Section IV

STRATEGIES THAT ADDRESS CULTURALLY RESPONSIVE EVALUATION

7. A GUIDE TO CONDUCTING CULTURALLY RESPONSIVE EVALUATIONS

Henry T. Frierson, Stafford Hood, and Gerunda B. Hughes

Culture is a cumulative body of learned and shared behavior, values, customs, and beliefs common to a particular group or society. In essence, culture makes us who we are.

In doing project evaluation, it is also important to consider cultural context in which the project operates and be responsive to it. How can an evaluation be culturally responsive? An evaluation is culturally responsive if it fully takes into account the culture of the program that is being evaluated. In other words, the evaluation is based on an examination of impacts through lenses in which the culture of the participants is considered an important factor, thus rejecting the notion that assessments must be objective and culture free, if they are to be unbiased.

Evaluation is based on an examination of impacts through lenses in which the culture of the participants is considered an important factor.

Moreover, a culturally responsive evaluation attempts to fully describe and explain the context of the program or project being evaluated. Culturally responsive evaluators honor the cultural context in which an evaluation takes place by bringing needed, shared life experience and understandings to the evaluation tasks at hand.

Why should a project director be concerned with the cultural context of a program undergoing evaluation? Simply put, as American society becomes increasingly diverse racially, ethnically, and linguistically, it is important that program designers, implementers, and evaluators understand the cultural contexts in which these programs operate. To ignore the reality of the existence of the influence of culture and to be unresponsive to the needs of the target population is to put the program in danger of being ineffective and to put the evaluation in danger of being seriously flawed.

Being sensitive and responsive to the culture of the participants and the cultural environment in which the programs exists is a process that should be an important component of program evaluation. Fortunately, cultural responsiveness as it relates to evaluation is gaining recognition

Cultural responsiveness is gaining recognition as a critical feature of the evaluation process.

as a critical feature of the evaluation process. This is particularly true for programs in which the participants' culture is acknowledged to have a major impact on program outcomes.

The Need for Culturally Responsive Evaluation

It may seem obvious to some, if not to most, professionals that cultural responsiveness should be an integral part of the project development and evaluation process. After all, who could argue against taking into account the cultural context when designing and conducting an evaluation? Doesn't everyone consider the cultural context? The answers to these questions are, respectively, "many" and "no." Apparently, not everyone agrees that implementing culturally responsive evaluation is a good idea. Essentially, there are two frequently stated arguments against using culturally responsive strategies and techniques in educational evaluations. First, there is the claim that evaluations should be culture free. Second, some individuals argue that while an evaluation should take into account the culture and values of the project or program it is examining, it should not, however, be *responsive* to them.

There are no culture-free evaluators, educational tests, or societal laws.

Let us examine the first argument. Just as surely as there are no culture-free evaluations, there are no culture-free evaluators, educational tests, or societal laws. Our values are reflected in our social activities, whether they are educational, governmental, or legal. The responsibility that educational evaluators have is to recognize their own personal cultural preferences and to make a conscious effort to restrict any undue influence they might have on the work.

The second argument, that educational evaluations should not be in the business of *responding* to the cultural contexts in which they are undertaken, is more troublesome. It is one thing to accept or recognize the reasonableness of the requirement to describe the cultural context. It is quite another to adopt evaluation strategies that are consonant with the cultural context(s) under examination. It is precisely this last point of view that is being advocated in this chapter. The field of educational evaluation has advanced over the past three decades, through its recognition of the role that fullness of description plays in a comprehensive evaluation process (e.g., Stake, 1967). In fact, it is becoming increasingly recognized that a responsive evaluation can greatly benefit the project and its stakeholders. Still, it remains all too rare that educational evaluation is designed to be responsive to the cultural context associated with the program or project that is being evaluated.

This chapter discusses strategies that have been found to be useful in conducting culturally responsive evaluation and to identify areas where further help is needed. We examine the role of culturally responsive evaluation at each of the critical phases of the evaluation process, showing how its principles can be applied to enhance good inquiry.

