American Journal of Evaluation
33(4) 496-514

DOI: 10.1177/1098214012445280
http://aje.sagepub.com

Development and Validation of the Cultural Competence of Program Evaluators (CCPE) Self-Report Scale

Krystall E. Dunaway[1], Jennifer A. Morrow[2], and Bryan E. Porter[3]

Abstract

No self-report measure of cultural competence currently exists in program evaluation. Adapting items from cultural competence measures in fields such as counseling and nursing, the researchers developed the Cultural Competence of Program Evaluators (CCPE) self-report scale. The goals of this study were to validate the CCPE and to assess differences in level of cultural competence among program evaluators based on various demographic variables. The sample consisted of 174 evaluators. Principal components analyses revealed three factors of the CCPE: cultural knowledge, cultural skills, and cultural awareness. The overall alpha of the CCPE was .88, and convergent validity was established via significant positive correlations between the CCPE and the Multicultural Counseling Inventory (MCI). Additionally, individuals who had received cultural competence training scored significantly higher on the CCPE, and receipt of cultural competence training was a significant predictor of scores on the CCPE.

Keywords

cultural competence, survey development, validation, statistics

As part of its *Guiding Principles for Evaluators,* the American Evaluation Association (AEA) requires that evaluators develop cultural competencies. Such competence is especially important, given the dramatically changing composition of the United States. The U.S. Census Bureau (2007) reports that racial and ethnic minorities, comprising approximately 100 million people, account for about one third of the nation's population. It is estimated that by 2025 ethnic minorities

[1] Virginia Beach City Public Schools, Virginia Beach, VA, USA
[2] University of Tennessee, Knoxville, TN, USA
[3] Old Dominion University, Norfolk, VA, USA

Corresponding Author:
Krystall E. Dunaway, Department of Educational Leadership and Assessment, Virginia Beach City Public Schools, 2512 George Mason Drive, Virginia Beach, VA 23456, USA
Email: krystall.dunaway@vbschools.com

will comprise 40% of all Americans and that by 2050, non-Whites will become the majority (Barrett & George, 2005; Stanhope, Solomon, Pernell-Arnold, Sands, & Bourjolly, 2005).

This growth in minority populations has led to the expectation that evaluators will work effectively with increasingly diverse groups. Our capacity to do this will depend on the acquisition of cultural competence (Hansen, Pepitone-Arreola-Rockwell, & Greene, 2000; Stanhope et al., 2005). Evaluation has historically been based upon Eurocentric perspectives and assumptions, which have limited applicability to racially and culturally diverse populations (Alkon, Tschann, Ruane, Wolff, & Hittner, 2001; Sue, Bingham, Porche-Burke, & Vasquez, 1999). This Eurocentric approach is characterized by the use of an *etic* perspective, which is a broad generic cultural awareness that is often too theoretical and abstract for use in real-world settings (Benavente, 2004; Dumas, Rollock, Prinz, Hops, & Blechman, 1999).

Evaluators would benefit from an *emic* perspective, which attempts to understand a phenomenon from the *native's* point of view. This perspective takes into account the values and traditions of different ethnic groups and focuses on the intrinsic cultural distinctions that are meaningful to the members of a given group (Alkon et al., 2001; Barrett & George, 2005). Program evaluators can avoid the dangers of an etic perspective by evaluating programs and assessing impacts through lenses in which culture is considered an important factor (Frechtling, 2002). Conducting evaluations using an emic perspective allows evaluators to make interpersonal connections and appropriate cultural judgments in the design and implementation of the evaluation. The challenge, however, is how one learns to do this. This article outlines the creation and validation of a self-report measure of cultural competence as well as an initial attempt to assess differences in cultural competence among program evaluators based on various demographic variables (e.g., age, gender, minority status, number of years of evaluation experience, and receipt of cultural competence training). It closes with conclusions and suggestions for future research using this new self-report measure.

What is Cultural Competence?

Generally, cultural competence can be defined as a dynamic process of framing assumptions, knowledge, and meaning from a cultural perspective different than one's own, and this allows professionals to work effectively in cross-cultural situations (Abernethy, 2005; Abrums & Leppa, 2001; Alkon et al., 2001; Stanhope et al., 2005). Specific to program evaluation, cultural competence refers to an awareness, understanding, and appreciation for cultural context when framing an evaluation, developing methodology, interacting with stakeholders, and interpreting results (SenGupta, Hopson, & Thompson-Robinson, 2004). This definition does not describe a static process; rather it incorporates the notion of responsiveness to culturally contextual factors (SenGupta et al., 2004).

