

CULTURAL RESPONSIVENESS IN STEM PROGRAM PLANNING,  
IMPLEMENTATION, AND EVALUATION

A Thesis

Presented to the Faculty of the Graduate School  
of Cornell University

In Partial Fulfillment of the Requirements for the Degree of  
Master of Psychology

by

Wanda Daniella Casillas

January 2012

© 2012 Wanda Daniella Casillas

## ABSTRACT

The recent increase in STEM and science-based programs targeted to minority students must be met with appropriate culturally responsive practices (Johnson, 2005). The current project was designed to investigate ways in which program evaluators and staff implement culturally responsive practices. Evaluators across the country and program staff in New York, California, and Texas were invited to participate in a multiphase concept mapping project to 1) *Brainstorm* culturally responsive practices, 2) *Sort* or organize these statements according to themes of their own choosing, and 3) *Rate* each statement on *importance* and *feasibility* with respect to their practice. We summarize results of the structured conceptualization effort in comparison to the theoretical literature, discuss statistical differences between perceptions of *Importance* and *Feasibility* of practices, and suggest activities that consolidate and align practices as conceptualized by each group in a way that makes principles actionable.

## BIOGRAPHICAL SKETCH

Wanda Casillas received her Bachelor's degree from the University of Texas at San Antonio and is currently a Ph.D. student in the Department of Human Development at Cornell University. Her research interests follow two lines of investigation that converge on a central concern for the learning, cognitive development, and performance of minority students in the sciences. These interests reflect her personal journey to becoming educated and lead an investigation into which paths to educational success align with the interests and talents of underrepresented minority students.

The first line of research draws from literature on scientific reasoning to frame a discussion for how individuals are believed to reason about science problems (causal reasoning) and proposes investigations for the role of cultural experience in scientific reasoning and learning in adolescents. Mrs. Casillas' second line of research focuses on the ecological space surrounding the cognitive development and performance of minority students through an investigation of appropriate program evaluation methods. In this research, Mrs. Casillas focuses on a systems approach to program evaluation which strengthens the cultural responsiveness of evaluation methods used in science-based programs which target minority student populations.

*This work is dedicated to my husband, David, whose unconditional support brought these efforts to fruition- may my efforts validate your truest sacrifices; And to my mother, may she rest in peace, for she gave me the strength, desire, and motivation to live fiercely and fight with courage in every challenge I encounter.*

## ACKNOWLEDGMENTS

I offer my sincerest gratitude to my committee, William Trochim, Ph.D., Barbara Koslowski, Ph.D., and Wendy Williams, Ph.D., for their unwavering support and patience.

Dr. Trochim provided any resources necessary to plan, implement, and complete the work contained herein and the work to follow. Dr. Koslowski has given me the tools necessary to navigate my professional development and my place in fields of interest to me, and Dr. Williams has challenged me to understand the expectations of others in addition to my own. My academic accomplishments would not be possible were these individuals not willing to concede that there are many ways to be accomplished and successful.

I also extend many thanks to the community of culturally responsive evaluators that have set a foundation for the work of generations to follow, especially Dr. Rodney Hopson for his mentorship through the *American Evaluation Association GEDI Fellowship Program* and to my fellow GEDI alum.

## TABLE OF CONTENTS

	Title	Page
	<i>Biographical Sketch</i>	iii
	<i>Dedication</i>	iv
	<i>Acknowledgements</i>	v
	<i>Table of Contents</i>	vi
	<i>List of Figures</i>	vii
	<i>List of Tables</i>	viii
I.	Introduction	1
II.	Background	2
III.	Methods	10
IV.	Analysis	20
V.	Results	26
VI.	Discussion	28
VII.	References	32
VIII.	Appendices	36

## LIST OF FIGURES

Number	Title	Page
1	Number of participants that completed each or multiple sections at once	13
2	Example of a similarity matrix for one participant	21
3	Point rating map of 97 culturally responsive statements	22
4	Fifteen cluster map of culturally responsive statements	24
5	Pattern match of Importance and Feasibility ratings	25

## LIST OF TABLES

Number	Title	Page
1	Type of programs participants primarily work with by phase of participation	11
2	Population participants primarily work with by phase of participation	12
3	Primary involvement of participants in each phase	14
4	Primary role of participants in each phase	14
5	List of fifteen clusters and statements that comprise each one	27

## *Chapter 1*

### **Introduction**

Science and technology are viewed as the foundations of the modern economy (Atkin & Black, 2003), but America has long-witnessed differential performance and participation in science fields among its populace. Racial and ethnic minorities have historically performed and participated at rates below non-minority counterparts throughout primary, secondary, and post-secondary schooling (Clark, 1999). This phenomenon is of particular interest regarding black and Hispanic students. Demographic trends since the 1970's indicate a growing population of minorities with predictions that Latinos will comprise 29% of the U.S. population by 2050, and Whites will become the minority, comprising 47% of the population (Passel, Cohen, & Pew Hispanic Center, 2008). Given that the largest growing subgroups of the U.S. have shown neither the interest nor the affinity for science fields desired by politicians, government policies have been enacted in order to target these subgroups through science-based education programs and maintain American competitiveness in the global economy (Atkin & Black, 2003). As a result, the federal government has invested over \$3 billion to increase minority participation (Committee on Prospering in the Global Economy of the 21st Century & Committee on Science, Engineering, and Public Policy, 2007). Increases in STEM initiatives necessitate corresponding measures of program success beyond financial accountability, and this can be achieved through program evaluation of outcomes and implementation. Evaluating for expected outcomes and process implementation is necessary not only for program accountability but for determining and duplicating successful practices and program models that result in success. Because STEM initiatives specifically attempt to increase minority participation in STEM fields, these programs unfold in contexts that require culturally responsive practices on behalf of both evaluators and

program staff.

Modern approaches to culturally responsive evaluation (CRE) practice represent much needed progress toward conceptualizing a philosophy of CRE and toward articulating frameworks for how to think about CRE. However, the extent to which "practices" are clearly identified and operationalized is, as of yet, insufficient for training evaluators and program staff in cultural responsiveness, and we provide evidence of such in this study. We sought to survey evaluators and program staff working in science-based initiatives around the country in order to 1) determine culturally responsive practices implemented by evaluators and program staff as stakeholders in the field, 2) assess differences in perceived *importance* and perceived *feasibility* of practices, and 3) offer suggestions to evaluators and program staff for how to enact culturally responsive practices in their professions based on analysis of statements produced by professional peers.

## *Chapter 2*

### **Background**

The last decade has seen a call for attention to context in program evaluation . The emphasis on *context* has taken many forms, but most relevant to this discussion is the focus on the *cultural context* of program participants (Hood, Hopson, & Frierson, 2005). Discussions range from discourse on *cross-cultural* evaluations (i.e. working with cultural groups outside the U.S.) (Conner, 1989; Hopson, 2003) to *cultural competence* (American Evaluation Association, 2011) to evaluative frameworks and good practices that are inclusive of diversity and context without a focus on culture, e.g. responsive evaluation (Stake, 1976), participatory evaluation (Cousins & Whitman, 1998), inclusive evaluation (Ryan, Greene, Lincoln, Matheson, & Mertens, 1998), democratic evaluation (House & Howe, 1999). However, most recently, the

educational evaluation and assessment literature has revitalized the discourse on *culturally responsive evaluation (CRE)*.

In support of the growing emphasis on cultural context, the American Evaluation Association, the premier organization for American evaluators, recently released a *Public Statement on Cultural Competence in Evaluation* (2011). In this document cultural competence is described as "a stance taken toward culture" (p.1). A few paragraphs describe the cultural competence of an evaluator as "a process of learning" (p.2) as opposed to arrival at an end point, a reflection on one's own cultural location and of those with whom an evaluator works, and the ability to access theories and methodologies to match the cultural context of a program. However, the statement does not describe actionable steps. In fact, some evaluators argue that being culturally *competent* in no way implies *responsiveness* (Juarez & Greene, 2011). On the other hand, CRE frameworks urge professionals to move beyond a passive understanding or recognition of cultural characteristics relevant to an evaluation context but instead to actively *respond* to the needs of a cultural group.

### **The movement toward culturally responsive evaluation**

At present culturally responsive evaluation (CRE), not specific to education programs or STEM-related initiatives, is discussed in two ways: either as a globalized "sensitivity" when evaluating programs that require some level of attentiveness to culture or as a culture specific, ethnographic account of an evaluation team (or evaluator) working with a particular program and a particular cultural group. In this review we focus on extracting important generalizable practices and principles which we categorized into themes. This provided a framework in which to begin operationalizing CRE behaviors in more detail and with specific application to STEM program planning and evaluating. These principles lack specificity but form a framework that is

further informed by the results of our current study. A few themes or principles are recurrent in the culturally responsive evaluation literature: a) build trust and communication, b) understand the accuracies and inaccuracies of the cultural group context, c) employ a multifaceted approach to data collection/choose appropriate methods, and d) stakeholder self-determination.

***Building trust and facilitating communication with a multicultural staff.*** Two popular CRE notions are the need to build trust and the ability to communicate with stakeholders through the inclusion of multicultural staff (Botcheva, Shih, & Huffman, 2009). Having a multicultural staff, and specifically, staff members that share experiences with the population in question, is believed to facilitate an increased understanding of cultural characteristics within the team and represent this understanding in planning and implementation. Having multicultural staff serves a similar purpose to engaging with a *cultural informant* (Slaughter, 1991), the use of which is also suggested to encourage communication and cultural understanding. When it is not feasible to recruit staff with the appropriate cultural experience cultural competency training is suggested (Fitzpatrick, 2007).

***Understand the accuracies/inaccuracies of cultural context.*** Hilliard (1989) makes another important contribution to our conceptualization of CRE as evaluators and program staff working in a culturally diverse context. This author suggests that researchers, evaluators, and program staff in an evaluation context make active attempts to understand the historical context and possible misinformation relevant to a cultural group of interest- that the information they hold to be true about a cultural group may be the result of miseducation, and thus professionals should be willing to reexamine what they believe about a cultural group. For example, minority students, due to historical underperformance in the math and sciences, are often believed to be incapable of learning content in these subjects (Noguera, 2001). This is one instance in which

Hilliard might argue that teachers and others fail to “unlearn” historical trends that have been misinterpreted as innate inabilities of ethnic minorities. From another perspective, Karen Kirkhart (2011) argues that evaluators should reflect on their own cultural position and, in effect, challenge what they believe to be true about their cultural position and how we relate to others as a result.

***Engage a multifaceted approach to data collection and choose appropriate methods.***

The strengths of using multiple methods in evaluation planning are clearly stated in the literature (Bledsoe, 2005; Bledsoe & Graham, 2005). For example, Botcheva, et. al. (2009) gives a detailed account about how researchers used a two-step survey process with variable success. However, quantitative survey data was augmented with content analysis of participant poems. The poem content was especially useful and validating to program participants because the immediate culture placed a great deal of value on creating poetry. Additionally, Hilliard (1989) argues that multifaceted approaches are facilitated by maintaining a multidisciplinary perspective. He emphasizes the multidimensional qualities of cultural context which cannot be properly addressed in any one theoretical domain or by the perspective of one expert in one field. Thus efforts to contextualize a participant culture, evaluation design, data collection, etc. should be the result of not only a psychologist, for example, but of sociologists and anthropologists working together to add dimension to our conceptualization of culture in general and the specific culture in question.

Another important task at the evaluation design planning stage is choosing culturally appropriate measures (Bledsoe, 2005; Botcheva, et. al., 2009; Fitzpatrick, 2007.) For instance, one study (Botcheva, et. al., 2009) described how an initial survey assessment resulted in useless data. The researchers in this study discovered that all the survey questions were centered from an

individualistic point of view whereas the predominant cultural perspective was to incorporate a group perspective or to contextualize oneself in a group. Thus the questions were rewritten to accommodate this perspective with the aid of cultural informants.

***Self-determination and stakeholder involvement in program theory development.*** One almost universally accepted quality of developing a culturally responsive evaluation is the inclusion of relevant stakeholders in program development. Some evaluators (Thurston, Graham, & Hatfield, 2003; Edwards, Seaman, Drews, & Edwards, 1995) suggest that self-determination is the anticipated result of involving stakeholders in program theory development. In this way stakeholders contribute and articulate for themselves what the goals and expected achievements of a program should be and explicitly include those goals in a logic model that is reflective of community values. In Fitzpatrick (2007), the author identifies actual behaviors for engaging with stakeholders during program theory development. For instance, the author suggests that evaluators ask questions to identify individuals invested in program success and should engage in trust-building by attending program sessions and holding informal meetings with invested individuals. Most evaluators favor a multi-faceted collaboration between all or some combination of participants, program managers, evaluators, funders, community informants, and others as program context dictates.

***STEM program evaluation and CRE.*** Though there is a dearth of literature on the culturally responsive evaluation of STEM programs, Johnson (2005) not only recognizes that STEM programs are often designed to include minority students and should require CRE practices but also attempts to describe how and where in the evaluation process to introduce such practices. Johnson begins with a summary of responses provided by culturally responsive evaluators of education programs that were asked how they address context in evaluation design.

