Strengthening Evaluation through Cultural Relevance and Cultural Competence

Karen E. Kirkhart, Syracuse University
kirkhart@syr.edu

Rodney K. Hopson, Duquesne University
hopson@duq.edu

Introduction to Case Scenario: Dialogue for Diversity and Social Change (DDSC)

From 2004 to 2006, a pilot program in a declining industrial city in the northeast corridor of the U.S. helped citizens develop commitments to each other and to the common good amidst what the program saw as decreased social cohesion and declining social capital. *Dialogue for Diversity and Social Change[1]* aimed to strengthen ethical leadership and to shift civic discussion and dialogue to a culture of shared exploration and renewed focus on relevant issues in the city, state, and nation. *DDSC* program founders considered this an innovative response and the essence of what they referred to as their "core practices"– providing a new kind of civic space in which reflection, connection, shared meaning-making, and substantive change can occur. Diverse groups of 8 to 10 citizens, who sensed a need to do more with their lives, met over dinner for seven consecutive weeks to discuss relevant social issues. Participants were referred by religious congregations, civic and community groups, public agencies and arts organizations. The program used poems, music, photographs, and art as a focus point of discussion and connection. Topics discussed at the meetings included race/diversity, economic disparity, environment, and materialism, to name a few. DDSC very much saw itself as addressing the declining civic engagement and social cohesion in the U.S., and in this declining steel, coal, and blue collar major mid-size city. By creating civic structures and dialogue in the northside section and other places where participants of the city reside, the program intended that a public, civic language would emerge that is less polarized and more reflective and would help to (re)connect communities and citizens across difference. The program furthermore intended to engage the many people who were not attracted to issues discussions, but sensed a need to connect more meaningfully with themselves, with others and with the common good. Participants may not have seen themselves as "public types," or they were already engaged, but they were equally weary of interest group disputes. Participant profiles varied; roles included belonging to a religious congregation, providing leadership on a non-profit board, working in a minimum wage job, or serving in a government office.

---

[1] A pseudonym.

Initial Discussion Questions

1. What elements of culture, at what levels, seem salient to this scenario at first glance?
2. How do your own cultural positions/contexts relate to the cultural elements in the scenario?
3. What perspectives and/or characteristics of culture are you assuming will *not* be as salient, based upon your initial impressions?

DDSC Case: Preparation for Stages 1- 3

Recently ended in fall 2006 due to lack of continued operational and financial support, DDSC was a program of the Light House Learning Center (LLC), an outgrowth of a 1917 settlement house that currently operates a variety of programs to help disadvantaged children increase their power of agency and understanding, through cutting-edge digital story-telling, literacy initiatives, and high-speed connections with other community partners. Formal governance of DDSC was through the LLC's 501c3 board. DDSC brought together stakeholders with divergent perspectives, who wanted to engage in trust-building through the seven week dialogue rounds, with opportunity to continue dialoguing after the seven weeks conclude.

By 2006, over 100 people of varied ethnic and social groups participated, with each participant averaging 5.6 sessions per 7-session round. Upon completion of their 7-session round, participants had the opportunity to participate in "Continued Dialogues," additional open-ended meetings with other completed participants. Of the total number of DDSC participants, the majority was white, one-quarter were African American, 5% were of Middle Eastern origin, and an additional 3% were of Asian or Latino/a background. Slightly over half of the DDSC participants (52%) were female.   Incomes ranged from under $15,000 to over $75,000, with the bulk in the middle income range, half (50%) from two-income households.   Most of the participants possessed college degrees and some had done post-graduate work, though fewer than 10% had had only modest formal education.

The program director and board chair of DDSC were the main initiators of the evaluation. Their interest was primarily in quantitative measures that help understand the knowledge, attitudes, and behavior of the participants during pilot stages and the initial year. A management/consulting firm with a specialty in "strategy development and planning services for nonprofits, foundations, community collaboratives, and government agencies" was hired to help understand these data during the initial start-up operation. In the beginning of the second year, a university center of evaluation joined the evaluation team to "get beneath" the survey data and to document participant stories and lives for future funding purposes. The evaluation team was made up of two staff (manager and staff associate) from the management/consulting firm, a member of the education faculty and a graduate student in liberal arts. The graduate student participated in a dialogue round prior to becoming a member of the evaluation team.

