

Masthead Logo

Western Washington University
Western CEDAR

Psychology Faculty and Staff Publications

Psychology

1992

The Role of the Researcher in Evaluating American Indian Alcohol and Other Drug Abuse Prevention Programs

Joseph E. Trimble, PhD

Western Washington University, joseph.trimble@wwu.edu

Fred Beauvais

Follow this and additional works at: https://cedar.wwu.edu/psychology_facpubs

Part of the [Multicultural Psychology Commons](#)

Recommended Citation

Trimble, PhD, Joseph E. and Beauvais, Fred, "The Role of the Researcher in Evaluating American Indian Alcohol and Other Drug Abuse Prevention Programs" (1992). *Psychology Faculty and Staff Publications*. 4.
https://cedar.wwu.edu/psychology_facpubs/4

This Article is brought to you for free and open access by the Psychology at Western CEDAR. It has been accepted for inclusion in Psychology Faculty and Staff Publications by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.

The Role of the Researcher in Evaluating American Indian Alcohol and Other Drug Abuse Prevention Programs

*Fred Beauvais
Joseph E. Trimble*

Preparation of this manuscript was supported in part by funds from the National Institute on Drug Abuse (Grant # P50 DA07074) and the National Institute on Alcohol Abuse and Alcoholism (Grant #AA 08302)

Abstract

This chapter defines the role and responsibilities of researchers who are asked to evaluate alcohol and other drug (AOD) programs in American Indian communities and settings. Building on the framework provided in the previous chapter, it identifies the various conceptual, methodological, and procedural problems that evaluators may encounter in settings that are culturally different from their own. Topics such as gaining access, measurement equivalence, report writing, and dissemination of results are given specific attention. The chapter also highlights those factors that can assist in "bridging the gap" between those

responsible for designing an evaluation protocol and those charged with designing and implementing prevention programs, and concludes that evaluation planning must be integrated into the planning of AOD programs in Indian communities.

Evaluation is an elastic word that stretches to cover judgments of many kinds. People talk about evaluation of a worker's job performance, evaluation of a movie script, evaluation of the sales potential of a new detergent. What all the uses of the word have in common is the notion of judging merit (Carol H. Weiss 1972, p. 1).

Introduction

As Candace Fleming (this volume) points out, there are many important cultural and socioeconomic dimensions that distinguish American Indians.*

Issues, Problems, and Pitfalls

Considering the wealth of information generated by ethnographers, epidemiologists, psychiatrists, and social workers, many Indian communities have been the target of a great deal of research, yet for many American Indians, the goals, methods, and procedures of science—and therefore of good program evaluation research—are unimportant, obscure, and unclear (Trimble 1977). Obviously, these efforts have not always worked out well. Many a well-intentioned social scientist has been and continues to be viewed as an outsider—"a predator who is using the Indian to further his career" (Maynard 1974, p. 402). A few researchers, largely because of their lack of attentiveness to a community's lifeways and thoughtways, have even been banned from continuing their work (cf. Manson 1989; Trimble 1988).

* The terms "American Indian," "Indian," and "Native American" are currently used interchangeably to refer to people native to the United States. Although the native people of Alaska differ from native people of the "lower 48" on a number of dimensions, they are generally included under these same terms. We will follow that convention.

Because of the growing concern surrounding the presence of outside non-Indian researchers, several Indian communities in North America have established rigid guidelines intended to regulate the research process and effort. Typically, researchers must present a proposal before a tribal, village, or community governing body, describing in detail the intent, nature, and benefits of the project. Delays occur more often than not. When and if the researcher is granted permission—a type of solicitor's license—conditions are attached. These may include (1) assignment of a tribal or village member to monitor the effort; (2) guidelines concerning respondent selection procedures; (3) the community's right to review and edit questionnaires, interview schedules, field notes, etc; (4) the community's right to review and edit research reports and to restrict or prevent the circulation and distribution of findings; and (5) ownership of the raw data and findings granted to the tribe or village. These conditions may be viewed as strident; however, it should be remembered that a few Indian communities forbid *any* outside-sponsored research from occurring within their boundaries. Hence, although the intricate and complicated elements of scientific research may not be well understood, Indian communities recognize all too well that the research process *can be* intrusive and the results invidious, divisive, and scandalous.

Whether an evaluator is invited by the community to design and conduct evaluation research or initiates the effort independently, several preliminary steps must be followed if the effort is to be effective and consistent with the community's lifeways and thoughtways. Adair and Deuschle (1970) strongly make the following recommendations:

1. Those members of the donor society concerned with planned change must have a comprehensive knowledge of the culture of those for whom the innovations are designed.
2. In addition, there must be constant awareness on the part of those planning change of their own culture (or subculture), its values, structures, predilections, and biases.
3. The political structure . . . must be understood and its leadership identified and worked through (pp. xiv–xv).

Finally, any researcher, regardless of training and orientation, has an obligation to honor the generosity, hospitality, and cooperation of the community of interest.

Heterogeneity of American Indians

The chapter by Fleming (this volume) provides an understanding of the context in which research and evaluation among American Indians takes place. Perhaps the most important lesson of that chapter is the heterogeneity of Indian people. Acculturative status, degree of identity, residential status, physiognomic characteristics, language preferences, and lifestyle preferences vary considerably between and among Indian and Alaska Native people. These factors must be taken into consideration when one is planning any type of intervention program and designing an evaluation model to assess its impact. Without this understanding, the evaluator will be operating only on stereotypes and will fail to connect with the local community.

