English; his words were translated into American Sign Language (ASL) by a hearing interpreter, and the hearing interpreter's signs were interpreted into MSL by a deaf interpreter who knew both ASL and MSL. This process was reversed for the focus group moderator to understand what the participants were saying.

One obvious challenge in trying to start off thought from marginalized lives is the need for in-depth knowledge of how to facilitate that participation for these cultural groups, access to resources to cover the expense of providing for meaningful participation, and patience to handle the added complexity of insuring that those with the least power can meaningfully participate in the provision of data. The staff and the focus group moderator expressed frustration about the communication—it seemed slow and difficult to control. One could guess that the reaction of court personnel to this type of communication system might be similar. This, in itself, was a valuable insight that we could share with court system personnel to elucidate the

potential for cultural conflicts emanating from different expectations concerning language use.

But, the payoff is in the quality of the data that results when diverse perspectives are represented. For example, we obtained the following data from the MSL group: The respondents reported having difficulty finding the right court room in which to appear. One respondent described arriving at the court house with a citation for failing to wear a seat belt. However, he has a hard time with English, doesn't speak Spanish, and uses Mexican Sign Language. Although people wrote notes for him as to how to find the court room, he could not read or understand the notes. He eventually gave up trying to find the room. If this were simply a problem of not being able to find the room, then we might be able to dismiss this incident. However, the consequence of not being able to find the room was that a bench warrant was issued for his arrest. When another problem with the police arose [because he jumped the turn stile at the train station instead of paying the proper fare], he was put in jail for five days. He is not sure, but he thinks it was because of the warrant. (Mertens, 1996, p. 11)

This episode suggests that his real crime was not being able to communicate in oral, signed, or written English. And inclusion of hard of hearing people illuminated other problems related to the attitudes of court personnel. For example, in one case, the judge became impatient with a woman who told him that she could not understand him. She said he became angry and immediately started using one syllable words, "...talking to me like I was a child." She could not understand him because she could not hear his voice; not because she was stupid. Some problems related to obtaining equal justice under the law can be solved by having appropriate technologies or interpreters available; others involve more intransigent attitudes of people in power.

Developing Data Collection Instruments and Facilitating Social Action

Involving members of marginalized groups can be one strategy that will lead to better data collection instruments. Oliver (1992) identified questions from the Office of Population Census and Surveys (OPCS) in Great Britain that represented a blame the victim theory, such as:

- What complaint causes your difficulty in holding, gripping, or turning things?
- Are your difficulties in understanding people mainly due to a hearing problem?
- Have you attended a special school because of a long-term health problem or disability?

He suggested that questions written from a transformative perspective would ask about the aspects of the social and physical environment that were serving as barriers to full access or personal freedom, rather than locating the problem of disability within the individual.

Chelimsky described one study in which the evaluators at GAO used such a strategy that involved surveying disabled people before conducting a survey to determine the effectiveness of the Americans with Disabilities Act. She said:

We used their responses in constructing both the design and the survey instruments, recognizing – based on what we had learned from them – the need to ask probing questions of business owners and operators not just about observable barriers, but also about invisible ones, such as whether a blind person with a guide dog might be refused entry to a cafe or restaurant. (USGAO/PEMD, May 1993, cited in Chelimsky, 1998, p. 47)

This practice of basing questions on the lives of those with personal insights and experience about the problems is in keeping with the tenets of transformative theory.

In my court access project, we designed the data collection forms with an eye towards facilitating transformative change. As part of the training programs for judges and other court personnel, we invited deaf and hard of hearing people and their advocates to attend the training workshops with representatives of the court systems in their state. The final session of the workshop involved small groups from each state working together to complete an action plan so that state-level teams could assess their current status in terms of court accessibility and make plans for future actions. Each item on the action plan form was the topic of a plenary session prior to the planning session so that the participants would have the most up-to-date information available on the topics to use in their planning discussions. The idea of planning together as a team with court personnel and representatives of the deaf and hard of hearing communities was repeatedly emphasized.