The purpose of the evaluation was outcome and impact-focused, largely driven by the hope of knowing how the program worked, what change occurred for its participants, and the impact of the program on the participants' social and community understandings. The program director and advisory board chair initiated the evaluation efforts with hopes of sustaining and developing DDSC both within the metropolitan city and beyond.

Stage 1: Prepare for the Evaluation
Stage 2: Engage Stakeholders
Stage 3: Identify Purpose of the Evaluation

Discussion Questions, Stages 1-3

1. What elements of background and context are important here? What more would you want to know?
2. Who was included on the evaluation team and what presumed skills and traits do they bring to the evaluation process?
3. Based on the stated purpose of this evaluation, who do you understand to be the major stakeholders?

DDSC Case: Preparation for Stages 4-6

By the end of the first year, two evaluation questions guided the evaluation undertaking: i) How is DDSC making a difference? and ii), How do the DDSC core practices (i.e., engaging diverse groups of citizens in sustained conversations; using poems, music, photographs, art and ordinary human voices as a focal point for connection; self-reflection; and the practice of a new civic engagement) matter? Significant foundation resources were invested in the operation of the program and in the second year, another major foundation invested evaluation resources to consider another question, iii) "How can we best evaluate complex, embedded learning experiences among participants and stakeholders in the program?" The design summary table on the following pages outlines evaluation questions, information sources, data collection strategies, procedures for gathering information, and data analysis and reporting procedures.

As the evaluation plan developed, the consulting/management firm and university center for evaluation derived three outcomes of interest from the three evaluation questions. As identified by the evaluation team, together with the Executive Director, the intended program outcomes included:

i)   Ability to identify change in relationships among diverse groups of people, including participants; participants take concrete steps to advance civic commitments; participants develop more than "just talk" but explore issues of conversations in their lives,

ii) Ability to influence local community, familial, and social networks toward social change; heightened community and civic engagement in metropolitan city by participants; establishment and maintenance of a meeting place for diverse groups of participants to convene and organize social activism, and

iii) Ability to make program decisions to develop new communities of conversation with participants in different context and structures; opportunity for program staff to identify "lessons learned" for implications to broader community.

DDSC took its evaluation seriously and was heavily evaluated for a program of its size. A variety of evaluation components, including ethnographic portraits, interviews, focus groups, and surveys were conducted either by evaluators from the university or the consulting/management firm. Significant dialogue between evaluation team members and DDSC's Program Director allowed for continual fine tuning of the evaluation process to capture as much of the most relevant information as possible.

Evaluation Design Summary Table *(Adapted from Fitzpatrick, Sanders & Worthen, 2004)*

Evaluation Questions	Information needed to answer the question	Information Sources	Data Collection Strategies	Procedures for gathering information			Data Analysis & Reporting Procedures
				Who?	**How?**	**When?**	
1: *(How) is DDSC making a difference?*	· Does change happen following dialogues? · What change occurs? ·What relationships are created? · What participation occurs in and beyond dialogues? · What practices/ activities make a difference?	Participants	Focus groups	Consulting firm/ University evaluators	Total of 3 participant focus groups held. Audiotaped.	Summer 05 – Spring 06	Focus group results transcribed, grouped by content areas. Quotations used to illustrate key ideas.
		Participants	Individual interviews	Consulting firm/ University evaluators	Participants interviewed by telephone. 10-20 mins. Tape recorded.	Summer 05	Recordings were transcribed, coded for content. Results summarized thematically. Narrative summary plus illustrative quotes.
		Participants	Surveys	Consulting firm/ University evaluators	Participants complete survey before starting DDSC, post-DDSC, and one year following completion.	Winter 05 – Summer 06	Quantitative analysis via descriptive statistics, reporting percentages of respondents per item option. Trends noted in data over time.
		Community members	Individual interviews	University evaluators/ Consulting firm	Key informants in community agencies that collaborated with DDSC interviewed about the impact of their collaboration. Community members who participated in DDSC interviewed about agency impact.	Summer 05 – Spring 06	Individual and group interviews transcribed, coded for content. Results summarized thematically, including illustrative quotes.