Assumptions and Chapter Organization

A few basic assumptions are in order. First, American Indians will continue to seek support for social programs that are aimed at alleviating many of the problems in their communities. Second, there will be a continuing expectation on the part of funding and sponsoring agencies that some sort of evaluative activity will be required if those programs are to receive continued funding. Finally, there has not yet been any detailed approach to conducting evaluation research that springs exclusively from Indian cultural values. Indeed, with this statement we may have encountered our first point of contention. Does the concept of evaluation even make sense from an American Indian point of view? This is, perhaps, too large a question to be resolved here.

Our intention here is to sidestep this larger question and pursue ways of applying the evaluative process, as it has evolved as a stepchild of conventional, academic-based scientific methods, to alcohol and other drug (AOD) abuse prevention programs in Indian communities, whether they exist in rural, village, or urban settings. This effort will necessarily involve a

number of compromises and may lead to unresolved issues. A further goal of this chapter is to bridge the gap between those responsible for the design and implementation of AOD misuse/abuse prevention programs and those charged with evaluating them. We will try to attend to the needs and concerns of both sides and to produce workable procedures for the evaluation process to be effective.

As an organizational structure for the discussion of program evaluation, we follow the development of a prevention program and present issues that are of particular relevance to Indian communities. At points we necessarily touch on topics that are not directly related to evaluation but that are important to an understanding of the entire process. In fact, it is difficult to separate program development from evaluation, given that, if done properly, they are intricately connected, even from the beginning. It is surprising, the number of requests we get to conduct an evaluation once the program is underway or even in its closing stages. In these cases, there is usually a sense of urgency, and one gets the feeling that the program staff have suddenly realized they have forgotten the evaluation component and need it to fulfill their contract. At this point, only a patchwork design is possible and, from the perspective of further knowledge, very little can be done.

Accessing Communities

Gaining access to work in Indian communities presents some unique challenges. The issues to be discussed are important, whether the evaluator is working in cooperation with a local agency or is totally new to the community.

Barriers to Access

In the past couple of decades, numerous Federal initiatives and programs have been launched to address the human services needs of Indian people. Unfortunately, most of these efforts have followed a predictable, unproductive course, which has led to a great deal of skepticism among Indian leaders.

What happens is that periodically, when awareness is raised about a particular problem on reservations, there is a move to create funding for new programs. The recent flurry of activity surrounding such issues as acquired immunodeficiency syndrome (AIDS) prevention, fetal alcohol syndrome, and child abuse and neglect are typical examples. Initially, there is a lot of enthusiasm for the effort, and programs are developed in many locations, often with the assistance of outside resources for planning and evaluation, such as those at colleges and universities. These programs normally operate with various levels of effectiveness until interest wanes at the funding source, monies run out, and the program dies. If patterns follow their usual course, we will see fewer and fewer initiatives in the coming years even though the problems themselves have not abated. Local human service providers, particularly at the tribal level, have witnessed endless rounds of this type of programming and quite understandably are leery of the efficacy or longevity of new programs. Those working with Indian communities must be aware of the source and nature of this skepticism.

There is also a variation on this theme that often occurs. On some reservations, there may be ongoing efforts to deal with a particular social, psychological, economic, or health-related problem. The efforts may not have high visibility, but there are local people who have invested a lot of time in them and have a personal commitment to the work they are doing. When a "new" initiative comes along to solve the problem, particularly an initiative that brings with it a good deal of outside involvement, there may understandably be a certain level of resentment. The local people who have been working on the problem may feel that their efforts have been ignored, and the implementation of a new program implies that what they have been doing has little or no value. Those who have endured several cycles of this will likely demonstrate little enthusiasm for new programs, and their lack of support could result in program failure. To avoid this, the program people and the evaluator need to have a thorough understanding of the history in the community they are working with and should work in tandem with existing service providers and resources (cf. Trimble and Hayes 1984).

A further barrier to access involves the wariness engendered by the past 400-year relationship between American Indians and the European and U.S. Governments. Fleming (this volume) has chronicled this often stormy and checkered relationship, and there is little doubt that a fully trusting relationship has yet to develop. It is important to recognize that much of the conflict Fleming describes is contemporary and very much on the minds of Indian people.

Gaining Access

Implementation of any program within a community requires some type of official sanction or alliance with a sponsoring agency. The evaluator needs to make certain that the appropriate clearances have been obtained and that the evaluation work is acceptable to the community. It often happens that the program staff feel their program is well accepted; however, should the evaluator have to venture outside the normal program boundaries for data collection, he or she may meet resistance if the right level of approval has not been obtained.

For example, in 1986, a researcher interested in collecting data from residents in a small Cree village in Canada's Manitoba Province knew in advance that the tribal leaders were reluctant to endorse his efforts. The tribal policies concerning the conduct of research by outsiders were fairly well known. The researcher stood to acquire a large Federal grant, but to obtain the funds he needed some semblance of support showing he had access to the village's adolescent population. Because the tribal leaders were not willing to endorse the project, the researcher solicited support from a non-Native local school principal. The principal complied, assuming that the tribal leaders supported the effort. The researcher received his funds and proceeded to conduct his work.

Some time later, the principal, while having lunch with a few tribal leaders, happened to mention his excitement about the research. In a matter of a few hours, the researcher and his assistants, together with their questionnaires, were summarily escorted to the reserve's boundaries. Shortly thereafter, officials at the researcher's university received a strongly worded letter

essentially declaring the reserve "off limits" for anyone affiliated with the institution with an interest in conducting future research in the community.