Preparing for the Evaluation

Preparing for the actual evaluation and assembling an evaluation team, is, of course, a critical stage in the evaluation process. At the outset, the sociocultural context in which the programs or projects are based must be taken into account. Situations where programs involve ethnically diverse participants and stakeholders call for the "creation of multi-ethnic evaluation teams to increase the chances of really hearing the voices of underrepresented students" (Stevens, 2000). Stevens reminds us that evaluators may, and often do, *listen* to what stakeholders say when they collect data on site from students, teachers, parents, and other participants or stakeholders. But the crucial question she asks is, do they *hear* what those individuals are saying? Stevens implies that the evaluator or evaluation team must have the "shared lived" experience to truly hear what is being said. At the very least, the evaluator or evaluation team should be fully aware of and responsive to the participants' and stakeholders' culture, particularly as it relates to and influences the program.

Multiethnic evaluation teams increase the chances of really hearing the voices of underrepresented students.

Given the important role of the evaluation team, care should be taken in selecting its members. Those members, whenever possible, should be individuals who understand or who at least are clearly committed to being responsive to the cultural context in which the project is based. Project directors should not, however, assume that racial/ethnic congruence among the evaluation team, participants, and stakeholders equates to cultural congruence or competence that is essential for carrying out culturally responsive evaluations (Thomas, 2001).

Engaging Stakeholders

When designing an evaluation that seeks to be culturally responsive, considerable attention must be given to the identification of the stakeholders. Stakeholders play a critical role in all evaluations, especially culturally responsive ones, providing sound advice from the beginning (framing questions) to the end (disseminating the evaluation results). It is important to develop a stakeholder group representative of the populations the project serves, assuring that individuals from all sectors have the chance for input. Indeed, those in the least powerful positions can be the most affected by the results of an educational evaluation. Students, for example, may qualify for consideration, as might their parents or care givers. When targeting an evaluation toward program improvement and decisionmakers' needs, it is easy to overlook the critical roles that students and parents might play in an educational evaluation.

Stakeholders play a critical role in all evaluations, especially culturally responsive ones.

In individual projects such as the Louis Stokes Alliance for Minority Participation and the Alliance for Graduate Education for the Professoriate, if participants' and stakeholders' perceptions and views are not taken into account from a cultural perspective, the evaluation may prove flawed, particularly if qualitative methods are employed. Moreover, even if quantitative methods are the primary methodological format, the various "voices" should be heard in the interpretation and presentation of the results. It is important that all key voices are accurately heard and listened to. If they are not, the entire evaluation process may be limited in its accuracy and opportunities for meaningful program improvement drastically reduced.

Identifying the Purpose(s) and Intent of the Evaluation

Another important step is to ensure that there is a clear understanding of the evaluation's purpose and intent. Generally speaking, as stated earlier, comprehensive program evaluation is designed to answer two basic questions: (1) Is the project being conducted as planned and is progress being made toward meeting its goals? and (2) Ultimately, how successful is the project in reaching its goals? To answer these questions, three basic types of evaluations are conducted: process, progress, and summative. The first two types of evaluations are called formative evaluations because they measure and describe program operations in order to "inform" project staff (and stakeholders) about the status of the program. Summative evaluations, on the other hand, reveal whether and to what extent the program achieved its goals and objectives.

Culturally responsive progress evaluations examine connections through culturally sensitive lenses.

Process evaluations examine the connections between and among program activities. Culturally responsive process evaluations examine those connections through culturally sensitive lenses. Careful documentation of the implementation of program activities is critical to making sense of the subsequent summative evaluation results. Having an evaluator or a team of evaluators that is culturally sensitive to the program environment will ensure that cultural nuances—large and small—will be captured and used for interpreting progress and summative evaluations.

Progress evaluations seek to determine whether the participants are progressing toward achieving the stated goals and objectives. Culturally responsive progress evaluations help determine whether the original goals and objectives are appropriate for the target population. In seeking to ascertain whether the participants are moving toward the expected outcomes, a culturally responsive progress evaluation can reveal the likelihood that the goals will be met, exceeded, or not exceeded given the program timeline and the results of the process evaluation.

Summative evaluations provide information about program effectiveness. Culturally responsive summative evaluations examine the direct effects of the program implementation on the participants and attempt to explain the results within the context of the program. For example, improved student achievement is influenced by and correlated with a variety of school and personnel background variables. Thus, to fully measure the effectiveness of the program and determine its true rather than superficial worth, it is important to identify the correlates of participant outcomes (e.g., student achievement, student attitudes) and measure their effects as well.

Framing the Right Questions

An important key to successful evaluation is to ensure that the proper and appropriate evaluation questions have been framed. For an evaluation to be culturally responsive, it is critical that the questions of significant stakeholders have been heard and, where appropriate, addressed.