The evaluators described a) issues concerning the relevance of the evaluation to stakeholders and the extent to which program staff value input from the cultural group in question, b) collaborative evaluation design and commonality of cultural characteristics among participants, staff, and evaluators, and c) addressing population culture in instrumentation. When asked about cultural sensitivity in evaluation data analysis, the same group of evaluators discussed a) translation and validity of instruments used, b) disaggregation of data with respect to various contextual variables (e.g. language proficiency, race/ethnicity, SES, etc.), and c) fair and balanced report writing that is consumable by multiple stakeholders.

In Johnson's own synthesis of the literature she identifies five contextual factors that affect an evaluation: project setting, participant characteristics, stakeholder involvement, instrument selection, and findings relevant to stakeholder groups. Based on these factors and the responses provided by workshop participants, Johnson (2005) proposes a framework for addressing cultural context in STEM evaluations. In many instances her suggestions are actionable behaviors but mostly represent vague principles common to the CRE effort in general. However, this work is unique in its effort to engage evaluators in a focus group-type discussion of CRE practices specifically relevant to STEM and in that it attempts to formally interject a CRE framework at every step of the evaluation process.

There are some issues with information about CRE practices in the literature. For instance, though many multicultural evaluators practice professionally, no existing examples of evaluations in which multicultural staff were actively recruited could be found. There is no discourse about best practices used to find and recruit qualified multicultural staff, least of all studies or observations about how an evaluation in which multicultural staff were recruited compares to an evaluation in which no active recruitment was used. Additionally, though

Kirkhart (2011) can speak elegantly about the value of reflective and reflexive approaches to evaluation in culture-specific contexts and provide examples of evaluations conducted from such a perspective, a clearly articulated method is lacking. Regarding stakeholder input, there is variability in how stakeholders are brought into program or evaluation planning and theory development. The majority of articles addressed stakeholder involvement to some extent, however, did not provide guidance on how to do so. Generally, discussion in the literature is inconsistent regarding the point in the evaluation and program planning process at which stakeholder input is or should be introduced. Thus there is no prescriptive guidance on how to decide the intensity or timing of stakeholder involvement in theory development across program domains, and discussions about how to illicit stakeholder input vary from project to project.

The most glaring concern with the current state of culturally responsive evaluation and program practice is that little evidence exists regarding which practices, when clearly identified, or which principles actually improve the quality of an evaluation. The literature is replete with context-specific examples of CRE implementation. However, the extent to which practices affected results or implementation of an evaluation is discussed in qualitative, ethnographic accounts. Even when accounts are convincing about the utility of CRE practices, the extent to which they can be extracted and applied to other program contexts is unclear. Thus practitioners wanting to learn about and adapt culturally responsive practices are left unsure of how to accomplish the task. Principles extracted from the literature provide an initial framework for assessing how CRE and culturally responsive program planning are conceptualized by practitioners- both evaluators and program staff; and Johnson's (2005) work is a positive example of how to begin to translate important concepts into behaviors. However, more specific direction is needed in order to implement principles/approaches. In the work discussed herein,

researchers acknowledge the importance of stakeholder input in the evaluation process and employed a research methodology that would validate the concern for stakeholder input while beginning a systematic inquiry into what constitutes culturally responsive practice. Concept mapping is a methodology which allows for multiple perspectives to be incorporated in response to a research question. For this reason, it was used in as the data collection methodology in this current study.

### **Using stakeholder input through concept mapping to investigate CRE practices**

Concept mapping has been used in program planning, evaluation, and in social science research for over 20 years (Kane & Trochim, 2007). This methodology is attractive and well-suited for the goals of this current study: a) to identify practices and principles of culturally responsive evaluation in STEM programs as articulated by persons that work in these arenas consistently and b) to investigate how professionals in the field perceive these practices and principles in terms of *importance* and *feasibility*. As indicated in the evaluation literature stakeholder input is a valuable part of evaluation practice. Once stakeholders are identified they can assist in framing evaluation questions, in conceptualizing program goals, and in redirecting iterative program/evaluation efforts. In a similar fashion, stakeholders in this research helped us to conceptualize cultural responsive evaluation practice and then provide us with feedback on the importance and feasibility of CRE. In our current study, stakeholders were identified as evaluators and program staff that work with STEM or other science-based initiatives. Participants in these programs and community members are other stakeholders that were identified but not included in this initial study due to limitations in resources and limited access to these populations.

Using stakeholders as the source of information is a unique approach to operationalizing

CRE concepts and to beginning an investigation of the utility of various CRE practices in future research. The concept mapping methodology described below allowed researchers in this current study to systematically access stakeholder input about the field of CRE and perform informative analysis about the state of CRE as perceived and practiced by professionals. The concept mapping methodology is described in detail below.

### *Chapter 3*

#### **Methods**

Concept Mapping is a structured conceptualization method (Trochim, 1989) which we employed to gather and organize information for the topic of interest in this project. *Concept Mapping* consists of three phases of participant involvement: *brainstorming* of statements in response to a prompt, *sorting* of statements that were gathered during the brainstorming phase, and *rating* of the statements on dimensions of interest. Trochim (1989) describes six steps in the implementation of all phases of *Concept Mapping*: a) preparation, b) generation of statements, c) structuring statements, d) representation of statements, e) interpretation of maps, and f) utilization of maps. The six steps are explicated below.

#### **Preparation**

*Sample recruitment.* Participants consisted of program evaluators identified through membership in the *American Evaluation Association* and program staff working with 4-H programs in Texas, New York, and California identified through the National 4-H public website. Individuals working either with diverse populations (ethnically diverse, differently-abled populations, sexual minorities, or other diverse populations as identified by participants) and/or with science-related initiatives were targeted to participate. However, as this study was conducted through email invitations and online, it was possible for unintended audiences to

access and complete any of the three phases of the project. Table 1 and Table 2 below describe the extent to which targeted sample populations were reached.

Table 1

*Type of programs participants primarily work with by phase of participation*

Phase	4-H	STEM	Other science-related initiative	None of the above	TOTAL
Brainstorming	2 (5.6%)	9 (25.0%)	7 (19.4%)	18 (50.0%)	36 (100.0%)
Sorting	0 (0.0%)	9 (47.4%)	1 (5.3%)	9 (47.4%)	19 (100.0%)
Rating <i>Feasibility</i>	0 (0.0%)	9 (42.9%)	2 (9.5%)	10 (47.6%)	21 (100.0%)
Rating <i>Importance</i>	1 (4.2%)	10 (41.7%)	2 (8.3%)	11 (45.8%)	24 (100.0%)

As indicated in Table 1, at least 50% of participants in each stage of the project identified as working with either 4-H programs, STEM programs, or some other science-related initiative as their primary focus. It is important to note that even though nearly 50% of participants in each phase stated that they did not work "primarily" with science initiatives, it is possible that they did work with them in some capacity. However, this data was not collected. Table 2 below indicates that at least 50% of participants at each phase worked primarily with a specifically diverse population. Again, even though nearly 50% of participants in each phase stated that they did not work "primarily" with a specifically diverse population, it is possible that they did work with them in some capacity and could therefore provide informative responses about their experiences with culturally responsive practices.

Table 2

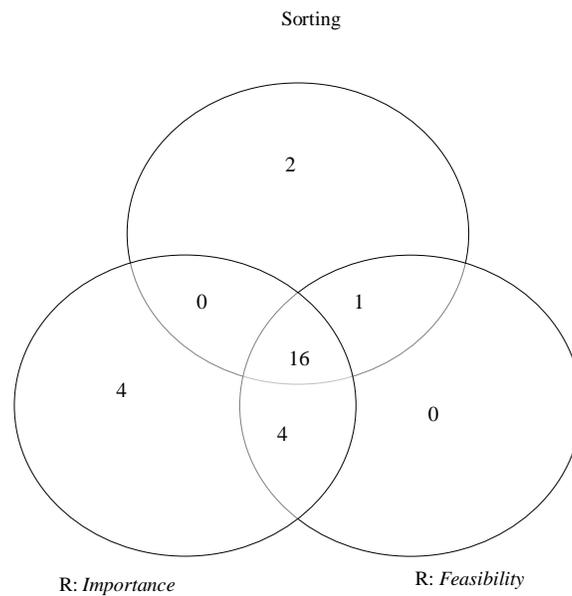
*Population participants primarily work with by phase of participation*

Phase	Ethnic/ racial minorities	Disabled persons	Sexual preference minorities	Low-income communities	Religious minorities	Women	No specific population	Other	TOTAL
	Brainstorming	10 (27.0%)	1 (2.7%)	0 (0.0%)	5 (13.5%)	0 (0.0%)	3 (8.1%)	15 (40.5%)	3 (8.1%)
Sorting	5 (26.3%)	0 (0.0%)	0 (0.0%)	2 (10.5%)	0 (0.0%)	2 (10.5%)	9 (47.4%)	1 (5.3%)	19 (100.0%)
Rating									
Feasibility	5 (23.8%)	0 (0.0%)	0 (0.0%)	2 (9.5%)	0 (0.0%)	2 (9.5%)	9 (42.9%)	3 (14.3%)	21 (100.0%)
Rating									
Importance	6 (25.0%)	0 (0.0%)	0 (0.0%)	1 (4.2%)	0 (0.0%)	2 (8.3%)	12 (50.0%)	3 (12.5%)	24 (100.0%)

Also, because phases of the project were open for participation at different times, participants could complete any combination of portions at once or return and complete individual sections at a different times. The *brainstorming* phase was completed several months before the *sorting* and *rating* phases and was the only section that was necessarily completed separately from the others. However, the *sorting* and *rating* phases were open at the same time, and each of these sections could be completed in one sitting if the participant chose to do so. Alternatively, a participant could complete one section and never return to complete another section or could complete one section and then return to complete another section at a later time. For those participants that participated anonymously it is not possible to identify those who completed one phase and returned to complete another phase later. Figure 1 below shows participation in each section and when participants completed multiple sections at once.

Figure 1

*Number of participants that completed each or multiple sections at once*



Another concern regarding the use of an open email invitation for project participation was that the email would reach individuals that were not somehow involved in evaluation activities (i.e. program staff with no participation in evaluations related to their program.) Table 3 below describes the extent to which participants were involved in evaluation activities of some form.

Table 3

*Primary involvement of participants in each phase*

Phase	Program planning/and or implementation	Planning and/or implementing an evaluation	Planning and/or implementation of both	TOTAL
Brainstorming	7 (17.5%)	12 (30.0%)	21 (52.5%)	40 (100.0%)
Sorting	1 (5.3%)	12 (63.2%)	6 (31.6%)	19 (100.0%)
Rating <i>Feasibility</i>	1 (4.8%)	13 (61.9%)	7 (33.3%)	21 (100.0%)
Rating <i>Importance</i>	1 (4.2%)	16 (66.7%)	7 (29.2%)	24 (100.0%)

Though some participants in each phase indicated that they were involved only in program activities and *not* evaluation activities, the majority of participants were experienced in evaluation planning and/or implementation. Table 4 describes participants' primary roles as either evaluators or as program staff, and indicates that nearly two-thirds of participants in each phase identified as "trained evaluators."

Table 4

*Primary role of participants in each phase*

Phase	Trained Evaluator	Program staff leading an evaluation	Program staff assisting with an evaluation	Program staff not working with an evaluation	Other	TOTAL
Brainstorming	14 (34.1%)	8 (19.5%)	4 (9.8%)	8 (19.5%)	7 (17.1%)	41 (100.0%)
Sorting	11 (57.9%)	1 (5.3%)	4 (21.1%)	0 (0.0%)	43(15.8%)	19 (100.0%)
Rating <i>Feasibility</i>	14 (66.7%)	1 (4.8%)	4 (19.1%)	0 (0.0%)	2 (9.5%)	21 (100.0%)
Rating <i>Importance</i>	17 (70.8%)	1 (4.2%)	4 (16.7%)	0 (0.0%)	2 (8.3%)	24 (100.0%)

## Generation of statements

***Focus Prompt.*** Initial preparation of the concept mapping exercise included generation of the focus prompt. The focus prompt is a partial statement intended to capture brainstorming ideas relevant to the focus of the concept mapping study. For this project, three versions of a focus prompt were generated by researchers and pilot tested with eight evaluators and graduate students studying evaluation to determine the most appropriate wording for the statement. The goal of the prompt was to elicit responses related to activities and behaviors more so than concepts or ideas about CRE. After reviewing the type of data that each prompt elicited, the following prompt was chosen for use in the study:

*" One specific thing I do to be culturally responsive in planning, implementation, and/or evaluation of social and education programs is...."*

This prompt, though it does not specifically reference STEM or science-based programs, resulted in more statements related to actions than other prompts. Additionally, regardless of the prompt the pilot sample was shown, statements tended to be general CRE concepts not related to STEM. That being the case, researchers proceeded to use the prompt that elicited practice or behavior relevant content.