Those not familiar with the history and structure of Indian communities may become confused by the issue of gaining access and legitimacy. For one who has worked in only non-Indian communities, the power structure and political scene can be baffling. Because most Indian tribes are sovereign political entities, it is usually necessary to gain approval for new programs and evaluation work from tribal or village councils. In larger tribes, this authority may be delegated to a council committee, often the health committee. We have seen any number of programs derailed because the tribe feels it has not given its sanction to the work. Occasionally an evaluator might believe the proper approval has been given (as in our example through the school system), only to find that people at the tribal level believe they also should have been approached. The remedy to this is a thorough knowledge of the local political and power structure.

Beyond the official levels of approval, one will often find informal gatekeepers who must be apprised of what is going to take place, and their consent must also be obtained, even if only informally. This is particularly important for programs that may be dealing with culturally sensitive material. In these instances, it is often necessary to spend a great deal of time informing local residents, particularly tribal elders, of the reasons for implementing the program, the content of the program, and the procedures that will be followed in the evaluation component. It is not uncommon for an evaluation effort to be stopped in midstream when local people protest the data collection procedures they find objectionable.

Both formal and informal approval may take an extended period of time, and this delay must be anticipated. It is characteristic of most Indian communities to make decisions by consensus—a process that cannot be rushed. If there is a sense that things are being rushed, the entire process may be delayed even further. One must also be aware that there are local political issues that must be dealt with before a final decision can be made. It is often tempting to try to become involved in the political arena in the

interest of moving things along; however, this is usually counterproductive, if not presumptuous. Indeed, as Fleming (this volume) points out, the political issues in Indian communities are subtle, complex, and usually inscrutable to those who are not from the community.

The ability to conduct program evaluation in a cross-cultural setting largely hinges on the nature of the agreements that are made in the beginning regarding the conduct of the evaluation. If the community people are made to feel they are an integral part of the process and if issues they feel are sensitive are dealt with appropriately, the evaluation will proceed smoothly. Moreover, community leaders must have a sense of ownership in the effort, especially because many will be providing the data gathered by the research team.

Cross-Cultural Methodological Concerns

Generating Researchable Ideas

Prevention programs usually begin with an individual or group that develops an idea it believes is unusual or, based on its experience, more workable in a particular community. One point we emphasize throughout our discussion, however, is that any new program ideas or procedures must be thoroughly checked for cultural congruence, and this must begin even at the initial stages of idea generation. A bad idea, or one that is at odds with local values and beliefs, has no chance of success. Furthermore, the task of evaluation is nearly impossible if there are major conflicts in how a problem is conceptualized. In formulating the idea and the evaluation plan, attention must be given to the knotty problems that emerge when working with different cultures.

Equivalence of Measures

Berry (1980) has made some distinctions in thinking about cross-cultural issues that can help evaluators gain congruence between

and among assessment approaches in the early stages of planning. He refers to three types of equivalence between cultures: conceptual, functional, and metric. We discuss the first two of these below and address metric equivalence further on.

To understand conceptual equivalence, one must recognize that every culture has developed ways of looking at the world that make sense to its members. This worldview, much of which is reflected in the language of the culture, has been shaped by environmental, historical, biological, and other factors that have marked that people's evolution as a unique group. While there may be commonalities in worldview, depending to some extent on the proximity of groups, there are also usually areas of significant differences. For instance, many American Indians differ from White people in their view of what mental illness is all about (Trimble and Hayes 1984). For some tribes, mental illness is the result of having in some way transgressed the rules of right living, and until this can be rectified through ceremony, the illness will continue; thus, it is a spiritual issue whose resolution is in the hands of a medicine person, or shaman. This contrasts with the White view that the person has been subjected to a pathological process that can be relieved through medication combined with the individual's efforts to change his or her behavior.

It can easily be seen how questions about mental health in one worldview would not make sense in the other. For instance, if one were to ask a very native-oriented (i.e., traditional) Indian what causes alcohol and other drug abuse and were to give that person the usual psychosocial options to choose from, most likely he or she would be at a loss to respond. For some tribal worldviews, the most appropriate response would be that the individual or a family member had broken some type of taboo and that the resulting lack of harmony with the spiritual world must be rectified through traditional medicine. Clearly, the typical kinds of questions that are asked about attitudes or beliefs among non-Indians probably would not elicit this same explanation.

The purpose or significance of an apparently similar behavior may differ across cultures; in other words, a behavioral act can be functionally equivalent yet have a different meaning. For

instance, for a White alcoholic there is usually some level of shame or guilt involved in abusive drinking when family members are present. For many American Indians, however, drinking with family members is a social event, and a refusal to drink is taken as a rejection of the other family members. Thus, if the evaluator is assessing the role played by the family in alcohol use, the questions used may be tapping different social meanings. If one were to ask an Indian if that person's family encourages him or her to drink, an affirmative response might indicate close family relationships, whereas a similar response from a White person could signal serious family dysfunction.

Lack of functional equivalence can occur at even more subtle levels. Indian children are often described as being very quiet and reserved, especially in a classroom headed by a non-Indian teacher. Quite often they are labeled as being withdrawn and unresponsive, the implication being that their behavior is a form of emotional dysfunction. However, this type of behavior in Indian children has a number of other explanations including showing respect for elders (i.e., the teacher), feeling shy in an unfamiliar situation, or being unwilling to speak up lest it be seen as an attempt to show superiority over other children. An assessment of classroom behavior using the usual indicators and interpretations could lead to erroneous explanations.

In our experience, it is still very common for evaluators to move into a cultural context where they have not had even cursory experience and to expect that all their concepts and measures will have the same meaning. In most instances, evaluators do not have the luxury of being able to establish conceptual or functional equivalence rigorously across cultures. This process is time-consuming and constitutes a long-term research program in and of itself. But this does not excuse the evaluator from making an honest effort to identify the most obvious and perhaps most important areas in which non-equivalence may occur.