It is critical that the questions of significant stakeholders have been heard and, where appropriate, addressed.

The questions that will guide an educational evaluation are crucial to the undertaking and ultimately to the success of the venture. Poorly framed questions rarely yield useful answers. Further, framing evaluative questions is *not* easily accomplished. In a culturally responsive evaluation, the questions will have been carefully considered not only by the evaluator and project staff, but by other stakeholders as well. It takes time and diligence to reach agreement on the questions to be pursued. One stakeholder group may care little about questions that are seen as vital by another group. However, it is crucial that all significant voices are heard.

Once an agreed-upon list of questions has been articulated to the satisfaction of the evaluation team and stakeholders, an epistemological task of great import comes to the fore, but again, it is not an easy task. They must ask, "What will we accept as evidence when we seek answers to our evaluative questions?" This, too, should be decided *before* embarking on a culturally responsive evaluation. It avoids subsequent rejection of evidence by a stakeholder who might say, for example, "This is interesting, but it really isn't hard data." Stakeholders often will be interested in the results that bear on one group over all others. If one particular group has not been involved or asked questions they consider as key, then the rest of the data may be viewed as suspect or irrelevant.

Questions regarding what constitutes acceptable evidence should be discussed before conducting the evaluation.

Discussions of what is important, and how we will know if we have acceptable evidence, are often messy and may be heated. The discussions, however, are always *necessary*. A more democratic approach to evaluation increases the need for competent evaluators who have a shared lived experience with the stakeholders. A democratic process also increases the likelihood that evaluative efforts will have all voices represented.

Designing the Evaluation

After the evaluation questions have been properly framed, sources of data have been identified, and the type of evidence to be collected has been decided, it is then time to identify the appropriate evaluation design.

There are a number of different evaluation designs that can be used to organize the processes of data collection and analysis and subsequently answer the evaluation questions. The evaluation design that you use does not necessarily need to be elaborate. It just needs to be appropriate for what you want to do.

As stated earlier, most comprehensive evaluation designs have both a qualitative and a quantitative component. Each component provides data in a format that is different from the other, but that can also be complementary to the other.

In addition, designs that incorporate data collection at multiple times provide an opportunity to examine the degree to which some aspect of the participants' behavior changed as a result of the project intervention(s). Furthermore, when comparison or control groups are incorporated into the pre-test/post-test design, evaluators are able to determine to what extent some aspect of participants' behavior changed relative to where it would have been had they not been subject to the project intervention(s).

Selecting and Adapting Instrumentation

Previous use does not guarantee cultural responsiveness.

Instrumentation provides the means for collecting much of the data for program and project evaluation. Therefore, it is very important that instruments be identified, developed, or adapted to reliably capture the kind and type of information needed to answer the evaluation questions. Also at issue is the validity of the inferences about the target population that are drawn from data collected using evaluation instruments. While it is preferable to use instruments that have some history, that have been tried out and have established validity and reliability, previous use does not guarantee cultural responsiveness. Oftentimes, measures that have been normed on a cultural group different from the target population are used in the evaluation process. In such instances, additional pilot testing of the instruments should be done with the cultural group or groups involved in the study to examine their appropriateness. If problems are identified, refinements and adaptations of the instruments should be made so that they are culturally sensitive and thus provide reliable and valid information about the target population.

Collecting the Data

The need to train data collectors in evaluation studies is great.

Culturally responsive evaluation makes substantial use of qualitative evaluation techniques. One of the tenets of qualitative methodology is that the individual who is collecting the data is the instrument. With that in mind, an instrument (or individual) that is an improper measure provides invalid data. Consequently, when collecting qualitative data directly from individuals, e.g., via interviews or observations, if those who are collecting and recording the data are not attuned to the cultural context in which the program is situated, the collected data could be invalid. While it may not appear to matter very much whether a person collecting student test papers in the classrooms is culturally responsive, cultural responsiveness does matter in many forms of data collection. In truth, it may indeed matter *how* the test papers are handed out to the students, *how* the test is introduced, and *what* the atmosphere is at the site where the students are being tested. The situation becomes far more complex in the collection of evaluative information through observations and interviews. The need to train data collectors in evaluation studies is great and, unfortunately, largely overlooked. Training them to understand the culture in which they are working is an even rarer event.