Cultural competence is best viewed in terms of what one is *becoming* as opposed to what one *is*; thus, acquiring cultural competence should never be treated as a one-time process resulting in expert status (Doutrich & Storey, 2004; McPhatter & Ganaway, 2003; Mendias & Guevara, 2001). A culturally competent individual can identify with one culture but understand the behaviors of another cultural group in relation to the cultural rules of *that* culture rather than his or her own (Guzman, 2003; Howard, 2002; Symonette, 2004).

Components of Cultural Competence

There are many models that describe how cultural competence is attained (Abernethy, 2005; Campinha-Bacote, 2002; McPhatter & Ganaway, 2003; Prochaska & DiClemente, 1982); arguably the most common paradigm of cultural competence consists of the three components of cultural awareness, cultural knowledge, and cultural skills (Figure 1). Building on this long-standing tradition of use, the current study employed this paradigm when developing the self-report measure.

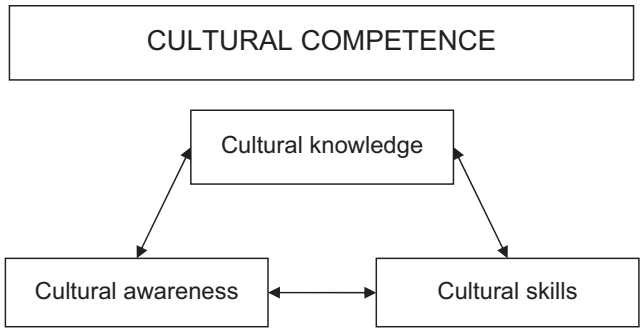

Figure 1. Cultural competence paradigm.

Cultural awareness includes the process of understanding one's own culture, biases, tendency to stereotype, reference-group membership, and power relations. Cultural knowledge includes learning about the attitudes, values, beliefs, and behaviors of cultural groups. Cultural skills focus on communication ability and training learners to be aware of cross-cutting cultural issues (Benavente, 2004; Betancourt, 2003; Pope & Reynolds, 1997; Sue, Arredondo, & McDavis, 1992; Wear, 2003).

These three facets are essential to culturally competent behavior and are also prerequisites to working effectively and ethically with individuals of all backgrounds (AEA, 2004). Additionally, the components are seen as interdependent contributors of equal importance in the pursuit of cultural competence; for example, increasing cultural awareness is a self-reflective process that does not increase cultural skills. And cultural knowledge can lead to stereotyping and oversimplification of culture (an etic perspective) if not coupled with cultural awareness. Cultural skills cannot logically be attained without proper cultural knowledge (Betancourt, 2003). Thus, unless all three components have been attended to, an individual cannot demonstrate cultural competence.

Why is Cultural Competence Important in Program Evaluation?

Cultural competence is important for three main reasons. First, all members of society develop and form a sense of self and others in the context of culture; in other words, each person's experiences are culturally bound (Carter, 2003). The presence of cultural competence alters potentially inappropriate culturally bound perceptions (e.g., racism, sexism) and prevents evaluators from considering their beliefs, customs, and behaviors as unique benchmarks by which to evaluate others (Beagan, 2003; Dumas et al., 1999; Greene, 1997; Guzman, 2003; Kirkhart, 1995).

Second, the questions participants are willing to answer, those with whom they are willing to share their perceptions, and the extent to which they are willing to participate throughout an evaluation are profoundly influenced by their perceptions of the evaluator (Hood & Cassaro, 2002). Therefore, it is important for evaluators to ask the question, "How do those with whom I am seeking to communicate perceive me?" The evaluator who considers this question is practicing *multilateral self-awareness* (Symonette, 2004). Such awareness is an instrumental component in the development of cultural competence, meaning that the individual is viewing oneself as "self in context" rather than simply as the one whom one sees oneself to be (Carter, 2003).