The key to this effort, while time-consuming, is also simple. It requires that extended discussions take place between evaluators and knowledgeable local people who are part of, or very familiar with, the local culture. The philosophy, goals, and

methods of the evaluation need to be thoroughly examined from both cultural perspectives to see if there are conflicts and to adjust the evaluation process to address the conflicts, if any exist.

Specifying a Program Theme

Once an idea has been developed, the entire scope and all the activities of the program must be conceptualized. At this point, the idea is broadened and may be broken down into separate or related components. Each component must once again be examined for its congruence with local culture and values. Discussions should involve local planners, decisionmakers, and program people. In fact, some of these people could acquire firsthand knowledge of the evaluation research as the planning and conceptualization unfold. They may even identify certain areas that might be problematic. The following example reinforces this point.

A recent attempt to create an AIDS prevention program in a reservation school system was nearly derailed due to a lack of understanding of the cultural values surrounding birth control. One goal of the program was to familiarize students with the use of condoms, and an extensive curriculum module was developed. But there were very strong local beliefs regarding the use of condoms, about which the evaluation planning team was unaware. For the local people, it was legitimate to teach condom use to prevent sexually transmitted diseases but it was not legitimate to teach methods of birth control. This appeared to be a contradiction to the program developers and evaluators, but it was totally consistent with local beliefs. The community people were very concerned about disease prevention but were not willing to compromise their beliefs about sexuality and fertility. In the end, it was necessary to come to some compromise about what could be taught in the schools as well as what attitudes could be measured in the evaluation component. Much of the sexual behavior material was deleted.

As each program objective is developed, there must be concurrent thinking about how each of its components will be evaluated. This not only makes evaluation an integral part of the

entire planning process but also helps to sharpen the thinking about the goals themselves. This interchange may also reveal potential points of conflict. The central task in evaluation is to determine what evidence would be sufficient to demonstrate that a program is either effective or ineffective. The usual approach to this task is to examine the program goals, seek outcomes that can be quantified, and, finally, identify measures that will yield numeric comparisons.

At every step in the planning process, however, a cross-cultural evaluation effort may encounter difficulty (cf. Lonner and Berry 1986). At the most basic level, there may be differences concerning the overall goal of a social intervention. For example, the stated goals of an AOD abuse prevention program on an Indian reservation may be to reduce the levels of AOD use among junior high school students and to delay the age of onset of AOD experimentation. This seems rather straightforward and would call for pre- and postmeasures (at the appropriate intervals) to assess rates of AOD use in the target population and determine the age of first use. Some people in the community, however, may see this as an irrelevant effort. They may be expecting the program to foster a return to more fundamental, tribal-specific values, which they believe would automatically bring a reduction in AOD use; thus, to them, measurement is unnecessary. In one respect, this disparity can be described as people simply operating at different levels of specification—one looking at broad program goals and the other looking at a more specific behavioral level. However, this disparity could also be signaling a basic disagreement about the entire purpose of the evaluation effort. A more native-oriented person may be implying that this is a moral or spiritual issue that is not amenable to quantification. Unless these differences in expectations are made explicit, conflict will continue throughout the evaluation process.

Resolution of such discrepancies in expectation is not easy and usually involves compromise. On the one hand, a native-oriented person may have to accept the need for quantitative measures of behavioral change. This type of information is valuable from one perspective and may be especially persuasive to those agencies providing the funding. In some instances a

funding source may even require the use of quantitative approaches as a condition for support. However, it is also reasonable to expect that there be some judgment about whether fundamental values are being considered. For instance, it would be appropriate to have tribal elders or healers inspect the content and process of the program and give their judgment about its value in addressing cultural values and beliefs. The important point is that these opinions and judgments need to be given equal weight in the evaluation process and in the final report. This will require a change in mindset on the part of both the evaluator, who is not a member of the culture, and the funding agencies, who may be reviewing the program for refunding. There continues to be a bias against this type of "soft data" (often referred to as "subjective") among many professionals and funding sources, yet at times this might be the evaluator's most important source of information. In the final analysis, the objective measures in the evaluation of a prevention program such as that described above may show significant changes in AOD use behavior, but the tribal elders might conclude that the program activities are so contrary to cultural beliefs that the program should be scrapped. Both types of information are legitimate evaluation "data," and if either is missing, accurate judgments about program effectiveness are not possible.

Another means of ensuring equivalence is the use of local people as part of the evaluation team. It is important that these people be deeply involved with the planning and that their views be given full consideration. Too often, local people are hired as program staff, but their ideas are not sought and they are not included in planning sessions. Therefore, evaluators must be aware that many Indians interact and communicate with one another in unique ways. In meetings where ideas are being shared and plans are being made, it is common for Indian people to withhold their comments until everyone else has spoken. It often happens that meetings are ended before the Indians in attendance have had an opportunity to present their views, and an important source of information is lost. Thus, it is appropriate to make a concerted effort to solicit input from Indian staff

members and to allow several minutes of silence to ensure that everyone has had the chance to speak.

A Word of Caution

Up to this point, we have given considerable attention to describing the "front end" of program development and evaluation because we believe this is where most problems will arise. Ironically, many evaluators/researchers spend the least amount of time working with these issues, believing that evaluation is primarily a technical process. We are disconcerted by the number of calls we receive from researchers who want us to "find an Indian population where I can implement this new program I just got funding for." One wonders how such a proposal ever survived the scientific grant peer review process! There is little or no understanding of the multiple points of possible incongruence or conflict that can arise at the conceptual level of program development and of the inappropriateness of bypassing the many levels of protocol. This insensitive approach also helps perpetuate the perception among Indian people that they are being "used" for research purposes, with little concern for their interests and welfare. The value flows only in one direction, seemingly to further only the goals and interest of the researcher.