There may not be much an evaluation team can do about the age, gender, race, and appearance of its members, but to deny that such factors influence the amount and quality of the data is imprudent. One thing that can be done to increase the probability of gathering evaluative information in a culturally responsive manner is for the project director to ensure that the principal evaluator and team members involved in the data collection know what they are hearing and observing.

Too often the nonverbal behaviors are treated as "error variance" in the observation and ignored.

Nonverbal behaviors can often provide a key to data interpretation among culturally diverse populations. One African American psychologist, Naim Akbar (1975 as cited in Hale-Benson, 1982), describes a few nonverbal behaviors in African American children. He notes that the African American child "expresses herself or himself through considerable body language, adopts a systematic use of nuances of intonation and body language, such as eye movement and position, and is highly sensitive to others' nonverbal cues of communication." When observing African Americans participating in the program under evaluation, much could be lost toward reaching "understanding." Too often the nonverbal behaviors are treated as "error variance" in the observation and ignored. The same can be true when interviewing an African American program participant and stakeholder. In one sense, the evaluators have to know the territory. For example, Floraline Stevens (2000) described how she and her colleagues overcame difficulties attendant to being responsive to culture during an evaluation project

within a large metropolitan school district. She pointed out that their extensive knowledge of the culture in the classroom and cultural background of the students overcame difficulties in collecting accurate data.

Lack of knowledge about cultural context is quickly evident when interview data are examined. Reviews of interview transcripts and observation protocol data that are done by reviewers without the ability to interpret meaning based on the (largely) unwritten rules of cultural discourse are likely to result in interpretations that are more frequently wrong than right. Similarly, subsequent discussions of flawed reviews limit communication and ultimately doom the possibility of shared understanding between participants and stakeholders of color and the evaluator who proves to be culturally nonresponsive.

Knowledgeable trainers, using the medium of videotaping, can and have produced considerable improvement in the skills of interviewers who must collect data in cultural settings unfamiliar to them. The training process can be very revealing for participants who seek to understand more about the nonverbal language they communicate and their own flawed communication habits. If interviewer training is entered with the spirit of openness and self-improvement, the results for collecting culturally responsive evaluative data can be considerable. Similar improvements in data collection and interpretation through observation can be achieved through intensive training and mentoring. Although the authors commend such training, in-service training is not the preferred solution. Greater and longer lasting improvements in the collection of culturally responsive evaluative data and in the conduct of program evaluations can be realized principally by recruiting evaluation data collectors and analyzers who already possess a shared lived experience with those who are being evaluated.

Analyzing the Data

One may conduct appropriate statistical techniques, such as analyses of variance, and examine test score distributions without much concern for the cultural context in which the data were collected, although that may actually be somewhat shortsighted. But the analysis of interview data and the interpretation of descriptions of behavior related to programs undergoing evaluation cannot be achieved without considerable sensitivity to, and understanding of, the cultural context in which the data are gathered.

Determining an accurate meaning of what has been observed is central in culturally responsive evaluation. Having adequate understanding of cultural context when conducting an evaluation is important, but the involvement of evaluators who share a lived experience may be even more essential. The charge for minority evaluators is to go beyond the obvious.

Knowing the language of a group's culture guides one's attention to the nuances in how language is expressed and the meaning it may hold beyond the mere words. The analyst of data gathered in a culturally diverse context may serve as an interpreter for evaluators who do not share a lived experience with the group being evaluated.

To this end, a good strategy is the creation of review panels principally comprising representatives from stakeholder groups to examine evaluative findings gathered by the principal evaluator and/or an evaluation team. When stakeholder groups composed of separate panels of parents, students, and community representatives, for example, review evaluative findings, the meaning of evaluative data is frequently fresh, and is not always aligned with confirming interpretations. Again, the results of the deliberations of review panels will not lend themselves necessarily to simple, easy answers. Our contention, however, is that they will more accurately reflect the complexity of the cultural context in which the data were gathered.

Disaggregation of collected data is a procedure that warrants increased attention.

Disaggregation of collected data is a procedure that warrants increased attention. Disaggregation of data sets is highly recommended because evaluative findings that dwell exclusively on whole-group statistics can blur rather than reveal important information. Worst still, they may even be misleading. For example, studies that examine the correlates of *successful* minority students rather than focusing exclusively on the correlates of those who fail are important. It can be enlightening to scrutinize the context in which data that are regarded as "outliers" occur. The examination of a few successful students, in a setting that commonly produces failure, can be as instructive for program improvement as an examination of the correlates of failure for the majority.