***Participant responses.*** The final draft of the focus statement was entered into the *Concept System Global* website, an internet based software that is specifically designed for use in concept mapping projects. Participants could access the statement prompt during a two-week open period from any internet capable computer. Individuals were invited to partake in the study via an email invitation containing a brief description of the study, and those that wanted more information or that agreed to participate followed a hyperlink embedded in the original email

invitation that directed them to the project webpage. Once at the project website, individuals could read a description of the project, read the consent form for the project, and choose to participate by continuing on to the next screen or could leave the site without participating. Participants were not required to formally consent as this project was exempted by the Cornell University IRB, but they were presented with a consent form to provide all project-relevant information.

There was no time limit to complete the brainstorming exercise and participants could revisit the site at multiple times if they chose. This study was different from a standard concept methodology (Kane & Trochim, 2007) in that statements were completed individually and anonymously as opposed to in a group. Participants were directed to list as many statements as they felt addressed the focus prompt and to list statements as separate, distinct ideas whenever possible. However, participants did not always successfully separate ideas (i.e. they used compound sentences, wrote in paragraph form as opposed to list form, etc.). Thus a process for separating ideas into distinct statements is described.

*Statement synthesis.* The process for synthesizing the statement set consisted of three steps. In Step 1 ninety-three original statements were adjusted to be grammatically correct as a completion of the focus prompt. In Step 2 compounded statements were identified and separated into distinct statements/ideas by using standard rules of grammar. For example, if more than one sentence was used to complete the focus prompt, each sentence was separated into its own distinct idea. If the individualized statements contained conjunctive terms such as “and” they were flagged as possible compound statements that may need to be further deconstructed. This resulted in a new set of 122 statements. However, in the final step the statement set was reduced by removing repetitive or unclear statements. For example, the response below was prepared for

analysis in the following way:

*...ask members of the population we serve to participate in planning the program.*

*This might include serving on an advisory committee, responding to a proposed activity, participating in a focus group related to the program.*

Step 1:

Is the response represented in paragraph form? If, yes then list each statement separately.

Statement 1:

*...ask members of the population we serve to participate in planning the program.*

Statement 2:

*This might include serving on an advisory committee, responding to a proposed activity, participating in a focus group related to the program.*

Step 2:

Does either statement represent compound ideas that should be separated? Yes, statement 2 describes three activities that can be separated into three statements.

Statement 2 becomes three statements:

*This might include serving on an advisory committee.*

*...responding to a proposed activity.*

*...participating in a focus group related to the program.*

Step 3:

Is each statement grammatically correct and does it complete the focus prompt?

Statement 1 does complete the thought, but each distinct idea in statement 2 needs to be adjusted to finish the thought. The new list of statements from the one original response becomes:

*...ask members of the population we serve to participate in planning the program.*

*...ask members of the population we serve to serve on an advisory committee.*

*...ask members of the population we serve to respond to a proposed activity.*

*...ask members of the population we serve to participate in a focus group related to the program.*

In this way, we preserved as much of the original text as possible while making statements useful for analysis and the next phase of the study. Note that in this example the original set of two statements became four statements.

After deconstructing compounded responses into separate coherent ideas, our original list of 93 statements grew to 122 statements. Research has shown (Kane & Trochim, 2007) that 100 statements is a reasonable set for participants to work with during the next anticipated phase of the research. Thus in the subsequent step we identified all statements with similar meanings and kept only one of them. We also eliminated any incomplete statements or statements that were irrelevant, nonsensical, etc. For example, the two statements below were very similar. Thus only one statement was kept.

*... ensure language in instrument development is culturally sensitive to stakeholders.*

*... review measures for culturally appropriate and accessible language*

*(questions, introductory blurb, etc.)*

The statement below was not relevant to the prompt and was removed.

*We tried very hard to adapt our program in order to ensure success.*

After the statement synthesis steps were completed, researchers were left with a final set of 97 statements to be used in subsequent phases. Appendix A contains the set of statements at each step in the statement synthesis process.

## **Structuring and rating the statements**

*Sorting.* The next data collection phase consisted of statement sorting. The new set of statements that resulted from statement synthesis was reentered into the Concept Systems Global website. Through an email invitation process similar to that described above, evaluators and program staff were invited to participate in this next phase of the project. Email invitations for participation in this phase of the project were sent to the same individuals and list serves as in the brainstorming phase of the project. It is possible that individuals participated in both phases after receiving email invitations. However, this data was not gathered as participation in both phases was anonymous. In the *sorting* phase, participants were shown the 97 brainstormed statements and were asked to create piles that group similar statements together. What makes statements "similar" is subjective, and participants are directed to use whatever underlying themes or concepts they choose. Participants then label each sort with respect to the theme or concept that they feel relates each statement to others in a given pile.

During this phase, researchers experienced challenges in participant recruitment and used multiple avenues for attaining data. The majority of participants partook in project activities by visiting the project website and following the instructions in the web program. For evaluators and program staff in the New York area where researchers were located, traditional concept mapping techniques were used. This consisted of printing the statements onto sort cards and organizing group events in which two or more individuals gathered and sorted the statements at the same time. Though discussion among participants engaging in this process is reasonable and often occurs, very little group discussion ensued in the one meeting that took place and is not expected to have an effect on data collected in this fashion. A few evaluators used the sort cards provided to complete the task individually instead of in a group. This process is not expected to

have an effect on the data collected, either. As recently found in a study by Rosas & Kane (in press), variation in data collection methods across 69 concept mapping studies did not affect the validity and reliability of study results. Individuals who had opted to use sort cards retained them for the rating phase described below.

***Rating statements.*** In a separate step, participants were also asked to provide two ratings for each statement- one rating based on how important they perceived a statement to be for their work as a culturally responsive evaluator or program staff member and one based on how feasible it is to implement each statement in their work. Ratings were conducted on a Likert-type scale where the number "0" represented *not important at all/not feasible at all* and the number "4" represented *very important/very feasible*. Participants who used the website to complete the study could choose to complete only the sort task, only one of the ratings, or any combination of the three tasks at once or return at different times to complete them. Participants who used printed sort cards to complete the study assigned each rating to a statement by writing an *F* for *Feasibility* or an *I* for *Importance* followed by their rating on each of the cards. Sort cards were returned to researchers for data entry.

#### *Chapter 4*

#### **Analysis**

Analyses were conducted on data collected during the sorting and rating phases of the current project as described in Kane & Trochim (2007). A two-dimensional multidimensional scaling (MDS) was conducted on the sort data which resulted in a set of coordinates for each statement, and a hierarchical cluster analysis was conducted using the resulting coordinates. Two sets of ratings were collected during the rating phase and were used to produce a *pattern match*

as described below. Additionally, a hierarchical linear model was constructed to assess whether trends identified between rating dimensions were statistically significant.

### Representation of statements

The purpose of a traditional concept mapping analysis is to represent the statements in terms of clusters of similar content and in terms of alignment with respect to ratings; and these analyses are conducted in *The Concept System* desktop software. The analysis conducted in *The Concept System* provides a structured, statistically-based organization of concepts represented in brainstormed statements from the perspective of relevant stakeholders. These analyses result in a *point map*, a *cluster map*, and a *pattern match*. For the purposes of this study representation of the statements in this way allowed us to compare themes with those in the literature, to identify which activities or statements are feasible, important, or both, and to identify general trends in the data for further analysis.

**Maps.** In a concept mapping methodology, the conceptual maps created are based on the statement sort data collected. A binary similarity matrix is created in which statements are grouped based on how often they were sorted together by participants. If two statements are sorted together by a participant a "1" is placed in the row by column intersection of the two statements. If they are not sorted together, a "0" is placed at the intersection. This can best be illustrated by Figure 2 below.

Figure 2

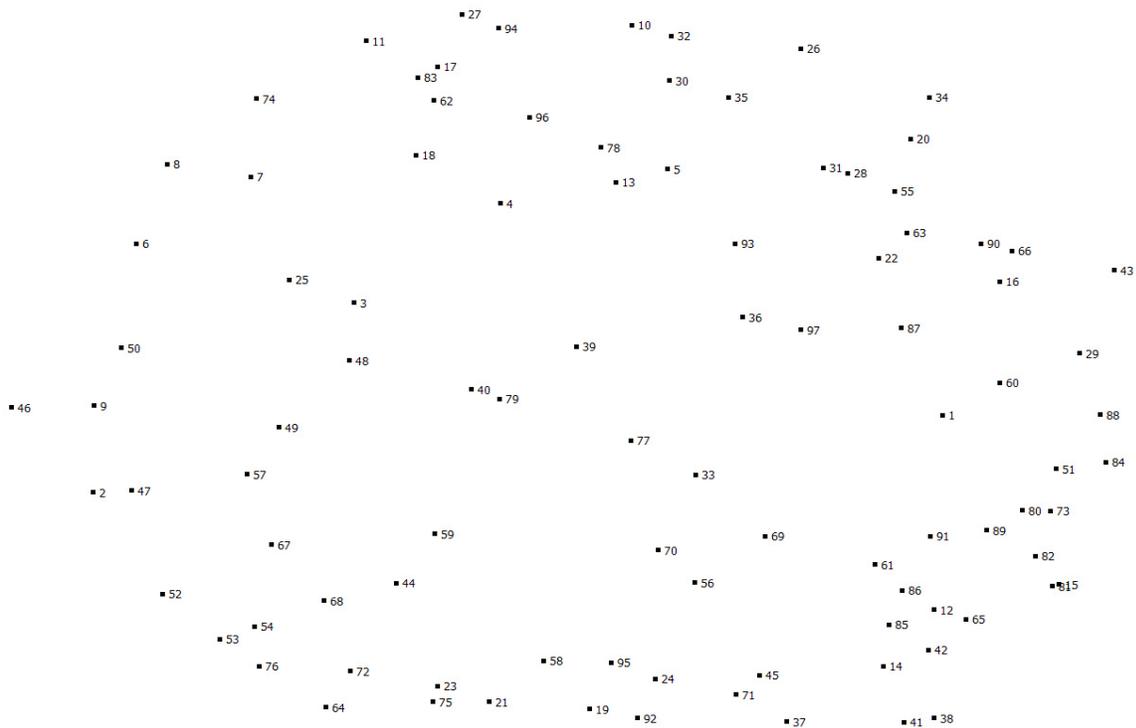
*Example of a similarity matrix for one participant*

	S1	S2	S3
S1	0	1	0
S2	1	0	0
S3	0	0	0

In this figure, statements S1 and S2 were sorted together by this participant. The matrices for all participants were aggregated in a multidimensional scaling (MDS) analysis. The two-dimensional solution that resulted provided a set of (x, y) coordinates for each statement. A plot of these coordinates is called a *point rating map*. Figure 3 below is a point rating map for the 97 statements in our study.

Figure 3

*Point rating map of 97 culturally responsive statements*



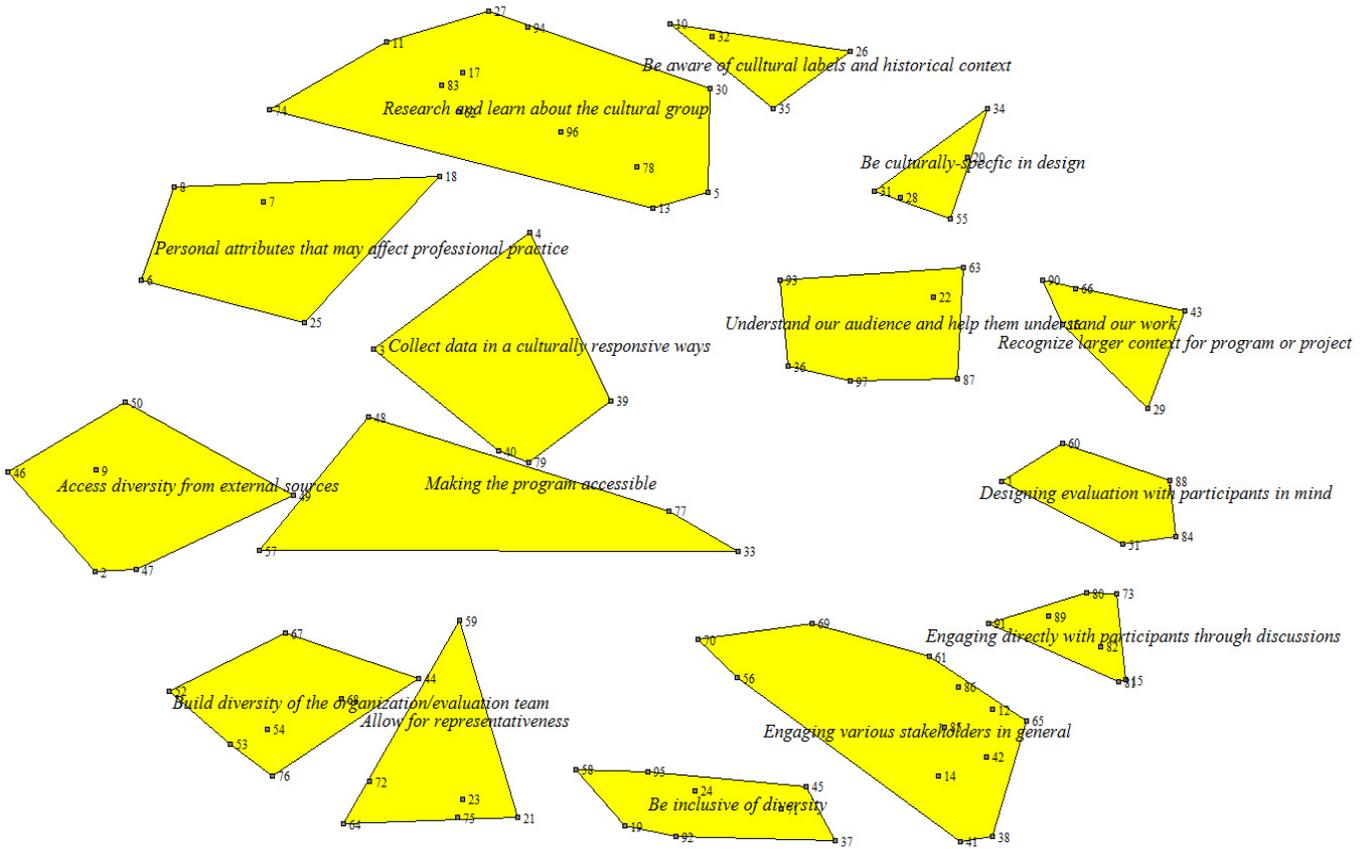
Data in this format reveals some potential clusters. It also allows us to see where a statement falls on the map in relation to other statements of potential interest. For our purposes in this current study, the point rating map is a necessary step for producing a cluster map. A hierarchical cluster analysis was performed on the MDS coordinates to produce clusters of

statements. This provides a visual, statistically determined map of which statements were sorted together most often by participants.