Evaluation Questions	Information needed to answer the question	Information Sources	Data Collection Strategies	Procedures for gathering information			Data Analysis & Reporting Procedures
				Who?	How?	When?	
Q2: *(How) do DDSC core practices matter?* Core practices: diverse groups; use of arts to facilitate discussion; self-reflection (shared looking, personal journeys); civic engagement	· What factors encourage participants to adopt a more generous stance toward social change? · What kind of impact did program have on the participant's social and community understandings? · What were participants' reactions to structure and goals of the program? · Do the arts create meaningful dialogue and spark ethical reflection? · What do dialogue practices influence?	Participants	Focus groups	Consulting firm/ University evaluators	Questions about core practices: diverse groups, facilitated discussion, shared looking, personal journeys. Audiotaped.	Winter 05 – Spring 06	Focus group results transcribed, grouped by content areas. Quotations used to illustrate key ideas.
		Evaluators	Participant observation	University evaluators	Evaluators enroll in DDSC, participate as group members through dialogue (& continued) rounds. Record notes on participant discussions.	Fall 04 – Spring 06	Field notes taken, written up for content to inform focus group surveys and interviews.
		Participants	Programmatic survey	Consulting firm	Written surveys on program practices administered to participants completing dialogue rounds.	Summer 05 – Spring 06	Descriptive statistics calculated by question. Percentage responding to each option. Narrative summary of modal response patterns.
		Participants	Ethnographic interviews	University evaluators	Repeated observation, dialogue, interviews over a period of several months. Audiotaped.	Winter 05 – Spring 06	Tapes transcribed. Content analysis. Results summarized narratively by theme.

Evaluation Questions	Information needed to answer the question	Information Sources	Data Collection Strategies	Procedures for gathering information			Data Analysis & Reporting Procedures
				Who?	**How?**	**When?**	
Q3: *How can we best evaluate complex, embedded learning experiences?*	· What learnings occurred? · How do learnings impact participant population and community? · How do we document learnings in complex participant spaces?	Program documents	Review mission statement	University evaluators	Review DDSC mission statement to extract characteristics of a complex, imbedded learning system, develop program theory.	Winter 05 – Spring 06	Key characteristics used in data collection, data analysis.
		Literature	Review literature	University evaluators	Read and summarize theory literature on democratic and responsive evaluation.	Summer 05 – Spring 06	Evaluation design grounded in culturally relevant theory.
		Participants	Ethnographic interviews	University evaluators	Two participants who had completed DDSC were interviewed over several months. Audiotaped.	Winter 05 – Spring 06	Tapes transcribed. Content analysis. Results summarized narratively by theme.
		Evaluators	Documents analysis	University evaluators	Proceedings from evaluation team meetings were recorded.	Winter 05 – Spring 06	Reflexive analysis of team meeting notes, proceedings. Data used to track evaluator involvement, impact, inform data collection. Program staff review data collected for program development, improvement.

Stage 4: Frame the Right Questions
Stage 5: Design the Evaluation
Stage 6: Select and Adapt Instrumentation

Discussion Questions, Stages 4-6

1. What/whose perspectives are represented in evaluation questions and what other questions might have been posed?
2. Whose perspectives were accepted as credible evidence? Credible to whom?
3. How well did the time frame in this study match the needs and rhythms of this context?

DDSC Case: Preparation for Stages 7-9

Based upon DDSC's mission emphasis that social issues and decline in civic engagement be addressed according to its core practices of sustained, facilitated conversations, it was clear that serious study of *multiple dynamics* was necessary at various levels of the evaluation work—not only among the program participants, but also among other stakeholders, organizations affiliated with the program, and among evaluation team members. The evaluation team reviewed the program's mission in order to understand how DDSC differed from other community initiatives and what elements distinguished it as a "complex, embedded, learning system."