Finally, cultural competence is a necessary and important skill for *everyone,* regardless of race, ethnicity, or gender. In other words, cultural competence should not just be a concern for those in the majority group. For example, Ladson, Lin, Flores, and Magrane (2006) found that Blacks are no more likely than non-Blacks to possess the knowledge, skills, and ability to negotiate encounters or situations with people from diverse cultures. And Abernethy (2005) found that cultural

competence is a vital skill for individuals working with people from *similar* backgrounds, because overidentification between evaluator and evaluatee can be just as detrimental as lack of understanding.

Program evaluation has embraced the idea that the field should be imbued with cultural competence, as evidenced by AEA's (2011) recent Public Statement on Cultural Competence in Evaluation. It remains unclear, however, how an evaluator can establish a culturally competent perspective and when this perspective would be appropriate in the evaluation process (Guzman, 2003). Methods for increasing the cultural competence of evaluations include: (1) considering the community for whom the evaluation plan is created, (2) pretesting survey instruments with different ethnic groups, (3) obtaining information about other attributes related to ethnicity beyond self-identification of ethnic group, (4) building a process check into the evaluation by holding ongoing discourse with the evaluation team concerning their experiences with participants, (5) using triangulation of multiple information sources, (6) including expert cultural or ethnic consultants on the evaluation team, and (7) creating research reports that contain full discussions of the sample and sampling methodology used (Alkon et al., 2001; Guzman, 2003; Okazaki & Sue, 1995; Taket & White, 1997).

There is currently no self-report measure of cultural competence for the field of program evaluation. Such a measure could serve as a first step in allowing evaluators to understand the importance of recognizing, appreciating, and incorporating culturally contextual factors into their practice (SenGupta et al., 2004). The importance of this issue is further underscored by the fact that cultural competence is not readily monitored by any governing body (e.g., AEA) beyond institutions of higher education. For instance, graduate students in evaluation programs have professors and advisors who are in a position to ensure that they are conducting evaluations in a culturally competent manner. However, once students graduate and begin conducting evaluations independently, no safeguards are in place to make sure that those evaluations are being carried out with cultural competence in mind. Against this background, the objectives of this study were (1) to develop a self-report measure of cultural competence that assessed the three components of the cultural competence paradigm for use as a training tool for evaluators and (2) to assess differences in level of cultural competence among evaluators based on various demographic variables. One goal and six research questions were addressed:

Goal

1) The new cultural competence self-report measure will exhibit moderate to high (>.70) reliability and validity.

Research Questions

1) Will level of cultural competence be higher among individuals with more years of evaluation experience?
2) Will there be a gender difference in level of cultural competence?
3) Will there be a difference in level of cultural competence based on minority status (minority vs. nonminority)?
4) Will there be a difference in level of cultural competence based on age?
5) Will level of cultural competence be higher among individuals who have received formal cultural competence training?[1]
6) What are the best demographic predictors of cultural competence?

Method

Participants

Because the researchers wanted to sample only individuals whose background and experience were relevant to the topic of cultural competence in evaluation (i.e., program evaluators), purposive

Table 1. Demographics of Participants

Variable		Sample %	AEA membership %
Gender	Male	25	33
	Female	75	67
Ethnicity[a]10	White	82	73
	Black/African American	6	7
	Asian	0	5
	Hispanic	6	5
	American Indian or Alaskan Native	2	2
	Native Hawaiian/Other Pacific Islander	1	<1
	Other	10	2
Highest degree	Doctorate	55	52
	Master's	39	42
	Bachelor's	6	7
Years of evaluation experience	Less than 5 years	30	33
	6–10 years	23	24
	11–15 years	17	16
	16 or more years	29	27
	No response	2	0
Age	20s or 30s	38	33
	40s	25	24
	50s	27	29
	60s or older	9	14
	No response	2	0
Nation of Origin[b]	USA	74	—
	Non-USA	19	—
	No response	7	—

Note: Due to rounding, percentages may not total 100. In addition to being asked their race, participants were asked if they were of Hispanic origin. Ten participants indicated that they were of Hispanic origin, and the breakdown is as follows: 6 White, 3 "Other," and 1 Black or African American.

[a]Participants could select more than one response.

[b]AEA reports "Primary Residence," which is not the same as nation of origin.

sampling was used (Shadish, Cook, & Campbell, 2002) for the validation of the self-report instrument. Specifically, heterogeneity sampling was utilized, in which the most diverse sample possible was attained. Detailed demographic information for the sample ($n = 174$) is included in Table 1. Also included is a comparison of the study's participant demographics to that of the overall AEA membership, as reported in a survey of over 2,500 AEA members (AEA, 2008). As the table indicates, the demographics of study participants are very similar to those of the overall AEA membership.