This current study resulted in the cluster map below. In *The Concept System* researchers decide the number of clusters they believe will accurately represent the data. In this study, analysis began with 25 clusters, and we combined one cluster at a time until the analysis arrived at a set of cohesive clusters that represented underlying themes of the statements within. This occurred when the cluster map contained 17 clusters. Researchers chose to begin with 25 clusters because it represented a reasonable number of possible activities or behaviors related to CRE and finished with 17 clusters because underlying themes of the 17 clusters were consistent with the CRE literature. However, careful analysis of the statements in each cluster revealed that two of the seventeen clusters contained statements that fit better, conceptually, into surrounding clusters. In this instance, researchers manually redrew cluster boundaries so that the two clusters were absorbed into neighboring clusters. Allowing the software to remove the two clusters automatically would have resulted in placement of the statements into clusters that did not make sense. It is possible that there were not enough sorts for the analysis to detect the extent to which these statements belonged in neighboring clusters. Additionally, three points on the map were moved slightly to prevent cluster overlap and allow for clarity in the visual appearance of the cluster map.

Figure 4

*Fifteen cluster map of culturally responsive statements*



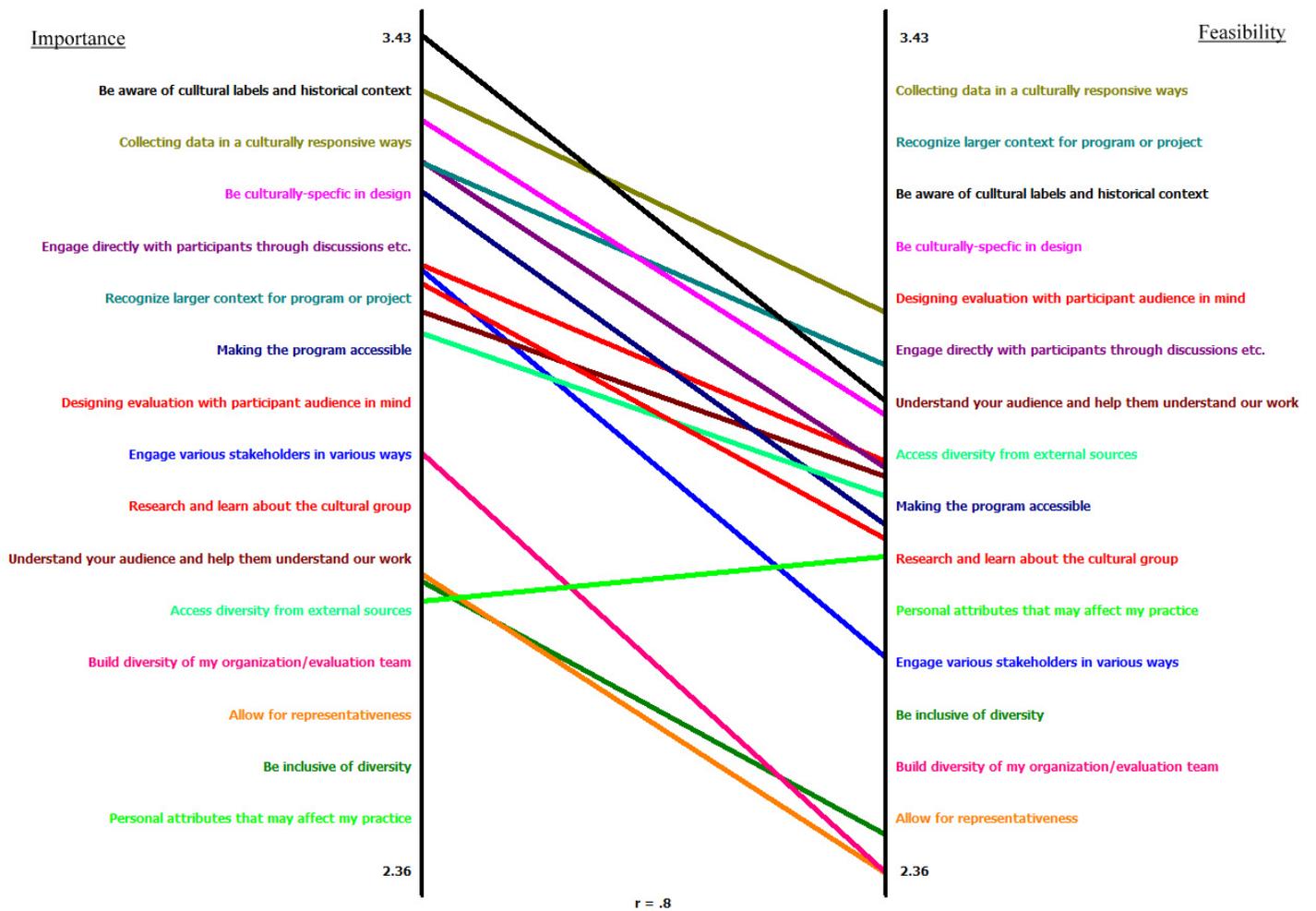
The cluster map analysis also produces suggestions for how to title each cluster based on titles assigned by participants during the sort task. In this project, researchers used their knowledge of the field of culturally responsive evaluation in conjunction with suggestions provided during the analysis to create cluster titles. A list of each cluster and examples statements within each cluster can be found in Table 5 where an interpretation of the map is discussed below. The complete list of statements for each cluster is in Appendix B.

**Pattern match.** In addition to creating maps, *The Concept System* produces pattern matches. Pattern matches are visual depictions of trends in the average ratings of clusters on a given factor. For instance, the figure below is a pattern match of the average *importance* ratings

for each cluster compared to the average *feasibility* ratings for each cluster across participants. It is evident from this pattern match that *importance* ratings trend higher on average than *feasibility* ratings. This appears true for 14 out of the 15 clusters. This pattern match revealed an important trend that we investigated further with a test for statistical difference.

Figure 5

*Pattern match of Importance and Feasibility ratings*



**Analysis for statistical significance.** The pattern match analysis revealed a persistent trend in which average ratings for *importance* appeared to be higher than average ratings for *feasibility* for most of the clusters. In order to test this, we constructed a mixed model in which

*feasibility* ratings were regressed on *importance* ratings for individual statements and for cluster averages. *Feasibility* ratings were significantly lower than *importance* ratings across clusters,  $F(14, 1716) = 1.63, p = .066$ , at the  $p < 0.10$  level.

## Chapter 5

### Results

#### Ratings

Findings revealed a statistical difference between dimensions of ratings and suggest that even though practitioners appreciate the value or importance of certain activities, they may not be aware of ways to implement these ideas or how to make CRE concepts actionable. Whereas the linear model reveals this difference, the pattern match allows us to see which cluster(s) may be driving this finding. We can see that Cluster 13 *Personal attributes that may affect my practice* is the only cluster in which participants perceive the statements/behaviors to be relatively equivalent in average *importance* and *feasibility*. All other clusters trend in a direction that indicates high *importance* and relatively lower *feasibility*.

#### Interpretation and utilization of the cluster map

In a concept mapping project, cluster maps can be used for any number of purposes, such as to frame evaluation questions or to determine program goals (Trochim, 1989). In this current work we interpret the cluster map as a hierarchy of global principles (represented by cluster titles) that are informed by more narrowly defined principles (statements) which may also be informed by examples of practices or behaviors (statements). However, given gaps in the way the field of CRE is conceptualized and put into practice our resulting taxonomy is incomplete. In some cases global principles are informed by more detailed principle statements but no examples of behaviors are provided. Table 5 provides an example of statements for each cluster.

Table 5

List of fifteen clusters and statements that comprise each one

Cluster name	Number of statements	Example of statements within each cluster
1. Designing evaluation with participants in mind	5	<i>"engage our clients in making explicit their values, assumptions, and philosophy."</i>
2. Engaging directly with participants through discussion	7	<i>"hold discussion with the group to get their insights on what is not working and how it might work better."</i>
3. Engaging stakeholders in general planning and theory development	12	<i>"ensure that a wide range of stakeholders receive and use results of the evaluation."</i>
4. Recognize larger context for programs or projects	5	<i>"to use term or concepts in the evaluation that related to their life experiences."</i>
5. Be culturally-specific in design	5	<i>"consider program outcomes that may differ from majority outcomes, e.g. outcomes that are culturally-specific."</i>
6. Understand our audience and help them understand our work	6	<i>"ensure that our evaluation is framed and focused on benefit to the community."</i>
7. Be inclusive of diversity	8	<i>"to plan activities that bring people from different backgrounds into constructive dialogue with each other."</i>
8. Making the program/evaluation accessible	4	<i>"to use information about barriers to participation to help me to make changes for future program activities."</i>
9. Allow for representativeness	6	<i>"to try to involve role models who are members of an underrepresented group (i.e. women, minorities, etc.) in professional development."</i>
10. Build diversity of the organization/evaluation team	7	<i>"find diverse board members as advisors."</i>
11. Access diversity from external sources	6	<i>"make use of the diversity of my co-workers."</i>
12. Collect data in culturally responsive ways	5	<i>"review measures for culturally appropriate and accessible language (questions, introductory blurb, etc.)"</i>
13. Understand personal attributes that may affect professional practice	5	<i>"spend time building and nurturing diverse friendships outside of work."</i>
14. Be aware of cultural labels and historical context	4	<i>"be aware of past injustice and harm that has resulted from evaluation or research with this population."</i>
15. Research and learn about the cultural group	12	<i>"to spend time with the population/organization in very informal ways in order to understand organizational norms, population behaviors, expectations, etc..."</i>

Themes that arise in the data gathered from stakeholders can be compared with themes and

principles salient in the CRE literature. Most importantly each of the themes or clusters from the cluster map is further explicated by the statements that comprise each cluster. In instances where themes from the literature resemble resulting clusters from the current study, statements also help to clarify what principles in the literature are suggesting. In this way the cluster map provides a systematic approach to stakeholder input on what CRE behaviors "look like" in everyday practice and help to translate existing theoretical literature into practice.

## *Chapter 6*

### **Discussion**

In our study several cluster titles resonate with themes in the literature. For example, Cluster 2 *Engaging directly with participants through discussion* and Cluster 3 *Engaging stakeholders in general* echo current discussions regarding stakeholder inclusion in program planning and evaluation. However, through this current work, we are able to provide nineteen statements within these two clusters combined that inform ways to address stakeholder input. For instance, clear suggestions include *"hold(ing) discussions with the group to get their insights on what is not working and how it might work better"* and *"establish agreements for their participation in reviewing the analysis report."* Cluster 7 *Be inclusive of diversity* is also similar to the idea of stakeholder inclusion but focuses on the diversity of stakeholders and suggests inclusion of diversity in general with statements such as *"plan activities that bring people from different backgrounds into constructive dialogue with each other."* Though a student of CRE may read about these ideas and conflate them, the current work provides properly nuanced categories to differentiate general stakeholder involvement to inform planning and evaluation from stakeholder input that specifically highlights diversity in the evaluation context.

Another cluster that resonates with the literature is Cluster 10 *Build diversity of my*

*organization* which is similar in notion to that of building trust through use of a multicultural staff. Additionally, statements within Cluster 12 *Collect data in culturally responsive ways* and Cluster 5 *Be culturally-specific in design* enumerate ways in which to address concerns about how to choose multifaceted and appropriate methods. Cluster 14 *Be aware of cultural labels and historical context* embodies discussions related to understanding accuracies and inaccuracies related to a culture. Cluster 14, however, is one example in which relevant statements remain vague and theoretical.