In keeping with the transformative paradigm's underlying assumptions, the evaluation and project activities were interwoven to facilitate a link to action from the evaluation results and the need to improve court access. Three strategies were used to facilitate this link: (1) the development of the workshop and curricular materials on the basis of the focus group data that had been used to identify the needs and problems experienced by deaf and hard of hearing people in the courts; (2) the composition of the state teams that attended each workshop to include representatives of the court and the deaf and hard of hearing communities; and (3) inclusion of the action plan preparation as part of the workshop activities (this was also used as a baseline for the follow-up survey). Through the action plan process, participants were asked to reflect on actions that were not limited to appropriate use of interpreters and technology, but also to consider legislation and attitudes as they influence achievement of equal justice.

Despite positive responses from the workshop participants about the value of the joint planning process, as an evaluator, I am also aware of the problems in bringing about change in a social system as complex as the American judicial system. In many cases, the individuals who attended do not have the power to make the changes that may be necessary. For example, the issue of a change of legislative requirements falls within the purview of the legislators, not the judiciary. Yet, it is the judge who will be sued in federal court if he/she adheres to state legislation that is not in compliance with federal legislation. So I do not make the claim that we have solved all the problems in the provision of equal justice in the courts for people who are deaf or hard of hearing through the application of the transformative paradigm to this evaluation project. I do, however, make the claim that by consciously working from within this paradigm, the right parties' voices can be brought into the conversation and the hope for real change is improved by this process.

CONCLUSIONS

Accepting the challenge of inclusion suggests the need to think differently about the design, implementation and communication of the results of our evaluations. The challenge is to approach the evaluation in a way that builds in a greater likelihood of utilization. However, the challenge goes further, and asks: What is the evaluator's role in facilitating utilization in a way that holds greater promise that the social ills that formed the basis for the need for the program can actually be ameliorated? How can we insure or increase the probability that the bridge between evaluation findings and social transformation can be crossed? My position is

this: As evaluators, we want people to use our findings for purposes of positive change. Therefore, the production of evaluation results and their use to transform societal ills cannot be separated. We are not the decision makers; however, we do contribute to the process of decision-making by the nature of our relationships with those in power to decide, by the way we relate our findings, and to whom and how we disseminate those findings. It is true that people in power can choose to ignore our findings, especially the ones they don't like. So, what power do we have? What other ways can we act on the knowledge we have gained by using networks of stakeholders, in addition to communicating with the decision makers?

Much of the talk about utilization adds up at the bottom line to the fact that evaluation empowers some interests, and not others. The challenge of inclusion that is really critical is who is empowered, and in what ways and for what purposes? These questions are inescapable, no matter whether you embrace evaluation as a technical activity or evaluation as contributing to societal transformation.

How can we design an evaluation that is used to change conditions in a positive way so that things actually get better for the most vulnerable, the least powerful? One way we do this is by meaningfully involving in the evaluation process those whose life experiences are the most full of hurt and pain; those who are sometimes labeled as the marginalized, disenfranchised, and least powerful. Finding ways to really listen to the life experiences of mothers who are facing the overwhelming task of finding and working at a job that does not provide a high enough salary to cover living expenses, health care, and child care. Finding ways to really listen to the African American children, their parents, and teachers in schools that have historically failed to convey even the most basic skills and face being placed in state receivership. Finding ways to really understand the experiences of deaf and hard of hearing people who have been denied equal justice under the law because of inadequate and inappropriate accommodations to their communication needs.

Including the voices of those that traditionally have been ignored is not enough – how do we get beyond that step? We know that the inclusion of these voices will mean more credible information about programs and program impacts on their lives. Transformation can only occur if this information is used to inform policies that effectively address the inequities that create the need for social programs.