Measures

After ensuring public use status or obtaining permission from authors, items from four self-report measures of cultural competence used in other fields (e.g., counseling, therapy, healthcare) were selected and altered (e.g., changing "counselor" or "nurse" to "program evaluator") to better suit the field of evaluation. Across the four measures there were similar items, so the researchers selected items with simple, brief wording over ones with complex, longer wording (e.g., "Program evaluation as a whole has failed to meet the needs of ethnic minorities" rather than "How would you react to the following statement? While program evaluation enshrines the concepts of freedom, rational thought, tolerance of new ideas, and equality, it has frequently become a form of oppression to subjugate large groups of people"). Similarly, the researchers selected items that were more targeted to

Table 2. Origin of Final Cultural Competence of Program Evaluators (CCPE) Questions

CCPE item	Origin	CCPE item	Origin
1	MAKSS #22	14	CAS #6
2	MAKSS #23	15	CAS #7
3	MAKSS #24	16	CAS #11[a]
4	MAKSS #27	17	MCCTS-R #1
5	MCCTS-R #11[a]	18	MCCTS-R #3
6	MCCTS-R #12[a]	19	MCCTS-R #4[a]
7	MAKSS #3	20	CAS #23[a]
8	MAKSS #4	21	MCCTS-R #20
9	MAKSS #8	22	MCCTS-R #30[a]
10	MAKSS #1	23	MAKSS #41[a]
11	MAKSS #7[a]	24	MAKSS #45
12	MAKSS #10	25	MAKSS #48
13	CAS #5	26	MAKSS #51[a]

[a]Slight wording change (e.g., "counselors" to "evaluators").

cultural competence as opposed to items that were more general (e.g., "How would you rate your ability to distinguish 'formal and informal' program evaluation strategies?"), given that the latter could be interpreted in a variety of ways by participants. Table 2 lists the origin of each final selected item. The items were combined, along with qualitative and demographic questions, to create the Cultural Competence of Program Evaluators (CCPE) self-report instrument. A description of each of the original four self-report measures follows.

Cultural awareness scale (CAS; Rew, Becker, Cookston, Khosropour, & Martinez, 2003). This 36-item instrument is designed to measure outcomes of a program to strengthen multicultural awareness among nursing faculty and students. The instrument consists of five subscales: (1) general educational experience, (2) cognitive awareness, (3) research issues, (4) behaviors/comfort with interactions, and (5) patient care/clinical issues. Alpha coefficients for the preceding subscales are .85, .79, .94, .71, and .77, respectively. Overall alpha for the CAS is .82. Items include, "When I have an opportunity to help someone, I offer assistance less frequently to individuals of certain cultural backgrounds" and "I respect the decisions of my patients when they are influenced by their culture, even if I disagree." All items are measured using a 7-point Likert-type scale (1 = *strongly disagree* to 7 = *strongly agree*).

The cultural competence self-assessment questionnaire[2] (CCSAQ; Mason, 1995). This 74-item measure is designed to assess the cultural competence of human services professionals. The instrument consists of three subscales: (1) knowledge of communities, which pertains to respondents' understanding of community dynamics, including racial composition, socioeconomic status (SES), support systems, and the cultural norms and values of people of color, (2) resources and linkages, which examines the availability of relevant information, materials, and resources for respondents' access and use, and (3) service delivery and practice, which examines respondents' understanding of appropriate treatment interventions, cultural strengths, historical accomplishments, family support systems, and methods of advocacy. Overall alpha for the CCSAQ is .80. Items include, "Do you know the social protocol within communities of color?" and "Do you feel safe within communities of color?" All items are measured using a 4-point Likert-type scale (1 = *not at all/not well/none/never* to 4 = *often/very well/many/regularly*).