The specificity of each statement varies. Statements such as *"hold(ing) discussions with the group to get their insights on what is not working and how it might work better"* and *"establish agreements for their participation in reviewing the analysis report"* or *"plan activities that bring people from different backgrounds into constructive dialogue with each other"* are specific practices rather than principles. However, these statements could be expanded further by asking *How might one facilitate discussion with a group?* or *What should an agreement look like?* Or one might ask *What kind of activities would accomplish constructive dialogue?* Regardless, these statements are a necessary first step in narrowing discussions that have remained broad in scope and must be better defined in order to progress the field. Continuing to narrow the scope of practice by asking questions in this way could make practices less applicable across program domains and more culture specific, which would defy the aims of this current study. Establishing what level of specificity is sufficient for providing guidance while allowing practices to be generalizable across contexts is a matter warranting further discussion in the field.

### **Methodological concerns**

Concept mapping is employed when researchers choose to combine qualitative data with quantitative data and a strong statistical approach. Though there is some flexibility in the

qualitative component of this methodology, portions of this study would benefit from more controlled sampling efforts. For example, sample recruitment of professionals working primarily with STEM or science-based programs and/or cultural minorities was not achieved to the extent desired. The language of the study invitation and the project website did not strongly discourage individuals who did not meet these criteria from participating. Additionally, the nature of web-based participation resulted in problems recruiting individuals. Whereas concept mapping is usually conducted as a group to alleviate tedium and encourage discussion of concepts, conducting concept mapping tasks in solitude at a computer proved to be a difficult endeavor for many participants. There is anonymous data showing that many individuals began tasks but did not complete them or entered the site and left without participating. In the future, this concern can be addressed by targeting participants with an invitation that allows them to create a login and complete tasks at will in any time frame that suits them rather than forcing them to complete a given task in its entirety each time they visit the site.

### **Future directions**

Our study provided some statistical evidence for the disconnect between theoretical approaches to culturally responsive evaluation and the implementation of CRE practice. Future studies should investigate perceived barriers to implementing CRE and gather information from professionals about what would make CRE principles actionable. Another important step is to conduct research on CRE practices that have been successfully articulated to investigate their effects on evaluation implementation and evaluation results.

### **Conclusions**

Concepts in the literature and ideas expressed by participants in this project converge on various points. It is not possible to discern if literature informs practice or if practice informs

literature, and it may not be a relevant concern, but the existence of a reciprocal relationship would lead us to expect similarities. Our assumption was that professionals that put theory in to practice on a daily basis would not necessarily provide new insights on CRE principles, but would be able to describe specific CRE activities or behaviors that exemplify concepts in the literature. In some cases this occurred but in others, ideas suggested by participants remained at the level of a global principle. Inconsistencies in the literature regarding how to put principles into practice are echoed in statements provided by participants, and our regression analysis provided statistically significant evidence at the  $p < 0.10$  level that participants' perceive implementation of CRE practice as problematic. It is not clear from this data why this perception exists, but it is an important finding nonetheless.

Among CRE evaluators, there has been a resistance to engage the evaluation field with traditional evidence of best practices when advocating for cultural responsiveness. This is evident in the dearth of research-based literature on what CRE activities actually affect evaluation results. A resistance to conducting research on evaluation and to questioning what constitutes evidence regarding culturally responsive evaluation has resulted in a lack of clearly articulated methods as is evidenced in this study by both the difference in *importance* and *feasibility* ratings and by the content of statements collected. It may well be a worthy and valid endeavor to question whether traditional forms of evidence, such as randomized controlled trials, apply to efforts to determine CRE best practices, but some form of evidence is necessary. The current study represents an initial effort to conduct research on culturally responsive evaluation efforts by first establishing that evidence-based CRE practice is in need. Now that evidence for this need exists, culturally responsive evaluators are charged with providing actionable, evidenced-based guidance for the practice of culturally responsive evaluation.

## References

- American Evaluation Association. (2011). *American Evaluation Association Public Statement on Cultural Competence in Evaluation*. Fairhaven, MA: Author.
- Atkin, J.M. & Black, P.J. (2003). *Inside science education reform: a history of curricular and policy change*. New York: Teachers College Press.
- Bledsoe, K. (2005). Cultural Reflections stemming from the evaluation of an undergraduate research program. In S. Hood, R. Hopson, & H. Frierson (Ed.), *The role of culture and cultural context: A mandate for inclusion, the discovery of truth, and understanding evaluative theory and practice* (pp. 217-235). Greenwich, CT: Information Age Publishing.
- Bledsoe, K. L., & Graham, J. A. (2005). The Use of Multiple Evaluation Approaches in Program Evaluation. *American Journal of Evaluation*, 26(3), 302-319.
- Botcheva L., Shih J., & Huffman, L.C. (2009). Emphasizing cultural competence in evaluation: A process-oriented approach. *American Journal of Evaluation*, 30(2), 176-188.
- Clark, J. V., & ERIC Clearinghouse for Science, Mathematics, and Environmental Education. (1999). *Minorities in science and math*. [Columbus, Ohio]: ERIC Clearinghouse for Science, Mathematics and Environmental Education.
- Committee on Prospering in the Global Economy of the 21st Century and the Committee on Science, Engineering, and Public Policy. (2007). *Rising above the gathering storm : energizing and employing America for a brighter economic future*. Washington, D.C.: National Academies Press.
- Conner, R. F. (1989). International and Cross-Cultural Viewpoints: New Perspectives on Evaluation. *American Journal of Evaluation American Journal of Evaluation*, 10(1), 58-62.

- Cousins, A. J., & Whitmore, A. E. (1998). Framing Participatory Evaluation. *New Directions for Evaluation*, 1998(80), 5-24.
- Edwards, E. D., Seaman, J. R., Drews, J., & Edwards, M. E. (1995). A Community Approach for Native American Drug and Alcohol Prevention Programs: A Logic Model Framework. *Alcoholism Treatment Quarterly*, 13(2), 43.
- Fitzpatrick, J. L. (2007). Evaluation of the Fun With Books Program: An Interview With Katrina Bledsoe. *American Journal of Evaluation American Journal of Evaluation*, 28(4), 522-535.
- Hilliard, A. G. (1989). Kemetic (Egyptian) Historical Revision: Implications for Cross-Cultural Evaluation and Research in Education. *American Journal of Evaluation American Journal of Evaluation*, 10(2), 7-23.
- Hood, S., Hopson, R. K., & Frierson, H. T. (2005). *The role of culture and cultural context : a mandate for inclusion, the discovery of truth and understanding in evaluative theory and practice*. Greenwich, CT: IAP.
- Hopson, R. K. (2003). *Overview of Multicultural and Culturally Competent Program Evaluation Issues, Challenges and Opportunities*. Oakland, 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 (Ed.), *The SAGE international handbook of educational evaluation* (pp. 431-448). Thousand Oaks, CA: Sage.
- House, E. R., & Howe, K. R. (1999). *Values in evaluation and social research*. Thousand Oaks, Calif.: Sage.
- Johnson, E. (2005). The use of contextually relevant evaluation practices with programs designed to increase participation of minorities in science, technology, engineering, and mathematics (STEM) education. In S. Hood, R. Hopson, & H. Frierson (Ed.), *The role of culture*

*and cultural context: A mandate for inclusion, the discovery of truth, and understanding evaluative theory and practice* (pp. 217-235). Greenwich, CT: Information Age Publishing.

Juarez, Gabriela & Greene, Jennifer C. (2011). To Be or Not to Be: Culturally Competent versus Culturally Responsive Evaluator. *Paper Presented at the 2011 American Evaluation Conference*, Anaheim, CA.

Kane, M., & Trochim, W. M. K. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks: Sage Publications.

Kirkhart, K. E. (2011). Culture and Influence in Multisite Evaluation. *New Directions for Evaluation*, 129(129), 73-85.

Noguera, P., & Wing, J. Y.,. (2006). *Unfinished business : closing the racial achievement gap in our schools*. San Francisco: Jossey-Bass, a Wiley imprint.

Passel, J. S., Cohn, D. V., & Pew Hispanic Center. (2008). *U.S. population projections, 2005-2050*. Washington, DC: Pew Research Center.

Rosas, S. R. & Kane, M. (2011). Quality and rigor of the concept mapping methodology: A pooled study analysis. *Manuscript in press*.

Ryan, K., Greene, J., Lincoln, Y., Mathison, S., & Mertens, D. M. (1998). Advantages and challenges of using inclusive evaluation approaches in evaluation practice. *American Journal of Evaluation*, 19(1), 101-22.

Slaughter, H. B. (1991). The Participation of Cultural Informants on Bilingual and Cross-Cultural Evaluation Teams. *Evaluation Practice*, 12(2), 149-57.

Stake, R. (1976). A theoretical statement of responsive evaluation. *Studies in Educational Evaluation*, 2(1), 19-22.

Stake, R. (1976). A theoretical statement of responsive evaluation. *Studies in Educational Evaluation*, 2(1), 19-22.

Thurston, W. E., Graham, J., & Hatfield, J. (2003). Evaluability Assessment: A Catalyst for Program Change and Improvement. *Evaluation & the Health Professions*, 26(2), 206-221.

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. *Evaluation and Program Planning*, 12(1), 1-16.

## Appendix A

Table A

*Original statement set collected prior to statement synthesis (N=93)*

Number	Statements as written by participants	Number	Statements as written by participants
1	encourage for measures to solicit open-ended feedback from participants	48	ensure language in instrument development is culturally sensitive to stakeholders
2	make use of the diversity of my co-workers	49	include questions of relevance to stakeholders
3	as often as possible, beta test the measures with an audience similar to the target population	50	engaging stakeholders with logic models or other related program theory schemes in ways that resonate with how they think about their program
4	review measures for culturally appropriate and accessible language (questions, introductory blurb, etc.)	51	ensuring multiple voices are included in the process of engaging stakeholders
5	use evaluation frameworks that draw upon cultural diverse values and world views, rather than rely on the currently popular, but culturally narrow, logic of "outcome measurement models"	52	pay attention to distributions of power when engaging stakeholders
6	spend time building and nurturing diverse friendships outside work.	53	assemble an evaluation team whose collective lived experiences are appropriate to context of evaluand
7	keep my mouth shut and listen as much as possible.	54	develop a stakeholder group that represents the lived experiences of the evaluand
8	apologize for my inevitable mistakes!	55	center my evaluation thinking and steps around the needs and nuances of the group I am working with
9	regularly examine my own social conditioning -- through critical reflection, conversations, reading, etc.	56	network with colleagues who are versed in diversity and attend diversity trainings.
10	consult regularly with trusted colleagues from backgrounds different than mine who I can count on to offer additional perspectives or point out oversights and mistakes.	57	think about how we could better target diverse populations with our marketing techniques and program implementation.

11	critically question my own assumptions, biases and world views throughout planning, doing and/or evaluating.	58	use venues known to be accessible and accessible to populations of interest
12	continuously learn about how social oppressions (structural and internalized) impact various constituency groups -- both "targeted" and "dominant"	59	Work with and through organizations that have established, credible ties with populations of interest
13	Engage in participatory program planning and participatory evaluation processes	60	rely on "cultural guides" -- people more knowledgeable about populations of interest than I am
14	pay primary attention to the centrality and nature of personal relationships (depth, breadth, quality, type, etc.) rather than to "programs" or "activities."	61	go to their meetings and seek input
15	conduct background research on the group's history	62	work with rural farm populations,
16	include group members in all phases of planning and implementation	63	train employees on diversity sensitivity
17	speak to a member of the group about potentially relevant topics	64	hire diverse employees
18	same as above	65	find diverse board members as advisors
19	Provide social and educational programs that addresses the concerns and needs of the population with which I will be working	66	We work to understand, respond to, and target the particular demographic and SES of the groups  Make contact with individuals who choose not to participate in a program (e.g. individuals referred or registered who do not actually attend a program) to try to identify what barriers prevented their participation. this information helps me to make changes for future program activities.
20	Use terms or concepts in the evaluation that related to their life experiences	67	
21	Try to learn as much as possible about the culture/people/group with which I will be working so I can understand the context and meaning of their comments	68	Encourage participants to share from family traditions, food, stories from their culture.
22	Consider cultural factors and account for them when designing evaluations.	69	Make sure there is time in program workshops to listen to participants and validate that there is much wisdom and many perspectives within the group.