This hope for real change provides a challenge to evaluators both in terms of how we do our work and how we frame ourselves as evaluators. How do we address issues of ideology, objectivity, politics, social justice, advocacy, discrimination, oppression, and inequities based on race, class, gender, disability, economic status, or sexual orientation? The evaluator's role is not the same role as an advocate, if the term advocate is assumed to mean a person who would distort the facts to support a foregone conclusion. The evaluator's role is not the same as the decision maker's role, if that means the person who has the power and right to decide on the course of action based on evaluation findings and other sources of information. But evaluators do face the challenge of determining how we can interact with stakeholders, including the least powerful, the advocates, and the decision makers so that we can play a meaningful part in a process designed to bring about real change.

These are the challenges that I see facing evaluators as we enter the next century. We live in an increasingly diverse society with increasingly serious social problems for the least powerful members of our society. When one group in our society has problems, they reverberate on other groups. It is not only the least powerful who are affected. People need to see connections between violence and powerlessness, and that this affects everyone in the form of unsafe living conditions and increased human suffering. I hope, as the evaluators present here struggle

with the process of responding to federal and state mandates for evaluation and the development of performance measurements, that thought will be given to these issues. As we prepare to enter the twenty-first century, evaluators from across the country and around the world will be struggling with the issue of exploring evaluation's role in bringing about fundamental changes in evaluation theory and practice that can lead to improvements for all members of society.

NOTES

1. Related to this point, Gordon and Meroe (1991) wrote: "This is a precarious and dangerous situation because too many times 'objectivity' has served as a mask for the political agenda of the status quo, thus marginalizing and labeling the concerns of less empowered groups as 'special interests'." (p. 28, cited in Banks, 1995, p. 12)

REFERENCES

- Alcoff, L., & Potter, E. (Eds.). (1993). *Feminist epistemologies*. New York: Routledge.
- Banks, J. A. (1993). The canon debate, knowledge construction, and multicultural education. *Educational researcher*, 22(5), 4-14.
- Banks, J. A. (1995). The historical reconstruction of knowledge about race: Implications for transformative teaching. *Educational researcher*, 24(2), 15-25.
- Canadian Research Institute for the Advancement of Women. (1996). *Feminist research ethics: A process*. Ottawa, Ontario: Author.
- Chelimsky, E. (1998). The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation*, 19(1), 35-56.
- Connell, P.J., Kubisch, A.C., Schorr, L.B., & Weiss, C.H. (1995). *New approaches to evaluating communities initiative: Concepts, methods, and contexts*. Washington, D.C.: Aspen Institute.
- DeKoning, K., & Marion, M. (Eds.). (1996). *Participatory research in health*. London: Zed books.
- Fine, M. (Ed.). (1992). *Disruptive voices*. Ann Arbor: University of Michigan Press.
- Gill, C. (1998). Invisible ubiquity: The surprising relevance of disability issues in evaluation. Plenary presentation at the American Evaluation Association's annual meeting in Chicago, IL, Nov. 5, 1998.
- Gitlin, A. (Ed.). (1994). *Power and method*. NY: Routledge.
- Greene, J. (1997). Evaluation as advocacy. *Evaluation Practice*, 18(1), 25-35.
- Greene, J., Lincoln, Y. S., Mathison, S., Mertens, D. M., & Ryan, K. (1998). Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation*, 19(1), 101-122.
- Harding, S. (1993). Rethinking standpoint epistemology: What is 'strong objectivity'? In L. Alcoff & E. Potter (Eds.), *Feminist epistemologies* (pp. 49-82). New York: Routledge.
- Hill-Collins, P. (1990). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. New York: Routledge.
- House, E. R., & Howe, K. R. (1998). The issue of advocacy in evaluations. *American Journal of Evaluation*, 19(2), 233-236.
- Humphries, B., & Truman, C. (Eds.). (1994). *Re-thinking social research*. Aldershot, UK: Avebury.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs*. Thousand Oaks, CA: Sage.
- Kirkhart, K. E. (1995). Seeking multicultural validity: A postcard from the road. *Evaluation practice*, 16(1), 1-12.
- Lather, P. (1992). Critical frames in educational research: Feminist and post-structural perspectives. *Theory into Practice*, 31(2), 1-13.