Measurement Issues

We now turn our attention to measurement issues and begin with the concept of metric equivalence as described by Berry (1980). The discussion that follows, however, focuses on conceptual problems, not on techniques of measurement; excellent descriptions of the latter are available elsewhere (Oetting et al. in press; Sudman and Bradburn 1974).

The items or scales that measure constructs often operate differently across cultures. A few years back, one of us was developing a multi-item scale to measure social deviance. The scale worked well for White youth and had a very high reliability. When the same scale was used with Indian youth, however, the reliability was much lower. Some investigation revealed that one of the items on the scale was consistently interpreted differently

by Indian youth. In fact, it had the opposite meaning for them than it had for White youth. The item, which asked students to rate how often they got into fights, read "I fight" and included the possible responses of "a lot," "some," "not much," or "not at all." White youth interpreted this item in the way we intended—i.e., "I get into fights." Many Indian youth, on the other hand, interpreted it to mean, "I will fight for what I believe is right." Not only did this clearly affect the scale's reliability among Indian youth but, if one were to compute a scale score, it would appear that Indian youth, on average, had higher levels of deviance than their White counterparts.

Another problem with metric equivalence occurs when the relationship between variables is not the same across cultures. In a current study, we have preliminary evidence that anger may operate differently in Indian and White students. For White students, higher levels of anger are related to higher levels of AOD use, whereas for Indian youth, higher levels of anger seem to lead to lower AOD use. It may well be that Indian young people experience a sense of anger at the many injustices they see in their world, and this acts as a positive motivator in their personal behavior. But without knowing that this difference exists between Indian and non-Indian youth, it would be easy to make the wrong interpretation about what anger means, and this might lead to interventions that are ineffective or harmful. This difference between Indian and non-Indian youth might also lead to erroneous conclusions when certain interventions are evaluated. Certain types of interventions could provoke this positive sense of anger among Indian youth. However, an inappropriate evaluation measure could lead to the notion that the program is harmful—i.e., that it leads to aggression and to a greater propensity to use alcohol and other drugs.

Although it is often ignored, establishing metric equivalence should be a standard task for evaluators. As previously mentioned, it is not enough to identify measures that have been used in other studies to measure a concept under consideration. It must be demonstrated that the selected instrumentation is both valid and reliable for the population on which it will be used. In addition to the usual reliability and validity studies, it is useful

to analyze the factor structure of the measures and constructs being used. Besides establishing metric equivalence, factor structure analysis can help in examining problems that may also exist with functional and conceptual equivalence. In fact, the above-mentioned problems in measuring deviance and anger were discovered through a structural analysis procedure.

The question is often raised as to whether it is best to use "off-the-shelf" measures or to construct new measures when doing evaluation in a cross-cultural milieu. There is no one answer to this question, given that problems can be encountered with each approach. But unless there is evidence that an existing measure has already worked in the population being evaluated, it is usually necessary to establish reliability and validity with that group. This is not to say, however, that all measures are inherently culturally biased and cannot be used, either in part or in whole, with other populations. For example, a scale to assess incidence and prevalence rates of AOD use, assuming no language barriers, should be accurate regardless of which population is being measured.

The construction of new items and scales is not a task that should be taken lightly. Many evaluators underestimate the difficulty of scale construction, and this difficulty is multiplied when the new scales are applied across cultures. One of the most common errors is not to test the scale before using it for evaluation. Pilot testing is an absolute requirement and should involve a debriefing procedure in which potential subjects can talk about their interpretation of the items. The many ways of interpreting a seemingly straightforward question are quite surprising. In assessing the level of perceived social acceptance among young Indian students, we have routinely used the item, "Other kids like to play with me." In the context of elementary school students, this item seems reasonable. However, at one point we tried the item with older students and received some unprintable responses. In retrospect, we felt particularly foolish for not having recognized the sexual connotations. In fact, as we looked back over our data from previous studies, we concluded that it was not even a particularly good item with younger students.

Many of the more mature ones may have seen the double meaning and responded capriciously.

One aspect of measurement that is often neglected in evaluations is the use of qualitative data. Too often, evaluators focus exclusively on quantitative methods and miss the richness of information that can accrue from other approaches. Qualitative methods, including ethnography, are particularly useful in dealing with the cultural diversity of Indian tribes, as discussed by Fleming (this volume). It is through these approaches that the subtle differences in worldviews can be discovered, and the evaluator can identify new dimensions that are not obvious or amenable to quantification. Fortunately, there is an increasing acceptance of qualitative methods from both the research community and the funding agencies.

The one caution that is in order is that sound qualitative methods be used. Evaluators who are not familiar with this area are often unaware that there are well-accepted, rigorous procedures for qualitative research that can yield reliable information. Too often, qualitative research is viewed simplistically and amounts to little more than having extended conversations with people. The best strategy for an evaluator who is inexperienced in these methods is to identify and use trained ethnographers who are familiar with the local culture.

There is one final set of points to be made in this section. Over the years, researchers working with different cultures have resorted to the use of measurement tools that are based on norms and the testing orientation of those with a Western perspective. All too often, these researchers encounter problems in administration, scoring, and, assuredly, interpretation. Critics abound, though, and a number of cross-cultural researchers have commented on the cultural inappropriateness of measurement approaches (Irvine and Berry 1983). Trimble and colleagues (1983), wondering "why some investigators, almost blindly and with utmost diligence, continue using measurement traditions" (p. 268), go on to identify seven common pitfalls in cross-cultural testing:

1. Psychological constructs are viewed as synonymous with locally derivable criteria, which may or may not be consistent with the implied intent of the construct.
2. The establishment of several types of equivalence is not considered essential.
3. It is assumed that once tests are purged of verbal material, leaving only nonverbal stimuli, they are more "culture-fair."
4. Norms gathered in one culture are used to evaluate the performance of individuals in other cultures.
5. People from around the world may have variable and different modes of responding to test items.
6. Such testing generally tends to infer deficits on the basis of test score differences.
7. Nearly all psychological tests are culturally isomorphic to the West, which can be characterized as sophisticated and "test wise."