23	draw on theorizations of culture to inform reflective and reflexive practice.	70	Ask members of the population we serve to participate in planning the program. This might include serving on an advisory committee, responding to a proposed activity, participating in a focus group related to the program.
24	plan activities that bring people from different backgrounds into constructive dialogue with each other.	71	We work to better know and understand the groups that are considered underserved by the program. We follow up with attempting to better understand the needs of that audience and ensure that we recruit from that population whenever possible.
25	ensure that cultural factors are considered (as they pertain to context and assumptions) in logic model development.	72	encourage participants to identify methods, outcomes and measurements that are meaningful to them.
26	strive for balanced/mixed groups (in a undergraduate service-learning cohort) - gender, ethnic/racial	73	Be aware of the entire audience
27	identify role models from underrepresented groups (in publications, materials, websites) for STEM careers	74	Recruiting in a variety of ways...paper, on web, and in person with youth.
28	consider needs of secondary underrepresented audiences (students of teachers who are participating in professional development)	75	Ensure that our programs are marketed to the entire county.
29	try to involve role models who are members of an underrepresented group (women, minorities, etc.) in professional development	76	I routinely conduct focus groups or interviews with key informants around program goals and objectives, methods utilized, and what works with which subsets of the target audience.
30	focus on making sure diverse individuals are involved as participants, planners, and evaluators	77	work directly with the administration of the program at the participating sites to understand their resources and capacity to run the project. Ideally, this includes discussion with youth participants in an open forum with onsite staff. We tried very hard to adapt our program in order to ensure success.
31	think hard about how a project may impact college students of color and their feelings regarding the project	78	have them bring foods from their culture which have been modified to be healthier so they can see it is possible to operate in a healthier manner within their

32	be aware of past injustice and harm that has resulted from evaluation or research with this population	79	culture.  allow participants to share their thoughts and practices to make sure I have a good understanding of my actual audience.
33	respect tribal elders and strive to follow appropriate protocol.	80	make sure I have well-researched the group I am working with so I am aware of what should and should not be done.
34	reflect on my own cultural location (age, gender, ethnicity, social class, disability, education, sexual orientation, faith) noticing which elements are particularly relevant to this context	81	be sure that some of people that conduct programs are representative of the populations they are working with (i.e. African American educators working with African-American populations), speak the same language, are authentic
35	consider program outcomes that may differ from majority outcomes; e.g., outcomes that are culturally specific	82	Determine if the program or effort is accessible to people with disabilities. This includes online accessibility.
36	listen carefully to how evaluation has been used and is perceived in this context in the past, listening especially for who benefited and who lost out from prior evaluations	83	we work to develop specific programs aimed at diversifying students in Geosciences.
37	seek culturally responsive evaluation theory to guide my evaluation practice	84	is to consider who my audience is and what there background is in relation to my program. Then I attempt to proceed with my program from a perspective that they can relate to. Hold discussions with the group to make sure that I understand the 'problem' that the program/policy being evaluated/studied from their historic and cultural perspective; to get their insights on what is not working and how it might work better; and how to 'mine'/gather data for authentic community insights, experiences and responses for the eval.
38	avoid framing cultural variables in terms of problems or deficits	85	

39	reflect on the values and assumptions underlying the evaluand, so that I can avoid reinforcing cultural stereotypes	86	learn as much as I can about the target population's culture, current social-economic status and the factors affecting them and then working with them to make sure that the evaluation questions and methods I intent to implement are appropriately nuanced. I also establish agreements for their participation in the reviewing the analysis and report.
40	assess the need for multilingual evaluators on the team	87	engaging participants in planning phase and sharing evaluation results Is to make explicit our practice values, assumptions and philosophy and engage our clients in making their explicit as well. Together, we then define the ways in which we will work together throughout the engagement including frames for research, instrument design, data collection, analysis and decision-making.
41	pay attention to how majority privilege operates to marginalize "others"	88	I work to plan, implement and evaluate programs WITH members of the culture rather than FOR members of the culture.
42	avoid making assumptions about cultural labels and groupings, remembering that cultural location is fluid and defined by multiple identifications	89	I often put together a project/program steering committee that includes members of the culture and those who are culturally responsive.
43	do my homework on the history of the community or organization that I will be entering.	90	Spend time with the audience to get a better sense of the culture. I try to observe and participate in a way that helps me feel the culture rather than simply read about it or ask about it.
44	ensure that the evaluation is framed and focused on benefit to community and stakeholders	91	Ask hands-on workshop participants to bring questions in advance about their goal(s) in attending the workshop. This gives me an insight into the minds of the participants
45	develop and use a stakeholder panel or advisory group assist in the analysis of data of particular cultural meanings of analysis	92	I have attempted to spend time with the population/organization in very informal ways into order to understand org. norms, population behaviors, expectations, etc... I read about the history of the
46	ensure that a wide range of stakeholders receive and use results of evaluation	93	

population/area that the population resides in. My baseline survey has questions regarding "lifestyles," experiences, and work exp. + basic demographics

47 ensuring that data collection is done at multiple points in time and is appropriate to the notions and habits of time of stakeholders

---

Table B

*Statement set after steps 1 and 2 of synthesis*

Number	Statements that have been made grammatically correct and disaggregated	Number	Statements that have been made grammatically correct and disaggregated
1	encourage for measures to solicit open-ended feedback from participants.	62	work with rural farm populations.
2	make use of the diversity of my co-workers.	63	train employees on diversity sensitivity.
3	as often as possible, beta test the measures with an audience similar to the target population.	64	hire diverse employees.
4	review measures for culturally appropriate and accessible language (questions, introductory blurb, etc.)	65	find diverse board members as advisors.
5	use evaluation frameworks that draw upon cultural diverse values and world views, rather than rely on the currently popular, but culturally narrow, logic of "outcome measurement models."	66	work to understand the particular demographic and SES of the groups
6	spend time building and nurturing diverse friendships outside work.	67	work to respond to the particular demographic and SES of the groups
7	keep my mouth shut and listen as much as possible.	68	and work to target the particular demographic and SES of the groups
8	apologize for my inevitable mistakes!	69	make contact with individuals who choose not to participate in a program (e.g. individuals referred or registered who do not actually attend a program) to try to identify what barriers prevented their participation.
9	regularly examine my own social conditioning -- through critical reflection, conversations, reading, etc.	70	to use information about barriers to participation to help me to make changes for future program activities.
10	consult regularly with trusted colleagues from backgrounds different than mine who I can count on to offer additional perspectives or point out oversights and mistakes.	71	to encourage participants to share from family traditions, food, stories from their culture.

11	critically question my own assumptions, biases and world views throughout planning, doing and/or evaluating.	72	make sure there is time in program workshops to listen to participants.
12	continuously learn about how social oppressions (structural and internalized) impact various constituency groups -- both "targeted" and "dominant."	73	validate that there is much wisdom and many perspectives within the group.
13	engage in participatory program planning and participatory evaluation processes.	74	to ask members of the population we serve to participate in planning the program.
14	pay primary attention to the centrality and nature of personal relationships (depth, breadth, quality, type, etc.) rather than to "programs" or "activities."	75	to ask members of the population we serve to serve on an advisory committee.
15	conduct background research on the group's history.	76	to ask members of the population we serve to respond to a proposed activity.
16	include group members in all phases of planning and implementation.	77	to ask members of the population we serve to participate in a focus group related to the program.
17	speak to a member of the group about potentially relevant topics.	78	to work to better know and understand the groups that are considered underserved by the program.
18	(same as above) speak to a member of the group	79	to follow up with attempting to better understand the needs of that audience.
19	provide social and educational programs that addresses the concerns and needs of the population with which I am working.	80	ensure that we recruit from that population whenever possible.
20	to use terms or concepts in the evaluation that related to their life experiences.	81	encourage participants to identify methods that are meaningful to them.
21	try to learn as much as possible about the culture/people/group with which I will be working so I can understand the context and the meaning of their comments	82	encourage participants to identify outcomes that are meaningful to them.
22	consider cultural factors and account for them when designing evaluations.	83	encourage participants to identify measurements that are meaningful to them.
23	to draw on theorizations of culture to inform reflective and reflexive practice.	84	to be aware of the entire audience.
24	to plan activities that bring people from different backgrounds into constructive dialogue with each other.	85	recruit in a variety of ways (e.g. by paper, on web, and in person with youth.)

25	ensure that cultural factors are considered (as they pertain to context and assumptions) in logic model development.	86	ensure that our programs are marketed to the entire county.
26	to strive for balanced/mixed groups (in an undergraduate service-learning cohort) - gender, ethnic/racial.	87	to routinely conduct focus groups or interviews with key informants around program goals and objectives, methods utilized, and what works with which subsets of the target audience.
27	identify role models from underrepresented groups (in publications, materials, websites) for STEM careers.	88	to routinely conduct focus groups or interviews with key informants around methods utilized.
28	consider needs of secondary underrepresented audiences (students of teachers who are participating in professional development).	89	to routinely conduct focus groups or interviews with key informants around what works with which subsets of the target audience.
29	to try to involve role models who are members of an underrepresented group (women, minorities, etc.) in professional development.	90	to work directly with the administration of the program at the participating sites to understand their resources and capacity to run the project.
30	focus on making sure diverse individuals are involved as participants, planners, and evaluators.	91	to include discussion with youth participants in an open forum with onsite staff.
31	think hard about how a project may impact college students of color and their feelings regarding the project.	92	to have them bring foods from their culture which have been modified to be healthier so they can see it is possible to operate in a healthier manner within their culture.
32	to be aware of past injustice and harm that has resulted from evaluation or research with this population.	93	to allow participants to share their thoughts to make sure I have a good understanding of my actual audience.
33	respect tribal elders and strive to follow appropriate protocol.	94	make sure I have well-researched the group I am working with so I am aware of what should and should not be done.
34	reflect on my own cultural location (age, gender, ethnicity, social class, disability, education, sexual orientation, faith) noticing which elements are particularly relevant to this context.	95	be sure that some of people that conduct programs are representative of the populations they are working with (i.e. African American educators working with African-American populations)
35	consider program outcomes that may differ from majority outcomes; e.g., outcomes that are culturally specific.	96	be sure that some of people that conduct programs are authentic.
36	listen carefully to how evaluation has been used and is	97	be sure that some of people that conduct programs

	perceived in this context in the past, listening especially for who benefited and who lost out from prior evaluations.		98	to determine if the program or effort is accessible to people with disabilities (e.g. this includes online accessibility.)
37	seek culturally responsive evaluation theory to guide my evaluation practice.		99	we work to develop specific programs aimed at diversifying students in Geosciences.
38	avoid framing cultural variables in terms of problems or deficits.		100	consider who my audience is.
39	reflect on the values and assumptions underlying the evaluand, so that I can avoid reinforcing cultural stereotypes.		101	consider what there background is in relation to my program.
40	assess the need for multilingual evaluators on the team.		102	attempt to proceed with my program from a perspective that they can relate to.
41	pay attention to how majority privilege operates to marginalize "others".		103	hold discussions with the group to make sure that I understand the 'problem' that the program/policy being evaluated/studied from their historic and cultural perspective.
42	avoid making assumptions about cultural labels and groupings, remembering that cultural location is fluid and defined by multiple identifications.		104	hold discussions with the group about how to 'mine'/gather data for authentic community insights, experiences and responses for the eval.
43	do my homework on the history of the community or organization that I will be entering.		105	hold discussions with the group to get their insights on what is not working and how it might work better.
44	ensure that the evaluation is framed and focused on benefit to community and stakeholders.		106	to learn as much as I can about the target population's culture, current social-economic status and the factors affecting them.
45	to develop and use a stakeholder panel or advisory group to assist in the analysis of data of particular cultural meanings of analysis.		107	work with them to make sure that the evaluation questions and methods I intend to implement are appropriately nuanced.
46	ensure that a wide range of stakeholders receive and use results of the evaluation.		108	establish agreements for their participation in reviewing the analysis and report.
47	ensure that data collection is done at multiple points in time that are appropriate to the notions and habits of time of the stakeholders.			

48	ensure language in instrument development is culturally sensitive to stakeholders.	109	engage participants in planning phase.
49	to include questions of relevance to stakeholders.	110	share evaluation results.
50	to engage stakeholders with logic models or other related program theory schemes in ways that resonate with how they think about their program.	111	to make explicit our practice values, assumptions, and philosophy.
51	ensure that multiple voices are included in the process of engaging stakeholders.	112	engage our clients in making their values, assumptions, and philosophy.
52	pay attention to distributions of power when engaging stakeholders.	113	to define, together, the ways in which we will work together throughout the engagement.
53	assemble an evaluation team whose collective lived experiences are appropriate to context of evaluand.	114	to include frames for research, instrument design, data collection, analysis and decision-making.
54	develop a stakeholder group that represents the lived experiences of the evaluand.	115	to work to plan, implement and evaluate programs WITH members of the culture rather than FOR members of the culture.
55	center my evaluation thinking and steps on the needs and nuances of the group I am working with.	116	to put together a project/program steering committee that includes members of the culture and those who are culturally responsive.
56	network with colleagues who are versed in diversity and attend diversity trainings.	117	to spend time with the audience to get a better sense of the culture.
57	think about how we could better target diverse populations with our marketing techniques and program implementation.	118	try to observe and participate in a way that helps me feel the culture rather than simply read about it or ask about it.
58	by using venues known to be accessible and accessible to populations of interest.	119	to ask hands-on workshop participants to bring questions in advance about their goal(s) in attending the workshop.
59	work with and through organizations that have established, credible ties with populations of interest.	120	to spend time with the population/organization in very informal ways into order to understand org. norms, population behaviors, expectations, etc...
60	rely on "cultural guides" -- people more knowledgeable about populations of interest than I am.	121	to read about the history of the population/area that the population resides in.
61	to go to their meetings and seek input.	122	conduct a baseline survey that has questions regarding "lifestyles," experiences, and work experience and basic demographics.