- Madison, A. M. (1992). *Minority issues in program evaluation*. (New Directions for Program Evaluation, Vol. 53). San Francisco: Jossey-Bass.
- McLaren, P. L., & Lankshear, C. (Eds.). (1994). *Politics of liberation*. New York: Routledge.
- Mertens, D. M., Farley, J., Madison, A. M., & Singleton, P. (1994). Diverse voices in evaluation practice: Feminists, minorities, and persons with disabilities. *Evaluation Practice*, 15(2), 123–129.
- Mertens, D. M. (1995). Identify and respect differences among participants in evaluation studies. In Shadish, W. R., Newman, D., Scheirer, M.A., & Wye, C. (Eds.), (1995). *The American Evaluation Association's guiding principles* (pp. 91–98). San Francisco: Jossey-Bass.
- Mertens, D.M., & McLaughlin, J. (1995). *Research methods in special education*. Thousand Oaks, CA: Sage.
- Mertens, D. M. (1996). Evaluation report on the third and fourth focus groups: Late-deafened and Mexican-American deaf adults in Los Angeles. Report prepared for the American Judicature Society, Chicago, IL.
- Mertens, D. M. (1998). *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. Thousand Oaks, CA: Sage.
- Mertens, D. M. (In press). Deaf and hard of hearing people in court: Using an emancipatory perspective to determine their needs. In C. Truman, D. M. Mertens, & B. Humphries (Eds.), *Research in an Unequal World*. London: Taylor & Francis.
- Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal*, 32(1), 3–33.
- Oliver, M. (1992). Changing the social relations of research production? *Disability, Handicap, & Society*, 7(2), 101–114.
- Reason, P. (Ed.) (1994). *Participation in human inquiry*. London: Sage.
- Reinharz, S. (1992). *Feminist methods in social research*. New York: Oxford University Press.
- Schorr, L.B., & Kubisch, A.C. (1995, September). *New approaches to evaluation: Helping Sister Mary Paul, Geoff Canada, and Otis Johnson, while convincing Pat Moynihan, Newt Gingrich and the American public*. Presentation at the Annie E. Casey Foundation Annual Research Evaluation Conference: Using research and evaluation information to improve programs and policies.
- Schwandt, T. A. (1997). The landscape of values in evaluation: Charted terrain and unexplored territory. In D. Rog & D. Fournier (Eds.), *Progress and future directions in evaluation: Perspectives on theory, practice, and method* (pp. 25–39). San Francisco: Jossey-Bass.
- Shadish, W. R., Newman, D., Scheirer, M.A., & Wye, C. (Eds.). (1995). *The American Evaluation Association's guiding principles*. San Francisco: Jossey-Bass.
- Stanfield, J. H. II (1993). Methodological reflections: An introduction. In J. H. Stanfield & R. Dennis, (Eds.), *Race and ethnicity in research methods* (pp. 3–15). Thousand Oaks, CA: Sage Publications.
- Stanfield, J. H. II (1998). Social action research as vehicles of empowered social change in multiethnic and ethnically different communities. Plenary presentation at the American Evaluation Association annual meeting in Chicago, IL, Nov. 4, 1998.
- Stanfield, J.H., & Dennis, R.M. (Eds.). (1993). *Race and ethnicity in research methods*. Newbury Park, CA: Sage.
- Truman, C., Mertens, D., & Humphries, B. (Eds.). (in press). *Research in an unequal world*. London: Taylor & Frances.
- Villegas, A. M. (1991). *Culturally responsive pedagogy for the 1990's and beyond* (Monograph). Princeton, NJ: Educational Testing Service.
- Weiss, C. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–34.
- Whitmore, E. (1998). *Understanding and practicing participatory evaluation*. (New Directions in Evaluation, Vol. 80) San Francisco: Jossey-Bass.
- W.K. Kellogg Foundation. (1998). *W.K. Kellogg Foundation Evaluation Handbook*. Battle Creek, MI: Author.