Collecting Data and Information

Care must be taken to ensure that the actual data collection is done with consideration and respect for all elements of the community. With the emergence of numerous social service programs, it often happens that multiple evaluations may be going on at the same time in any one community. At a certain point, systems become saturated with surveys and other assessments, and resistance develops. This is not an issue solely with Indian communities, although we are aware that many people on reservations feel they are under extreme scrutiny, which gives rise to certain levels of resentment and suspicion. Schools in particular feel the pressure to respond to a variety of social problems, which they feel reduces their capacity to educate. One teacher on a southwestern reservation recently commented, "Anymore we have to teach kids everything from tying their shoes to how to put on a condom, and we barely have time to teach them to read." It will be increasingly important to keep the amount of evaluation assessment to a minimum, and such assessment must be done in a way that creates the least intrusion on the system.

One way of reducing the amount of assessment that is done for evaluation is to ensure that the distinction between basic research and evaluation is maintained. Evaluators are essentially researchers, and it is difficult for them not to ask for the maximum amount of information when a given problem is presented. However, what is "interesting" to an evaluator/researcher often has only marginal relevance to program effectiveness. This is not always an easy judgment to make, given that evaluation must often respond to both questions, "Does the program work?" and "Why does it work?"—the latter being necessary to see if the program will generalize. However, if the survey protocol becomes too long, it will jeopardize the entire data collection process. It is also important to recognize that many reservation schools have serious financial and staffing deficits, which result in poor reading skills among students. A survey package that works well for urban youth may be too complex for Indian youth.

Another assessment difficulty that may arise is that some members of the community may feel they were not consulted about the content of the assessment tools. This can never be totally averted, and there is always the potential for objections. This is particularly common in Indian communities, where decisions are made by consensus and procedures are not sanctioned until everyone has had an opportunity to voice an opinion. The evaluator must ensure that a reasonable amount of time has been given for widespread review of procedures, especially the items to which young people will be asked to respond.

The evaluator needs to visit the actual site where the program to be evaluated will be conducted, and to become familiar with the people the program will serve and with their daily operations and problems. Without this understanding, the data collection plan may be totally unrealistic. Many Indian communities are geographically isolated and lack resources that are commonly assumed to exist.

A site visit not only will give the evaluator a feel for local conditions but also will help prevent a common problem: system overkill. It is easy to build an ideal evaluation plan that, in practice, is not feasible. For example, a plan that calls for interviewing the parents of all the children in a program might work

well in a city but not in a location that encompasses hundreds of square miles. In addition, many grassroots programs employ people whose personal and financial resources are severely overtaxed; thus, to ask them to commit more personal time to engage in evaluation tasks is totally unrealistic.

Data Analysis, Report Writing, and Dissemination of Findings

Data Analysis

Data analysis is generally a technical task that is often left to the evaluation team, which then presents the results to the program staff and other local groups. In a cross-cultural setting, it is advisable to include knowledgeable local people even in the early phases of analysis; it is an opportunity to gain greater insight into the data. In a recent evaluation of a mental health training program on a reservation, one outcome variable was change in the communication patterns between the staff in various mental health agencies in the communities. Based on the data, the evaluator determined that communication patterns had changed only slightly for the better in some respects but that there were certain channels that showed no change; thus, the evaluator concluded that this aspect of the program was ineffective and should be eliminated as a program goal. However, a local agency representative, who was part of the evaluation team, indicated that much of the communication in this community was governed by traditional family and clan relationships and that the approach taken in the training program had little chance of changing this pattern. Part of the issue had to do with traditional and therefore acceptable paths of communication between older and younger generations, but there was also an element of longstanding feuds between certain factions within the community. This interpretation had clear implications for any conclusions drawn about program effectiveness; communication patterns could be improved, but a different approach would be needed.

Report Writing

Fleming (this volume) discusses the negative impact that many research and evaluation studies have on the reputation of ethnic/racial communities. In large part, this is due to the way in which program proposals and evaluation reports are written. Social programs, by nature, are designed to address social ills; therefore, written descriptions focus heavily on the negative aspects of communities. When ethnic/racial communities are involved, this type of reporting—over time—tends only to reinforce negative stereotypes.

In the worst case, study results can be blatantly used to denigrate a community. Some years back, a local bordertown newspaper obtained the results of a survey of alcohol and other drug use that was given on a reservation, and it sensationalized the results. Although there were no overt racial statements, the intent was clear and the Indian community experienced a great deal of shame. Social problems do exist in ethnic/racial communities; however, it is necessary to place them in context, and any evaluation report should reflect that context. For example, AOD use problems in Indian communities largely reflect socioeconomic conditions and are not related to any inherent cultural characteristics. Indian people are becoming increasingly impatient with the litany of social ills that are ascribed to them, and an evaluation report that presents a balanced picture will get a much better reception and is more likely to be used. In a word, the report should be written and presented respectfully.