---

Table C

*Statement set after completion of synthesis steps*

Number	Statements after removing redundant or incoherent ideas	Number	Statements after removing redundant or incoherent ideas
1	encourage for measures to solicit open-ended feedback from participants.	50	rely on "cultural guides" -- people more knowledgeable about populations of interest than I am.
2	make use of the diversity of my co-workers.	51	to go to their meetings and seek input.
3	as often as possible, beta test the measures with an audience similar to the target population.	52	train employees on diversity sensitivity.
4	review measures for culturally appropriate and accessible language (questions, introductory blurb, etc.)	53	hire diverse employees.
5	use evaluation frameworks that draw upon culturally diverse values and world views, rather than rely on the currently popular, but culturally narrow, logic of "outcome measurement models."	54	find diverse board members as advisors.
6	spend time building and nurturing diverse friendships outside work.	55	work to understand, target, and respond to the particular demographic and SES of the groups.
7	keep my mouth shut and listen as much as possible.	56	make contact with individuals who choose not to participate in a program (e.g. individuals referred or registered who do not actually attend a program) to try to identify what barriers prevented their participation.
8	apologize for my inevitable mistakes!	57	to use information about barriers to participation to help me to make changes for future program activities.
9	consult regularly with trusted colleagues from backgrounds different than mine who I can count on to offer additional perspectives or point out oversights and mistakes.	58	to encourage participants to share from family traditions, food, stories from their culture.
10	critically question my own assumptions, biases and world views throughout planning, doing and/or evaluating.	59	to make sure there is time in program workshops to listen to participants.

11	continuously learn about how social oppressions (structural and internalized) impact various constituency groups -- both "targeted" and "dominant."	60	validate that there is much wisdom and many perspectives within the group.
12	engage in participatory program planning and participatory evaluation processes.	61	to ask members of the population we serve to participate in a focus group related to the program.
13	pay primary attention to the centrality and nature of personal relationships (depth, breadth, quality, type, etc.) rather than to "programs" or "activities."	62	to work to better know and understand the groups that are considered underserved by the program.
14	include group members in all phases of planning and implementation.	63	to center my evaluation and follow up with attempting to better understand the needs of that audience.
15	speak to a member of the group about potentially relevant topics.	64	ensure that we recruit from that population whenever possible.
16	to use terms or concepts in the evaluation that related to their life experiences.	65	to encourage participants to identify methods and outcomes that are meaningful to them.
17	try to learn as much as possible about the culture/people/group with which I will be working so I can understand the context and the meaning of their comments	66	to be aware of the entire audience.
18	to draw on theorizations of culture to inform reflective and reflexive practice.	67	recruit in a variety of ways (e.g. by paper, on web, and in person with youth.)
19	to plan activities that bring people from different backgrounds into constructive dialogue with each other.	68	ensure that our programs are marketed to the entire county.
20	ensure that cultural factors are considered (as they pertain to context and assumptions) in logic model development and evaluation design.	69	to routinely conduct focus groups or interviews with key informants around program goals and objectives, methods utilized, and what works with which subsets of the target audience.
21	to strive for balanced/mixed groups (in an undergraduate service-learning cohort) - gender, ethnic/racial.	70	to work directly with the administration of the program at the participating sites to understand their resources and capacity to run the project.
22	consider needs of secondary underrepresented audiences (e.g. students of teachers who are participating in professional development).	71	to include discussion with youth participants in an open forum with onsite staff.

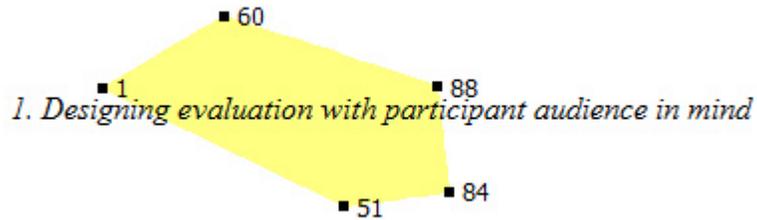
- |    |   |    |  |
|----|---|----|--|
| 23 | to try to involve role models who are members of an underrepresented group (i.e. women, minorities, etc.) in professional development.  | 72 | to have them bring foods from their culture which have been modified to be healthier so they can see it is possible to operate in a healthier manner within their culture.             |
| 24 | focus on making sure diverse individuals are involved as participants, planners, and evaluators.  | 73 | to allow participants to share their thoughts to make sure I have a good understanding of my actual audience.  |
| 25 | think hard about how a project may impact college students of color and their feelings regarding the project.   | 74 | make sure I have well-researched the group I am working with so I am aware of what should and should not be done.  |
| 26 | to be aware of past injustice and harm that has resulted from evaluation or research with this population.  | 75 | be sure that some people that conduct programs are representative of the populations they are working with (i.e. African American educators working with African-American populations) |
| 27 | reflect on my own cultural location (age, gender, ethnicity, social class, disability, education, sexual orientation, faith) noticing which elements are particularly relevant to this context. | 76 | be sure that some of the people that conduct programs are authentic and/or speak the same language.  |
| 28 | consider program outcomes that may differ from majority outcomes, e.g. outcomes that are culturally specific.   | 77 | to determine if the program or effort is accessible to people with disabilities (e.g. this includes online accessibility.)   |
| 29 | listen carefully to how evaluation has been used and is perceived in this context in the past, listening especially for who benefited and who lost out from prior evaluations.                  | 78 | consider what their background is in relation to my program.   |
| 30 | seek culturally responsive evaluation theory to guide my evaluation practice.   | 79 | attempt to proceed with my program from a perspective that they can relate to.   |
| 31 | avoid framing cultural variables in terms of problems or deficits.  | 80 | hold discussions with the group to make sure that I understand the 'problem' that the program/policy is being evaluated/studied from their historic and cultural perspective.          |
| 32 | reflect on the values and assumptions underlying the evaluand, so that I can avoid reinforcing cultural stereotypes.  | 81 | hold discussions with the group about how to get 'mine'/gather data for authentic community insights, experiences and responses for the  |

33	assess the need for multilingual evaluators on the team.	82	evaluation.
34	pay attention to how majority privilege operates to marginalize "others".	83	hold discussions with the group to get their insights on what is not working and how it might work better.
35	avoid making assumptions about cultural labels and groupings, remembering that cultural location is fluid and defined by multiple identifications.	84	to learn as much as I can about the target population's culture, current social-economic status and the factors affecting them.
36	ensure that the evaluation is framed and focused on benefit to community and stakeholders.	85	work with them to make sure that the evaluation questions and methods I intend to implement are appropriately nuanced.
37	to develop and use a stakeholder panel or advisory group to assist in the analysis of data of particular cultural meanings of analysis.	86	establish agreements for their participation in reviewing the analysis and report.
38	ensure that a wide range of stakeholders receive and use results of the evaluation.	87	share evaluation results.
39	ensure that data collection is done at multiple points in time that are appropriate to the notions and habits of time of the stakeholders.	88	to make explicit our practice values, assumptions, and philosophy.
40	to include questions of relevance to stakeholders.	89	engage our clients in making explicit their values, assumptions, and philosophy.
41	to engage stakeholders with logic models or other related program theory schemes in ways that resonate with how they think about their program.	90	to define, together, the ways in which we will work together throughout the engagement.
42	ensure that multiple voices are included in the process of engaging stakeholders.	91	to include frames for research, instrument design, data collection, analysis and decision-making.
43	pay attention to distributions of power when engaging stakeholders.	92	to work to plan, implement and evaluate programs WITH members of the culture rather than FOR members of the culture.
44	assemble an evaluation team whose collective lived experiences are appropriate to the context of the evaluand.	93	to put together a project/program steering committee that includes members of the culture and those who are culturally responsive.
			to spend time with the audience to get a better sense of the culture.

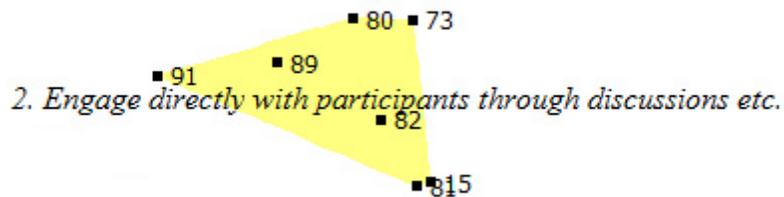
45	develop a stakeholder group that represents the lived experiences of the evaluand.	94	try to observe and participate in a way that helps me feel the culture rather than simply read about it or ask about it.
46	network with colleagues who are versed in diversity and attend diversity trainings.	95	to ask hands-on workshop participants to bring questions in advance about their goal(s) in attending the workshop.
47	to think about how we could better target diverse populations with our marketing techniques and program implementation.	96	to spend time with the population/organization in very informal ways in order to understand organizational norms, population behaviors, expectations, etc...
48	by using venues known to be accessible and specifically accessible to populations of interest.	97	conduct a baseline survey that has questions regarding "lifestyles," experiences, work experience and basic demographics.
49	work with and through organizations that have established, credible ties with populations of interest.		

---

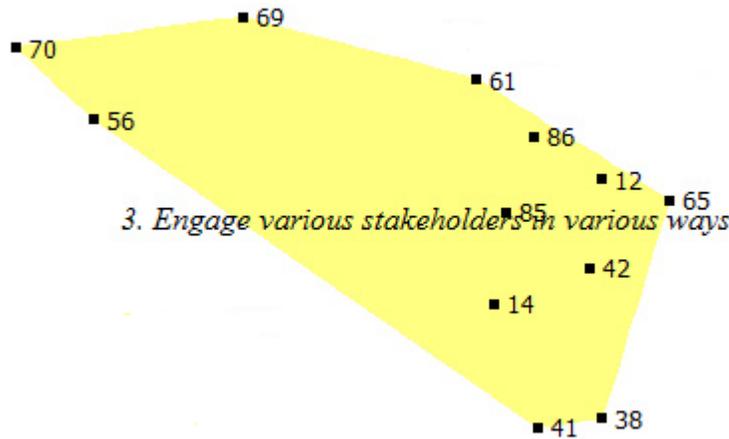
## Appendix B



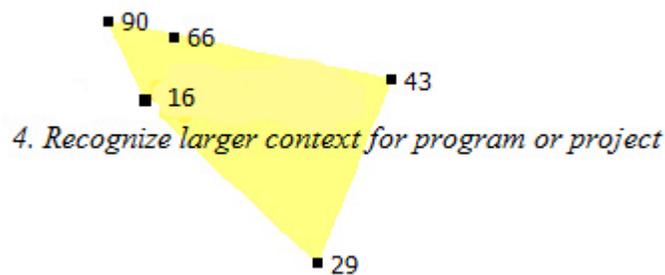
- 1. encourage for measures to solicit open-ended feedback from participants.
- 51. to go to their meetings and seek input.
- 60. validate that there is much wisdom and many perspectives within the group.
- 84. work with them to make sure that the evaluation questions and methods I intend to implement are appropriately nuanced.
- 88. engage our clients in making explicit their values, assumptions, and philosophy.



- 15. speak to a member of the group about potentially relevant topics.
- 73. to allow participants to share their thoughts to make sure I have a good understanding of my actual audience.
- 80. hold discussions with the group to make sure that I understand the 'problem' that the program/policy is being evaluated/studied from their historic and cultural perspective.
- 81. hold discussions with the group about how to get 'mine'/gather data for authentic community insights, experiences and responses for the evaluation.
- 82. hold discussions with the group to get their insights on what is not working and how it might work better.
- 89. to define, together, the ways in which we will work together throughout the engagement
- 91. to work to plan, implement and evaluate programs WITH members of the culture rather than FOR members of the culture.

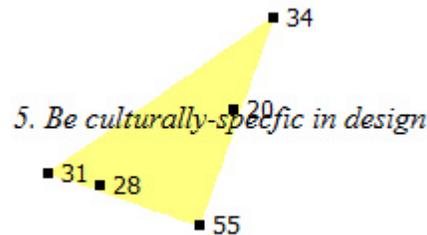


- 12. engage in participatory program planning and participatory evaluation processes.
- 14. include group members in all phases of planning and implementation.
- 38. ensure that a wide range of stakeholders receive and use results of the evaluation.
- 41. to engage stakeholders with logic models or other related program theory schemes in ways that resonate with how they think about their program.
- 42. ensure that multiple voices are included in the process of engaging stakeholders.
- 56. make contact with individuals who choose not to participate in a program (e.g. individuals referred or registered who do not actually attend a program) to try to identify what barriers prevented their participation.
- 61. to ask members of the population we serve to participate in a focus group related to the program.
- 65. to encourage participants to identify methods and outcomes that are meaningful to them.
- 69. to routinely conduct focus groups or interviews with key informants around program goals and objectives, methods utilized, and what works with which subsets of the target audience.
- 70. to work directly with the administration of the program at the participating sites to understand their resources and capacity to run the project.
- 85. establish agreements for their participation in reviewing the analysis and report.
- 86. share evaluation results.