Findings revealed that DDSC made a difference in the following ways for participating individuals and agencies:

❖ Improved level of civic engagement/action.
❖ Reinforced practice of using art as a reminder or example of how others see things or of a personal belief and as a resource for exploring different ideas.
❖ Increased or enhanced participants' receptivity to opportunities to commit or get involved in contributing to the common wealth.
❖ Reinvigorated, refocused and/or reenergized participants in relation to maintaining their civic commitment or involvement.

Core Practices such as diverse group composition, use of the arts to facilitate discussion, self-reflection, and civic engagement affected participants in the following ways:

❖ Most participants found discussions centered on arts / humanities materials helpful or extremely helpful.
❖ Focal points (change maker profiles, visual images, poems, etc.) were particularly useful in stimulating conversation and evoking reactions.
❖ The following overall benefits from participation in DDSC were reported by at least 50% of participants:
  · *Renewed sense of possibilities*
  · *Challenged participants to make a specific commitment*
  · *Provided space for ethical reflection*

❖ Group diversity provided different perspectives on issues and allowed for diverse interactions around issues that would not normally have occurred.

In June, 2006, the evaluation team presented its findings from the two year evaluation study to the Executive Director and a few advisory board members in an LLC office. In addition to a presentation, a summary report of key questions answered and accompanying appendices was provided to assist and potentially leverage support for continued dialogues.

Stage 7: Collect the Data
Stage 8: Analyze the Data
Stage 9: Disseminate and Use the Results

Discussion Questions, Stages 7-9

1. What additional data collection procedures might have been useful to consider in designing a culturally responsive evaluation?
2. Given findings as briefly summarized, what aspects of cultural context might add meaning to guide recommendations?
3. Were results shared in a culturally congruent way?

Strengthening Evaluation through Cultural Relevance and Cultural Competence

Karen E. Kirkhart
Syracuse University
kirkhart@syr.edu

Rodney K. Hopson
Duquesne University
hopson@duq.edu

Multicultural Validity

All evaluative understandings and judgments are grounded in culture. Multicultural validity refers to the correctness or authenticity of understandings across multiple, intersecting cultural contexts (Kirkhart, 1995). It focuses attention on how well evaluation captures meaning across dimensions of cultural diversity, and it scrutinizes the accuracy or trustworthiness of the ensuing judgments of merit and worth. Like validity in general, it is a multifaceted construct, permitting one to explore the many ways in which culture impacts meaning and understanding. Multicultural validity may be argued and understood in terms of methodology, consequences, interpersonal relationships, lived experience, and theory (Kirkhart, 2005). Each justificatory perspective directs attention to a different type of evidence to support validity. Figure 1 summarizes the five justifications. Methodological justifications of multicultural validity direct attention to the validity of measurement and design elements. Interpersonal justifications scrutinize relationships among the researcher(s) and the researched. Experiential justifications examine validity in terms of the lived experience of program participants. Invoking theoretical justifications of multicultural validity leads to scrutiny of theoretical foundations. Consequential justifications examine the impacts or sequelae of evaluation to reflect on validity. Table 1 provides examples of arguments used to support validity claims under each of these justifications. These arguments interact; they are not independent.

Failure to address culture threatens the validity of evaluative understandings and actions. Threats are specific reasons why inferences may be partly or completely wrong. The five perspectives that define supporting justification can also point to errors of either omission or commission that threaten validity. Table 2 summarizes validity threats corresponding to each of the five justifications. The likelihood of any given threat occurring depends on context, and background knowledge is required to appreciate how a specific threat operates.