At several points in this chapter we have suggested that an evaluation may have two purposes. At the local level, people need to know whether the program has value for them, that is, is it culturally congruent, well received by the community, and consistent with local values and norms? There is also a need for technical data that support the report's conclusions and may be used to answer more specific questions. These two purposes suggest the need for two types of reports. It often happens that technical reports are never used at the local level because they are too complex and do not respond directly to the need to make decisions. In the absence of a more comprehensible document, the community may be left with the feeling that the evaluation

was a wasted effort, and it may develop a negative attitude toward evaluation in general. Thus, a report written in nontechnical language specifically to address the local need is appropriate (Trimble 1977).

Dissemination

The issuing and dissemination of the evaluation report will often have to be handled carefully, particularly if the report contains sensitive information. In one sense, this is a question of who owns the data. Strictly speaking, the report is a part of the intervention program, which, in turn, is an extension of the community agency sponsoring it. Given this, the community or its representatives would have the final decision about the dissemination of results. There must also be a recognition, however, that this type of information can be useful in other communities and therefore needs to be published in some form.

Several approaches can be used to reduce controversy over publication. First, as we have stated before, these negotiations should occur very early in the evaluation process, and some general agreement should be reached. In some cases, even if prior agreement has been reached, there may be some unanticipated results that community people find sensitive and would not like to see publicized. Usually, a compromise can be reached through negotiation, whereby some information may be deleted or left in a report for internal purposes only. It is also useful to allow local people to preview the report to determine whether there are any conclusions that could be more accurately interpreted in light of local culture, values, or beliefs.

Whenever there is concern over report content, the manner in which the report is released can be extremely important. In 1980, a very sensitive report on alcoholism in a Native Alaska village was released to the general media (cf. Manson 1989). In addition to a number of other serious errors in protocol, the information from the study was presented at a press conference thousands of miles away from the village where the study was conducted. This precluded any participation by local people and allowed the whole situation to be presented out of context. Once again, an Indian community experienced a great deal of shame

because the information released implied that nearly all of the Indian adults in the community were alcoholic. Although the actual situation was quite different, there was no way to moderate what was presented.

It is good practice to have local people involved in any release of information, either in person or through a cover letter signed by an agency representative. This once again demonstrates the need for community people to be intimately involved with any evaluation effort. It not only ensures that the most accurate information is presented, but also precludes the perception that the community is once again the subject of outside interventions and is not capable of resolving local issues.

Social Policy Implications

As we pointed out earlier, there is tremendous diversity among American Indians and a strong movement to preserve that diversity. While this imparts a sense of pride and integrity among Indian people, it also makes the process of evaluation very complex. To be effective, a program should be tailored to the tribe where it is offered; therefore, a program that may be effective for one tribe may be culturally irrelevant for others. If the focus is on the content of a particular program, there may be little generalizability from one tribe to another. However, we have discussed a number of process variables that have more widespread application among Indian tribes. Perhaps the most important issue is that of spending adequate time in the early stages of the evaluation ensuring that there is congruence between the values and expectations of the program staff, the program audience, and the evaluator. There must be extensive time spent ensuring that the concepts and measures used in the evaluation are culturally appropriate.

Another common theme in the evaluation of Indian programs is the need to agree on the use and dissemination of evaluation results. There must be a continuing effort to place the focus of evaluation on the value of the program and not on the value of the program staff. Too often, evaluation is seen as a way of cutting programs rather than improving them. It is important to realize

that employment on many reservations is a critical issue. Unemployment rates are exceptionally high and jobs are greatly prized; on some reservations where housing is scarce, getting a job also means having access to a house. If program staff feel that an evaluation of their work could possibly jeopardize their job and housing, there may well be resistance. Thus, it is incumbent on both program administrators and the evaluator to create a climate of trust and a sense that the evaluation is solely for the purpose of improving the services offered.

A more general issue involves the creation of an ethic among tribal services that evaluation is an integral part of the work that is being done and not just an adjunct that is tacked on to program services. This can be accomplished by striving to involve all program staff in the process and by having them contribute to the data collection. Often evaluation is perceived as some elaborate technical enterprise that can be conducted and understood only by experts or outsiders. In truth, at one level, evaluation is nothing more than pulling information together to form some sort of judgment about the efficacy of a program. Using this broad framework, anyone on the program staff can contribute to the evaluation effort. There is often an extensive amount of qualitative and anecdotal data that can bolster an evaluation report. For example, in Indian communities the perception and judgment of elders is often critical to a program's long-term success. If the older people see it as valuable and congruent with traditional values, the program stands a much better chance of being widely accepted in the community. On the other hand, a negative perception by elders can be sufficient to terminate a program; however, without some way of detecting the attitudes of elders, it may never be known what led to the program's demise. This is the type of information that may not emerge from the formal evaluation process yet may well be accessible to the local staff.

Similarly, many programs have unintended consequences, either positive or negative, that may not be anticipated in the formal data-gathering system. Unless all staff members are alert to these possibilities and are encouraged to share them with the evaluator, such data may never come to light.

An example of an unanticipated outcome occurred some years ago on a southwestern reservation during a program intended to educate people about diabetes. Over the course of a year, community health representatives went out to remote homes and disseminated information concerning the diagnosis and treatment of diabetes. After a time, the clinic staff noted that a number of people were coming in for health care unrelated to diabetes. What had happened is that the representatives had created relationships with many of these families and had made it much more acceptable for people to come into the clinic for health problems that previously had been ignored. The important point here is that the staff were alert to this outcome and were willing to share it with the evaluators, who in turn made this information an integral part of the evaluation report.