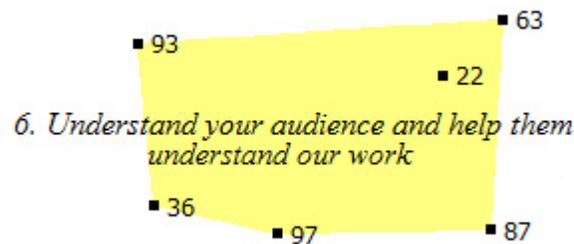


- 16. to use terms or concepts in the evaluation that related to their life experiences.
- 29. listen carefully to how evaluation has been used and is perceived in this context in the past,

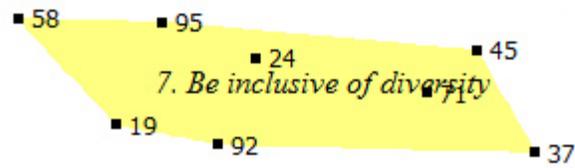
listening especially for who benefited and who lost out from prior evaluations.  
 43. pay attention to distributions of power when engaging stakeholders.  
 66. to be aware of the entire audience.  
 90. to include frames for research, instrument design, data collection, analysis and decision-making.



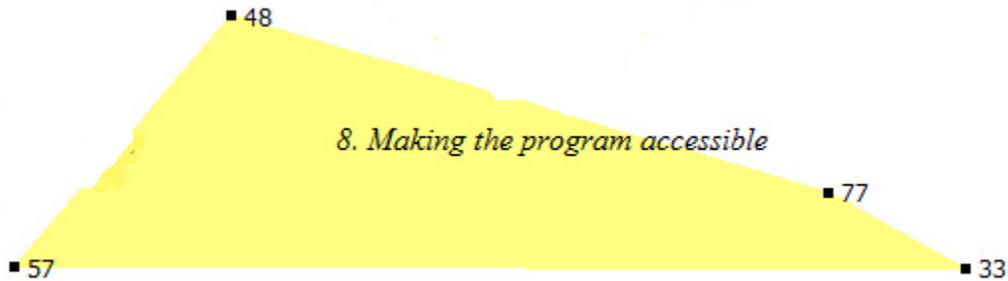
20. ensure that cultural factors are considered (as they pertain to context and assumptions) in logic model development and evaluation design. (20)  
 28. consider program outcomes that may differ from majority outcomes, e.g. outcomes that are culturally specific. (28)  
 31. avoid framing cultural variables in terms of problems or deficits. (31)  
 34. pay attention to how majority privilege operates to marginalize "others". (34)  
 55. work to understand, target, and respond to the particular demographic and SES of the groups.



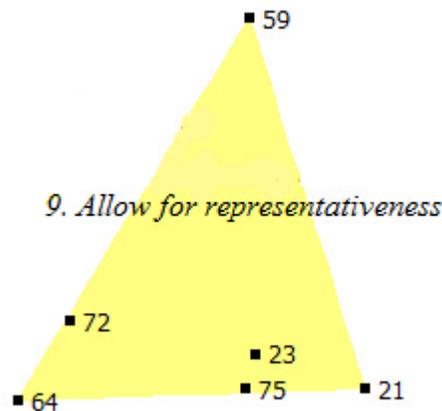
22. consider needs of secondary underrepresented audiences (e.g. students of teachers who are participating in professional development).  
 36. ensure that the evaluation is framed and focused on benefit to community and stakeholders.  
 63. to center my evaluation and follow up with attempting to better understand the needs of that audience.  
 87. to make explicit our practice values, assumptions, and philosophy.  
 93. to spend time with the audience to get a better sense of the culture.  
 97. conduct a baseline survey that has questions regarding "lifestyles," experiences, work experience and basic demographics.



- 19. to plan activities that bring people from different backgrounds into constructive dialogue with each other.
- 24. focus on making sure diverse individuals are involved as participants, planners, and evaluators.
- 37. to develop and use a stakeholder panel or advisory group to assist in the analysis of data of particular cultural meanings of analysis.
- 45. develop a stakeholder group that represents the lived experiences of the evaluand.
- 58. to encourage participants to share from family traditions, food, stories from their culture.
- 71. to include discussion with youth participants in an open forum with onsite staff.
- 92. to put together a project/program steering committee that includes members of the culture and those who are culturally responsive.
- 95. to ask hands-on workshop participants to bring questions in advance about their goal(s) in attending the workshop.



- 33. assess the need for multilingual evaluators on the team.
- 48. by using venues known to be accessible and specifically accessible to populations of interest.
- 57. to use information about barriers to participation to help me to make changes for future program activities.
- 77. to determine if the program or effort is accessible to people with disabilities (e.g. this includes online accessibility.)



21. to strive for balanced/mixed groups (in an undergraduate service-learning cohort) - gender, ethnic/racial.

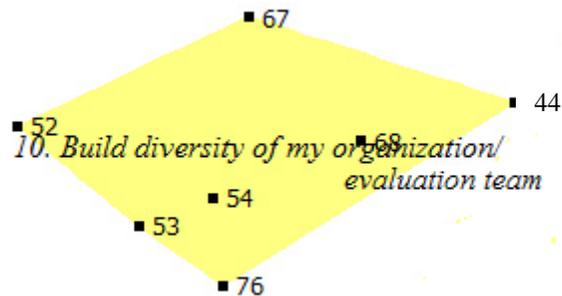
23. to try to involve role models who are members of an underrepresented group (i.e. women, minorities, etc.) in professional development.

59. to make sure there is time in program workshops to listen to participants.

64. ensure that we recruit from that population whenever possible.

72. to have them bring foods from their culture which have been modified to be healthier so they can see it is possible to operate in a healthier manner within their culture.

75. be sure that some people that conduct programs are representative of the populations they are working with (i.e. African American educators working with African-American populations)



44. assemble an evaluation team whose collective lived experiences are appropriate to the context of the evaluand.

52. train employees on diversity sensitivity.

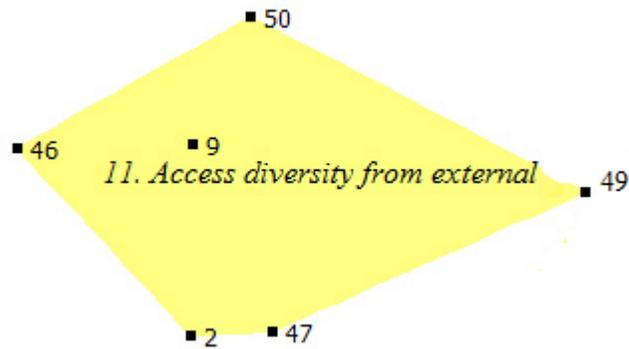
53. hire diverse employees.

54. find diverse board members as advisors.

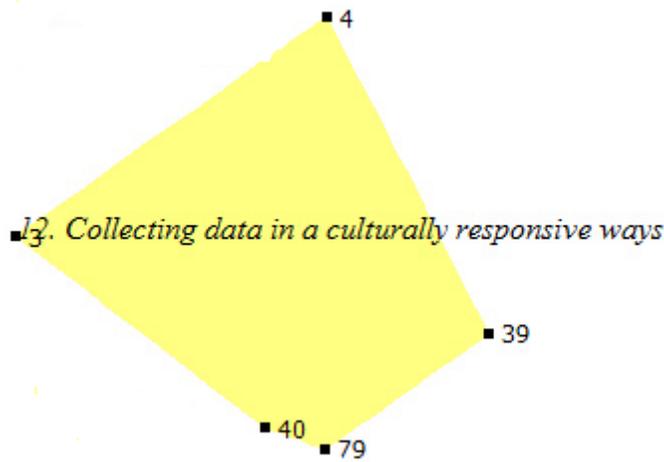
67. recruit in a variety of ways (e.g. by paper, on web, and in person with youth.)

68. ensure that our programs are marketed to the entire county.

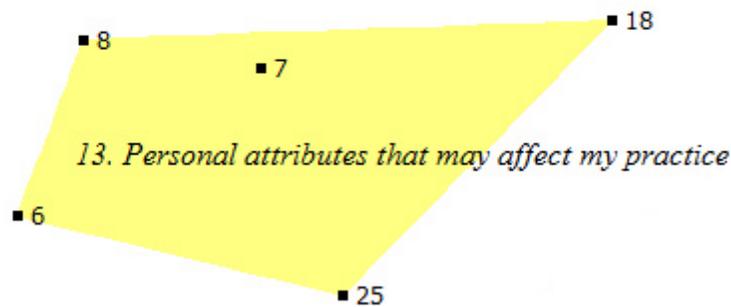
76. be sure that some of the people that conduct programs are authentic and/or speak the same language.



- 2. make use of the diversity of my co-workers.
- 9. consult regularly with trusted colleagues from backgrounds different than mine who I can count on to offer additional perspectives or point out oversights and mistakes.
- 46. network with colleagues who are versed in diversity and attend diversity trainings.
- 47. to think about how we could better target diverse populations with our marketing techniques and program implementation.
- 49. work with and through organizations that have established, credible ties with populations of interest.
- 50. rely on "cultural guides" -- people more knowledgeable about populations of interest than I am.



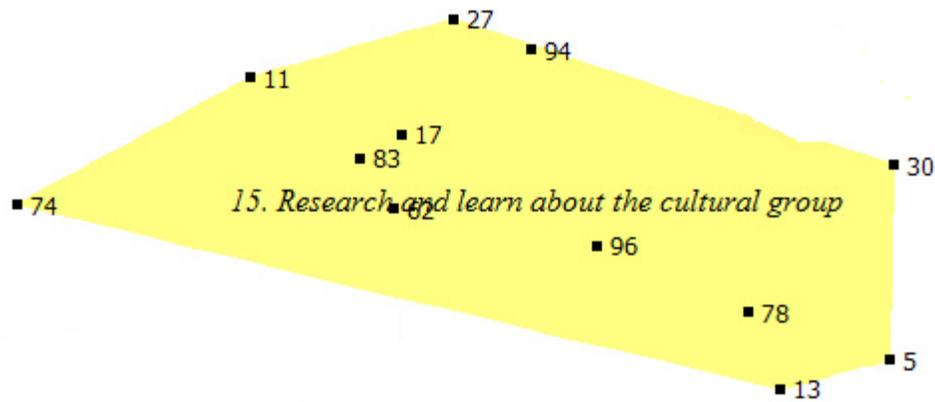
- 3. as often as possible, beta test the measures with an audience similar to the target population.
- 4. review measures for culturally appropriate and accessible language (questions, introductory blurb, etc.)
- 39. ensure that data collection is done at multiple points in time that are appropriate to the notions and habits of time of the stakeholders.
- 40. to include questions of relevance to stakeholders.
- 79. attempt to proceed with my program from a perspective that they can relate to.



- 6. spend time building and nurturing diverse friendships outside work.
- 7. keep my mouth shut and listen as much as possible.
- 8. apologize for my inevitable mistakes!
- 18. to draw on theorizations of culture to inform reflective and reflexive practice.
- 25. think hard about how a project may impact college students of color and their feelings regarding the project.



- 10. critically question my own assumptions, biases and world views throughout planning, doing and/or evaluating.
- 26. to be aware of past injustice and harm that has resulted from evaluation or research with this population.
- 32. reflect on the values and assumptions underlying the evaluand, so that I can avoid reinforcing cultural stereotypes.
- 35. avoid making assumptions about cultural labels and groupings, remembering that cultural location is fluid and defined by multiple identifications.



5. use evaluation frameworks that draw upon culturally diverse values and world views, rather than rely on the currently popular, but culturally narrow, logic of "outcome measurement models."

11. continuously learn about how social oppressions (structural and internalized) impact various constituency groups -- both "targeted" and "dominant."

13. pay primary attention to the centrality and nature of personal relationships (depth, breadth, quality, type, etc.) rather than to "programs" or "activities."

17. try to learn as much as possible about the culture/people/group with which I will be working so I can understand the context and the meaning of their comments.

27. reflect on my own cultural location (age, gender, ethnicity, social class, disability, education, sexual orientation, faith) noticing which elements are particularly relevant to this context.

62. seek culturally responsive evaluation theory to guide my evaluation practice.

to work to better know and understand the groups that are considered underserved by the program.

74. make sure I have well-researched the group I am working with so I am aware of what should and should not be done.

78. consider what their background is in relation to my program.

83. to learn as much as I can about the target population's culture, current social-economic status and the factors affecting them.

94. try to observe and participate in a way that helps me feel the culture rather than simply read about it or ask about it.

96. to spend time with the population/organization in very informal ways in order to understand organizational norms, population behaviors, expectations, etc.