Though culture belongs at the center of any conversation about validity, in practice it has often been excluded. Multicultural validity moves considerations of culture to the center of validity arguments.
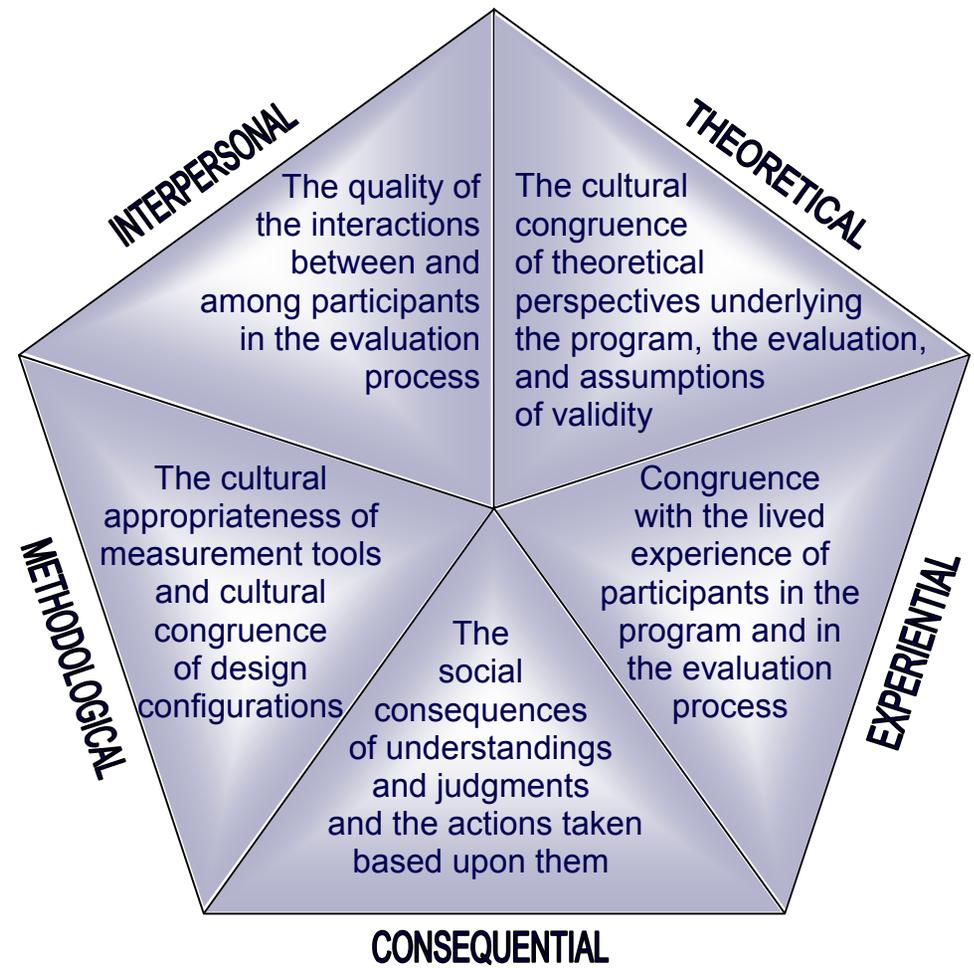
Figure 1

Table 1: Summary of Justifications of Multicultural Validity

Justifications	Examples	Illustrative Probe Questions
Methodological Validity is supported by the cultural appropriateness of measurement tools and cultural congruence of design configurations	• Measurement tools have been developed for a particular ethnic group and validated for a particular use. • The sampling frame insures inclusion of diverse cultural perspectives appropriate to the program being evaluated and its context. • The study design employs a time frame appropriate to the context. • Evaluation questions represent a range of perspectives, values and interest.	• Whose values were represented in the evaluation questions you chose? • What procedures did you use to gain multiple perspectives? • How did the sources of information included in the evaluation permit more than one perspective to come forward? • Did participants who provided evaluation data represent the full range of consumer diversity? • In what ways were the data collection tools you used congruent with the project itself?
Interpersonal Validity is supported by the quality of the interactions between and among participants in the evaluation process	• Evaluators respect local norms and authority in entering the community to undertake evaluation. • Evaluators take time to build relationships and understandings as part of the early process of planning and design development. • Evaluators reflect on their own cultural positions and positions of authority with respect to other participants in the evaluation process. • Meaningful roles are established for stakeholder participation and barriers to full participation are addressed.	• What roles were created for stakeholders to participate in this evaluation? • Was the time frame for the evaluation sufficient or did it feel rushed? • What steps were taken to establish trust with participants in the program? • Did evaluators understand their position vis-à-vis the local community (were they seen as insiders or outsiders?) and the program itself (what authority did they hold?) • Were some participants better represented in the evaluation than others? If participation in the evaluation was unequal, how were barriers to evaluation participation addressed? • Were the data collected confidential? Were they anonymous? What were procedures for maintaining either or both?
Theoretical Validity is supported by culturally congruent theoretical perspectives	• Evaluators select culturally appropriate evaluation theory to frame their epistemology, methods and procedures. • Program theory is grounded in multiculturally valid social science research. • Program theory is grounded in the cultural traditions and beliefs of program participants. • Validity theory itself is examined for culturally-bound biases and limitations.	• Was there a theory base underlying the evaluand? • Did the program theory take culture into account? • Was there a theory base underlying the evaluation—e.g., Culturally Responsive Evaluation (CRE)? How well did it address culture? • How was the validity of this evaluation argued (i.e., what were the warrants of validity claims?)