The multicultural counseling awareness, knowledge, and skills survey (MAKSS; D'Andrea, Daniels, & Heck, 1991). This 60-item questionnaire measures the effectiveness of cultural competence training on counselors' cross-cultural awareness, knowledge, and skills. In previous research (D'Andrea et al., 1991), this scale has exhibited high reliability on the three subscales of awareness ($\alpha = .75$, 20 items), knowledge ($\alpha = .90$, 20 items), and skills ($\alpha = .96$, 20 items). Items include, "Ambiguity and stress often result from multicultural situations because people are not sure what to expect from each other" and "The human service professions, especially counseling and clinical psychology, have failed to meet the mental needs of ethnic minorities." All items utilize a 4-point Likert-type response scale (1 = *strongly disagree* to 4 = *strongly agree*), with higher scores indicating greater cultural competence.

The multicultural counseling and training survey–revised (MCCTS-R; Holcomb-McCoy, 2000). This 32-item survey is designed to measure the perceived multicultural competence of professional counselors. The instrument consists of three subscales: (1) multicultural knowledge, (2) multicultural awareness, and (3) multicultural terminology. Alpha coefficients for the preceding subscales are .95, .85, and .97, respectively. Some items include, "I nonverbally communicate my acceptance of culturally different students" and "I can discuss how culture affects the help-seeking behaviors of students." All items utilize a 4-point Likert-type response scale (1 = *not competent/not able to perform at this time* to 4 = *extremely competent/able to perform at a high level*), with higher scores representing higher levels of cultural competence.

Multicultural counseling inventory (MCI; Roysircar, 2004). This 40-item instrument was administered to participants in its entirety along with the CCPE in order to establish convergent validity of the new instrument. It measures multicultural counseling competencies and includes four subscales: multicultural counseling skills, multicultural awareness, multicultural counseling relationship, and multicultural counseling knowledge. Alpha coefficients for the preceding subscales were .77, .51, .75, and .72, respectively. Items include, "I perceive that my race causes the clients to mistrust me" and "I am able to quickly recognize and recover from cultural mistakes or misunderstandings." All items are measured using a 4-point Likert-type scale ranging from *very inaccurate* to *very accurate*. The MCI was chosen as the validity measure because it contains all subscales of the cultural competence paradigm (skills, awareness, and knowledge). Also, the primary author of the MCI would not allow individual items or subscales to be used or adapted for the development of other instruments, which made it impossible to use selected items from this scale in the creation of the CCPE. Table 3 contains a description of each measure's function in the study, as well as references.

Procedure

After the researchers altered the selected questions from the four already-established cultural competence measures to make them suitable for use with evaluators, the new cultural competence self-report measure was pretested using the Delphi technique of instrument creation (Colton & Covert, 2007). This was a way to obtain the opinion of experts without bringing them together face to face. Initially, the new measure was reviewed by two experts in survey research and program evaluation and then revised based on their feedback. The revised measure was then e-mailed to four experts in program evaluation and cultural competence whose disparate specializations ranged from diversity and race to survey design and research methods. These four individuals reviewed the measure independently for readability, sentence length, wording, clarity, response categories, cultural appropriateness, bias, and time frame/tense. The researchers again revised the instrument after receiving feedback from the reviewers.

Table 3. Function of Each Measure in the Current Study

Measure	Acronym	Use in Current Study	Reference
Cultural Awareness Scale	CAS	Selected items used to create the CCPE	Rew et al. (2003)
The Cultural Competence Self-Assessment Questionnaire	CCSAQ	Selected items used to create the CCPE	Mason (1995)
Multicultural Counseling Awareness, Knowledge, and Skills Survey	MAKSS	Selected items used to create the CCPE	D'Andrea et al. (1991)
The Multicultural Counseling and Training Survey–Revised	MCCTS-R	Selected items used to create the CCPE	Holcomb-McCoy (2000)
Multicultural Counseling Inventory	MCI	Administered to study participants in tandem with the CCPE	Sodowsky, Taffe, Gutkin, and Wise (1994)

After receiving Institutional review board (IRB) approval, the researchers collected data via online surveying of program evaluators during February and March 2009. An invitation for participation and link to the online survey were posted on the AEA listserv, known as EVALTALK, and e-mailed to participants in the 2008 AEA conference and members of the Southeast Evaluation Association (SEA).[3] EVALTALK is an online discussion forum available to all AEA members and consists of approximately 5,000 evaluators who work in either academia or in the industry. SEA, with approximately 150 members, is an organization specifically for evaluators in the southeastern United States.