In sum, the data rarely speak for themselves, but rather are given voice by those who interpret them. The voices that are heard are not only those who are participating in the project, but also those of the analysts who are interpreting and presenting the data. Deriving meaning from data in program evaluations that are culturally responsive requires people who understand the context in which the data were gathered.

Disseminating and Utilizing the Results

Evaluation results should be viewed by audiences as not only useful, but truthful as well.

Dissemination and utilization of evaluation outcomes are certainly important components in the overall evaluation process. Moreover, a critical key is to conduct an evaluation in a manner that increases the likelihood that the results will be perceived as useful and, indeed, used. Culturally responsive evaluations can increase that likelihood. Hence, evaluation results should be viewed by audiences as not only useful, but truthful as well (Worthen, Sanders, and Fitzpatrick, 1997).

Information from good and useful evaluations should be widely disseminated. Further, communications pertaining to the evaluation process and results should be presented clearly so that they can be understood by all of the intended audiences.

Michael Q. Patton (1991) pointed out that evaluation should strive for accuracy, validity, and believability. Patton (1997) further stated that evaluation should assure that the information from it is received by the "right people." Building on his cogent observation we would add that the "right people" are not restricted to the funding agency and project or program administration and staff, but should include a wide range of individuals who have an interest or stake in the program or project.

The dissemination and use of evaluation outcomes should be thought through early when preparing an evaluation, that is, during the evaluation-planning phase. Moreover, the use of the evaluation should be firmly consistent with the actual purposes of the evaluation. Further, the purpose of the evaluation should be well defined and clear to those involved in the project itself.

As we talk about dissemination, our discussion comes full circle, and we return to the earliest steps in evaluation design, the evaluation questions. These questions themselves are always keys to a good evaluation—those that would provide information that stakeholders care about and on which sound decisions can be based must always guide the work. The right questions, combined with the right data collection techniques, can make the difference between an evaluation that is only designed to meet limited goals of compliance and one that meets the needs of the project and those who are stakeholders in it. Applying the principles of culturally responsive evaluation can enhance the likelihood that these ends will be met, and that the real benefits of the intervention can be documented.

References

Gordon, E.W. (1998). Producing Knowledge and Pursuing Understanding: Reflections on a Career of Such Effort. AERA Invited Distinguished Lectureship. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA, 13 April.

Hale-Benson, J. (1982). *Black Children: Their Roots, Culture, and Learning Styles*, Revised Ed. Baltimore, MD: Johns Hopkins University Press.

Hood, S. (2000). Commentary on Deliberative Democratic Evaluation. In *Evaluation as a Democratic Process: Promoting Inclusion, Dialogue, and Deliberation*, edited by K. Ryan and L. DeStefano. New Directions for Program Evaluation, No. 85. San Francisco, CA: Jossey-Bass.

Kahle, J.B. (2000). Discussant Remarks. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals*, NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.

Kirkhart, K.E. (1995). Seeking Multicultural Validity: A Postcard From the Road. *Evaluation Practice*, 16 (1): 1-12.

Patton, M.Q. (1991). Toward Utility in Reviews of Multivocal Literatures. *Review of Educational Research*, 61(3): 287-292.

Patton, M.Q. (1997). *Utilization-Focused Evaluation: The New Century Text.* Thousands Oaks, CA: Sage Publication, Inc.

Stake, R. (1967). The Countenance of Educational Evaluation. *Teachers College Record*, 68: 523-540.

Stake, R. (1980). Program Evaluation, Particularly Responsive Evaluation. In *Rethinking Educational Research,* edited by W.B. Dockrell and D. Hamilton. London: Hodder & Stoughton.

Stevens, F.I. (2000). Reflections and Interviews: Information Collected about Training Minority Evaluators of Math and Science Projects. In *The Cultural Context of Educational Evaluation: The Role of Minority Evaluation Professionals,* NSF 01-43. Arlington, VA: National Science Foundation, Directorate for Education and Human Resources.

Thomas, V.G. (2001). Understanding and Framing Talent Development School Reform Evaluation Efforts. Talent Development School Reform Evaluation Guide. Washington, DC: Howard University. Center for Research on the Education of Students Placed at Risk (CRESPAR). Unpublished report.

Worthen, B.R., Sanders, J.R., and Fitzpatrick. (1997). *Educational Evaluation*, Second Ed. White Plains, NY: Longman, Inc.