Justifications	Examples	Illustrative Probe Questions
Experiential Validity is supported by the lived experience of participants	• Local citizens and program consumers contribute their wisdom to the evaluation process. • Evaluators reflect on their own history and cultural positions, seeking assumptions and "blind spots." • Evaluators employ a cultural guide to increase their understanding and appreciation of local culture. • Evaluative data are understood in terms of the realities of the people they represent.	• How well were program consumers represented as sources of information in this evaluation? Program providers? Community or public? • Was a "cultural guide" needed or used? Why or why not? • How did your own personal characteristics and cultural location impact the evaluation? • How did participants and /or providers of the program contribute to the interpretation of the data? Were findings "checked" with them?
Consequential Validity is supported by the social consequences of understandings and judgments and the actions taken based upon them	• History of evaluation in this community is acknowledged and addressed, especially if that history is oppressive, exploitive. • Mechanisms are identified and negotiated by which evaluation will give back to the community. • Evaluation improves the ability of the community to advance its goals and meet the needs of its members. • Evaluation promotes social justice.	• In what ways has evaluation historically interfered with or supported the program? • How does this evaluation itself support the goals of the program? • How does the evaluation relate to social justice (or does it)? • Did the evaluators build in any "give back" to the community?

References

Kirkhart, K. E. (1995). Seeking multicultural validity: A postcard from the road. *Evaluation Practice*, *16*(1), 1-12.

Kirkhart, K. E. (2005). Through a Cultural Lens: Reflections on validity and theory in evaluation. In S. Hood, R. K. Hopson & H. T. Frierson (Eds.) *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding in evaluative theory and practice* (pp. 21-39). Greenwich, CT: Information Age Publishing, Inc.

Ridley, C. R., Tracy, M. L., Pruitt-Stephens, L., Wimsatt, M. K., & Beard, J. (2008). Multicultural assessment validity. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological and educational applications* (3rd ed., pp. 22-33). New York: John Wiley & Sons.

Table 2. Summary of Threats to Multicultural Validity

Threats	Examples	Illustrations
Methodological Validity is threatened by using cultural inappropriate measurement tools, culturally incongruent design configurations, or by procedural errors.	Design incongruence	Selecting a research design that violates cultural norms (e.g., for American Indians, between-group designs comparing people, school or tribes are incongruent with deeply held values) or employs a time frame inappropriate to the context.
	Limited selection	Sampling frame fails to insure diverse representation within cultural subgroups.
	Construct invalidity of cultural variables	Cultural variables are inaccurately defined. This can occur through underrepresentation (e.g., race accepted as a simplistic marker for more complex set of phenomena) or construct irrelevant variance (e.g., attaching prejudicial stereotypes to cultural variables).
	Measurement invalidity	Measurement tools have been developed on majority populations and not validated for use in culturally-specific contexts. Interpretation uses majority norms.
	Language non-equivalence	Failure to translate into languages appropriate to context. Use of inaccurate translation procedures. Ignoring oral traditions and relying on written communication.
	Procedural violations of personal rights, protections	Coercion used to gain participation. Failure to establish and follow protocols for maintaining confidentiality, anonymity. Failure to respect ownership of knowledge or information.
	Limited framework	Evaluation questions are framed from a single perspective, failing to consider alternative values and interests (e.g., provider perspective reflected but not consumers or community). Restricted information sources provide limited range of answers (e.g., program participants but not those who found program culturally offensive).
Interpersonal Validity is threatened by flawed interactions between and among participants in the evaluation process.	Inappropriate entrance	Local norms and authority structures are bypassed, ignored or violated in entering the community to perform evaluation.
	Rushing the agenda	Evaluators move purposefully ahead in their activities without taking time to build rapport and relationship with community members