Participants who completed the survey had the option of completing a separate form containing their name and primary e-mail address if they wanted to enter a raffle to win one of ten $20 gift cards. The database for this information was kept separate from the survey database to maintain anonymity of the participants. Of the 174 participants who completed the survey, 95 (54.60%) entered their names and e-mail addresses into the separate form. The researchers randomly chose 10 gift card recipients from this pool of 95 names.

Results

The data were cleaned before any inferential analyses were conducted. Unexpectedly, every item from the MCI had missing data, ranging from 9.7% to 44.6%. The researchers decided not to replace the missing values for cases in which more than 15 MCI items were missing (thus excluding 23.5% of cases), as this would have falsely reduced the variability of these items (Tabachnick & Fidell, 2001). Instead, these items were left missing and were excluded from the subsequent correlational analysis. For cases with 15 or fewer missing MCI items, missing values were replaced with the group mean for years of evaluation experience (which was a continuous variable with values ranging from 1 to 40) for that item. For example, a missing value on MCI item #16 for a participant with 11 years of evaluation experience would be replaced with the mean value on MCI item #16 of other participants with 11 years of evaluation experience. Subsequently, missing data for the 40 MCI items ranged from 8.6% to 17.1%.

Factor Solution and Reliability of CCPE

Several principal components analyses (PCA) were conducted on the continuous variables contained in the CCPE, and internal consistency (i.e., Cronbach's α) and convergent validity were assessed. Measures of sampling adequacy revealed no issues with the factorability of the correlation matrix.

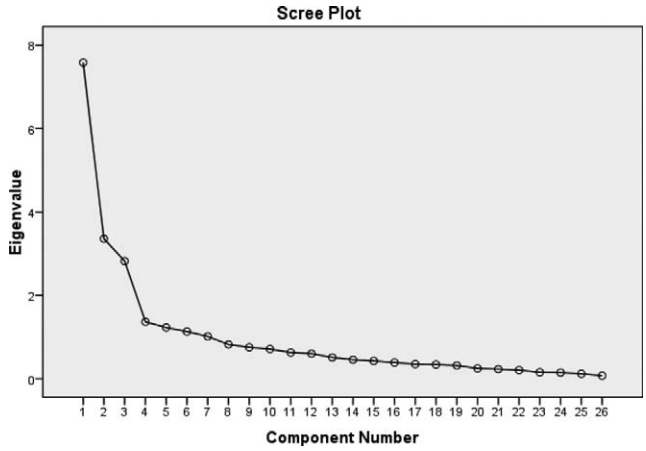

Figure 2. Scree plot of the unrotated factors.

Bartlett's test of sphericity was significant, $\chi^2(325) = 2341.45$, $p < .001$, and the Kaiser–Meyer–Olkin measure of sampling adequacy was .83, which is considered excellent (Pett, Lackey, & Sullivan, 2003).

Examination of the Scree plot (Figure 2) suggested a 3- to 5-component solution. These three solutions were tested, and the 4 and 5 component solutions contained factors with fewer than 3 items. Therefore, the researchers selected a three-component solution with varimax rotation. Varimax rotation was chosen because it aids interpretation when the components are to be used as dependent variables (Tabachnick & Fidell, 2001). Items that did not adequately load ($< |.30|$) on any of the components were deleted. The remaining items again underwent PCA, and items with ambiguous loadings (i.e., those that loaded on more than one component with values displaying less than a .200 difference) were deleted individually. This procedure was repeated until there were 26 items[4] that loaded at least .35 on one of the components, with no ambiguous loadings. Internal reliability was then calculated for each component.

The first component, *cultural skills* ($\alpha = .89$), contained 14 items and accounted for 24.59% of the variance. Reflected in items such as "How well would you rate your ability to analyze a culture into its component parts?" and "I can discuss program evaluation from a cultural/ethnic/racial perspective," this component revealed participants' feelings and behaviors when interacting with persons from different cultures. Items were measured on a 5-point Likert-type scale ranging from *very limited* or *strongly disagree* to *very good* or *strongly agree*, depending on question wording.

The second component, *cultural knowledge* ($\alpha = .80$), contained 7 items and accounted for 17.06% of the variance. This component included items that represented participants' perceptions of their knowledge regarding various culture-related terms. Items were measured on a 5-point Likert-type scale ranging from *very limited* to *very good*. Items included "What is your current knowledge of the term racism?" and "What is your current understanding of the term ethnocentrism?"