Decisions to retain or terminate programs are often based on factors other than the evaluation data. These decisions may be out of the control of the program staff and certainly are not in the purview of the evaluator. In one instance, a training program for health workers on a reservation was retained even though the data indicated that it was not particularly effective. The reason for retaining the program was that it was a central link in a more extensive career ladder plan being implemented by the tribe. There may be a host of other administrative or political reasons for retaining or terminating a particular program. However, even in these instances, the evaluation results are still useful. We believe that administrative decisions should be made in light of all possible information, and program evaluation is certainly a critical source. Even if the final decision differs from what is suggested by the evaluation, that decision has been made on an informed basis.

This issue brings up an important administrative decision regarding evaluation: Will the results be used to help alter programs while they are operating (formative evaluation) or will they be used solely at the end of the program to determine the program's effectiveness (summative evaluation)? Depending on the purpose, different strategies will be used, and people's expectations concerning the nature of the outcomes must be congruent with the purpose. In the above health training program example,

a formative evaluation would have been helpful in providing feedback that could have changed the program midstream to make it more effective. In the evaluation of another drug prevention program in which a formative evaluation was used, it was discovered through interviews that the students felt that the information they were being given was too simplistic and that they were being talked down to. Outside the classroom, the students were making jokes about the program, and it was clear that the information was ineffective. Once this was discovered by the evaluation team, the staff was able to alter the curriculum and make it more age appropriate.

At other times, it might be more appropriate to test the effectiveness of a program without trying to alter it midstream—that is, to design the evaluation so that it provides only a judgment at the end of a program concerning the program's efficacy. This would be appropriate in a situation in which the agency is trying to decide which one of a number of different existing AOD problem prevention programs would be useful in their community.

Although it has become increasingly common for Indian AOD programs to have an evaluation component, it is surprising how little use is made of the results. Currently, there is no central source where prior evaluations are available to those designing prevention programs. Unless the evaluation has been published in a professional journal, there is no access to the findings. This is unfortunate because a lot has been learned about prevention that could be useful to new programs. It would be helpful to establish a clearinghouse for evaluation results of Indian AOD use prevention programs. This would allow for more rational program planning as well as avoiding the frequent feeling that efforts are being duplicated without the benefit of valuable feedback regarding effectiveness.

Summary and Conclusions

Program evaluation for Indian AOD use/misuse prevention efforts is hardly a new idea; however, it has been applied very unevenly in the past. Some communities have taken it seriously

and implemented highly sophisticated evaluation plans; others, unfortunately, have believed that a highly subjective evaluation would appease program funding sources. Nonetheless, there are a number of barriers to performing good evaluations, and many are the result of the cross-cultural issues that are involved. It is our strong conviction that evaluation needs to be an integral part of programming, but we also believe that not enough effort has been put forth in trying to forge procedures that are culturally congruent and consistent with local lifeways and thoughtways. As long as there is reluctance to confront the many challenging questions posed by this pursuit, evaluation will be seen by tribes as an irrelevant exercise and by evaluators as an impossible task. We hope we have provided some guidance whereby good evaluation can be conducted and the many valuable lessons that are being learned every day in Indian communities can be explained and communicated.

References

- Adair, J., and Deuschle, K. *The People's Health: Anthropology and Medicine in a Navajo Community*. New York: Appleton-Century-Crofts, 1970.
- Berry, J. Introduction to methodology. In: Triandis, H., and Berry, J., eds. *Handbook of Cross-Cultural Psychology*, Vol. 2. *Methodology*. Boston, Mass.: Allyn and Bacon, 1980. pp. 1-28.
- Irvine, S., and Berry, J., eds. *Human Assessment and Cultural Factors*. New York: Plenum, 1983.
- Lonner, W., and Berry, J., eds. *Field Methods in Cross-Cultural Research*. Newbury Park, Calif.: Sage Publications, 1986.
- Manson, S. *Journal of the National Center* (entire edition) 2(3):7-90, 1989.
- Maynard, E. The growing negative image of the anthropologist among American Indians. *Human Organization* 33:402-403, 1974.
- Oetting, E.; Dinges, N.; and Beauvais, F. Self-report social and psychological measures in drug abuse research: Selecting measures that are valid and culturally appropriate. In: Trimble, J.; Bolek, C.; and Niemcryn, S., eds. *Conducting Cross-Cultural Substance Abuse Research: Emerging Strategies and Methods*. Newbury Park, Calif.: Sage Publications, in press.
- Sudman, S., and Bradburn, N. *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine, 1974.

- Trimble, J. The sojourner in the American Indian community: Methodological issues and concerns. *Journal of Social Issues* 33(4):159-74, 1977.
- Trimble, J. Putting the ethic to work: Applying social psychological principles in cross-cultural settings. In: Bond, M., ed. *The Cross-Cultural Challenge to Social Psychology*. Newbury Park, Calif.: Sage Publications, 1988. pp. 109-121.
- Trimble, J., and Hayes, S. Mental health intervention in the psychosocial contexts of American Indian communities. In: O'Conner, W., and Lubin, B., eds. *Ecological Approaches to Clinical and Community Psychology*. New York: Wiley, 1984. pp. 293-321.
- Trimble, J.; Lonner, W.; and Boucher, J. Stalking the wily emic: Alternatives to cross-cultural measurement. In: Irvine, S., and Berry, J., eds. *Human Assessment and Cultural Factors*. New York: Plenum, 1983. pp. 259-273.
- Trimble, J.; Manson, S.; Dinges, N.; and Medicine, B. American Indian concepts of mental health: Reflections and directions. In: Pedersen, P.; Sartorius, N.; and Marsella, A., eds. *Mental Health Services: The Cross-Cultural Context*. Beverly Hills, Calif.: Sage Publications, 1984. pp. 199-220.
- Weiss, C. *Evaluation Research: Methods for Assessing Program Effectiveness*. Englewood Cliffs, N.J.: Prentice Hall, 1972.