Threats	Examples	Illustrations
	Limited cultural communication	Evaluators do not speak the language(s) of the local community or are uninformed about oral and written traditions, including symbols and ceremonies.
	Violation of trust	Evaluators fail to maintain the transparency of process and communication required to establish and maintain trust in their integrity.
	Barriers to engagement	No meaningful roles are established to permit genuine engagement; participation is held at superficial levels of token representation.
	Differential power	Evaluators fail to consider their own cultural position and the accompanying dynamics of power that impact interpersonal communication.
Theoretical Validity is threatened by culturally incongruent theoretical perspectives	Majority evaluation theory	Majority evaluation theory is applied to culturally-specific contexts without critical reflection or adaptation.
	Culturally biased social science-based program theory	Program theory is based upon social science research that itself was culturally biased.
	Transformation bias in program theory	Social science research is inappropriately translated or applied to program theory in ways that are insensitive to local culture.
	Validity theory narrowly defined	Only majority understandings of validity theory are accepted as the standard of "scientific rigor."
Experiential Validity is threatened by disconnection from the lived experience of program participants, evaluation participants, and community members	Invalidation of lived experience	Lived experience is reframed or recast in such a way that original voice and meaning are distorted, compromised or obscured. Misappropriating the lived experiences of others.
	Exclusion of experiential evidence	Local citizens and/or program consumers are not invited to contribute their wisdom to the evaluation process.

Threats	Examples	Illustrations
	Personal history	Evaluators fail to reflect on the implications of their own lived experience, both in terms of cultural strengths and limitations.
	Lack of cultural knowledge	Evaluators fail to inform themselves appropriately of the history, background, knowledge, values, and traditions surrounding the evaluand.
	Acultural synthesis	Failure to interpret the data in terms of the lived realities of the people they represent.
Consequential Validity is threatened by failure to attend to the social consequences of evaluative judgments and the actions taken based upon them	Failure to track consequences	Failure to track the consequences of understandings and actions as a reflexive check on validity. Assuming that consequences cannot be known or traced to the evaluation. Assuming that consequences are irrelevant. Assuming that evaluation influence will be positive.
	Exploitation/non-reciprocation	Evaluation gathers information from the site, but it does not address ways in which it will give back to the program or community being evaluated.
	Disempowerment	Evaluation is designed in such a way that it does not improve the ability of the program or community to advance its goals or meet the needs of its members.
	Oppression	The evaluation exacerbates inequity or undermines social justice.
	Ahistorical perspective	The evaluation fails to examine the prior history of evaluation in relation to this program or community. This is a particularly serious omission if that history is oppressive, exploitive.

Revised 6/09KEK

References

Kirkhart, K. E. (2009, February). Establishing multicultural validity in evaluation. Workshop presented to the Graduate Education Diversity Internship Program, American Evaluation Association/Duquesne University. Washington, DC.

References

Barnouw, V. (1985). *Culture and personality* (4th ed.). Homewood, IL: The Dorsey Press.

Centers for Disease Control and Prevention (1999). Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report,* 48(RR11), 1-40.

Conner, R. F. (2004). Developing and implementing culturally competent evaluation: A discussion of multicultural validity in two HIV prevention programs for Latinos. In M. Thompson-Robinson, R. Hopson & S. SenGupta (Eds.), *In search of cultural competence in evaluation: Toward principles and practices, New Directions for Evaluation*, Number 102 (pp. 51-65). San Francisco: Jossey-Bass.

Frierson, H. T., Hood, S., & Hughes, G. B. (2002). Strategies that address culturally-responsive evaluation. In J. Frechtling (Ed.), *The 2002 user-friendly handbook for project evaluation* (pp. 63-73). Arlington, VA: National Science Foundation.

Hopson, R. K. (2003). Overview of multicultural and culturally competent program evaluation: Issues, challenges and opportunities. Woodland Hills, CA: The California Endowment.

Hopson, R. K. (2009). Reclaiming knowledge at the margins; Culturally responsive evaluation in the current evaluation moment. In K. Ryan & J. B. Cousins (Eds.) *The SAGE international handbook of educational evaluation.* Thousand Oaks, CA: Sage.

Jay, M., Eatmon, D., & Frierson, H. (2005). Cultural reflections stemming from the evaluation of an undergraduate research program. In S. Hood, R. K. Hopson & H. T. Frierson (Eds.) *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding in evaluative theory and practice* (pp. 201-216). Greenwich, CT: Information Age Publishing, Inc.

King, J. A., Nielsen, J. E., & Colby, J. (2004). Lessons for culturally competent evaluation from the study of a multicultural initiative. In M. Thompson-Robinson, R. Hopson & S. SenGupta (Eds.), *In search of cultural competence in evaluation: Toward Principles and Practices, New Directions for Evaluation*, Number 102 (pp. 67-80). San Francisco: Jossey-Bass.

Kirkhart, K. E. (1995). Seeking multicultural validity: A postcard from the road. *Evaluation Practice*, *16*(1), 1-12.

Kirkhart, K. E. (2005). Through a Cultural Lens: Reflections on validity and theory in evaluation. In S. Hood, R. K. Hopson & H. T. Frierson (Eds.) *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding in evaluative theory and practice* (pp. 21-39). Greenwich, CT: Information Age Publishing, Inc.

Kirkhart, K. E. (In press). Eyes on the prize: Multicultural validity and evaluation theory. *American Journal of Evaluation*, *31*(3).

LaFrance, J. (2004). Culturally competent evaluation in Indian Country. In M. Thompson-Robinson, R. Hopson & S. SenGupta (Eds.), *In search of cultural competence in evaluation: Toward Principles and Practices, New Directions for Evaluation*, Number 102 (pp. 39-50). San Francisco: Jossey-Bass.

LaFrance, J., & Nichols, R. (2010). Reframing evaluation: Defining an indigenous evaluation framework. *The Canadian Journal of Program Evaluation*, *23*(2), 13-31.

Manswell Butty, J. L., Reid, M. D., & LaPoint, V. (2004). A culturally responsive evaluation approach applied to the Talent Development School-to-Career Intervention Program. In V. G. Thomas & F. I. Stevens (Eds.), *Co-constructing a contextually responsive evaluation framework: The Talent Development Model of School Reform, New Directions for Evaluation*, Number 101 (pp. 37-47). San Francisco: Jossey-Bass.

Mathie, A., & Greene, J. C. (1997). Stakeholder participation in evaluation: How important is diversity? *Evaluation and Program Planning*, 20(3), 279-285.

Nieto, S. (1999). *Affirming diversity: The sociopolitical context of multicultural education* (3rd ed.). Boston: Allyn & Bacon.

Orlandi, M. A. (Ed.) (1992). *Cultural competence for evaluators: A guide for alcohol and other drug abuse prevention practitioners working with ethnic/racial communities*. U.S. Department of Health and Human Services, Office for Substance Abuse Prevention. DHHS Publication No. (ADM)92-1884.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Prilleltensky, I., & Nelson, G. (1997). Community psychology: Reclaiming social justice. In D. R. Fox & I. Prilleltensky (Eds.), *Critical psychology: An introduction* (pp. 166-183). London: Sage Publications.

Ridley, C. R., Tracy, M. L., Pruitt-Stephens, L., Wimsatt, M. K., & Beard, J. (2008). Multicultural assessment validity. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological and educational applications* (3rd ed., pp. 22-33). New York: John Wiley & Sons.

SenGupta, S., Hopson, R., & Thompson-Robinson, M. (2004). Cultural competence in evaluation: An overview. In M. Thompson-Robinson, R. Hopson & S. SenGupta (Eds.), *In search of cultural competence in evaluation: Toward Principles and Practices, New Directions for Evaluation*, Number 102 (pp. 5-19). San Francisco: Jossey-Bass.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (In press). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.