The third component, *cultural awareness* ($\alpha = .73$), included 5 items and accounted for 11.28% of the variance. Items represented participants' personal as well as global awareness of the role of culture and were measured on a 5-point Likert-type scale ranging from *strongly disagree* to *strongly agree*. Sample items included "Ambiguity and stress often result from multicultural situations because people are not sure what to expect from each other" and "I think my behaviors are influenced by my culture."

Overall, the rotated three-component solution of the CCPE accounted for 52.93% of the variance. Values greater than 50% are considered good (Tabachnick & Fidell, 2001). The entire scale of 26

Table 4. Correlations of CCPE Factors and MCI Factors

	MCI 1	MCI 2	MCI 3	MCI 4	Total MCI
CCPE 1	.45*** ($n = 137$)	.58*** ($n = 139$)	.47*** ($n = 140$)	.61*** ($n = 139$)	.67*** ($n = 134$)
CCPE 2	.37*** ($n = 141$)	.36*** ($n = 143$)	.36*** ($n = 144$)	.36*** ($n = 143$)	.46*** ($n = 138$)
CCPE 3	.17* ($n = 144$)	.14 ($n = 146$)	−.10 ($n = 147$)	.18* ($n = 146$)	.15 ($n = 141$)
Total CCPE	.50*** ($n = 133$)	.56*** ($n = 135$)	.46*** ($n = 136$)	.61*** ($n = 135$)	.68*** ($n = 130$)

Note: CCPE 1 = cultural skills; CCPE 2 = cultural knowledge; CCPE 3 = cultural awareness; MCI 1 = multicultural counseling skills; MCI 2 = multicultural awareness; MCI 3 = multicultural counseling relationship; MCI 4 = multicultural counseling knowledge.
*$p < .05$. ***$p < .001$.

items had an internal consistency of .88. Convergent validity of the CCPE was established by comparing the three components and total score of the CCPE to the four components and total score of the MCI via Pearson *r* correlations. Sample sizes for these correlations ranged from 130 to 146. The cultural skills and cultural knowledge subscales were significantly positively correlated ($p < .001$) with all four MCI subscales (skills, awareness, counseling relationship, and counseling knowledge) as well as the total score of the MCI. The cultural awareness subscale was significantly positively correlated ($p < .05$) with the MCI skills and counseling knowledge subscales. Additionally, the total CCPE score was significantly positively correlated ($p < .001$) with all four MCI components as well as the total score of the MCI. See Table 4 for correlations.

The PCA, along with additional qualitative questions,[5] resulted in the final version of the CCPE self-report scale (Appendix A).[6] This self-report instrument consists of five qualitative questions that probe participants' perceptions of qualities possessed by a culturally competent program evaluator and 26 questions that constitute the three subscales: cultural skills (14 items), cultural knowledge (7 items), and cultural awareness (5 items). Alpha coefficients for the preceding subscales were .89, .80, and .73, respectively. The overall alpha of the quantitative CCPE items was .88. All items are measured using a 5-point response scale (1 = *very limited/not aware/strongly disagree* to 5 = *very good/very aware/strongly agree*). The instrument also includes nine demographic questions (e.g., age, race, sex, highest level of education, years of experience in program evaluation).

After validating the CCPE, group differences based on demographic variables[7] were assessed. First, the assumptions of homoscedasticity, normality, linearity, and independence were checked in the multivariate analysis of variance (MANCOVA) and regression models. No violations were found. For the following analyses, the dependent variables were the three factors of the CCPE (cultural skills, cultural knowledge, and cultural awareness).

Influence of Years of Evaluation Experience on Self-Reported Cultural Competence (RQ 1)

To ascertain whether individuals with more years of evaluation experience had higher self-reported levels of cultural competence, a MANCOVA was conducted. The independent variable was years of evaluation experience (less than 5 years, 5–10 years, 11–15 years, 16–20 years, 21–25 years, over 25 years). The overall MANCOVA was nonsignificant, $F(3, 147) = 1.91$, *ns*.

Influence of Gender and Training on Self-Reported Cultural Competence (RQs 2 and 5)

To assess the impact of gender and receipt of cultural competence training on self-reported level of cultural competence, a 2 × 2 factorial MANCOVA was conducted. Gender (male or female) and receipt of cultural competence training as defined by "yes" or "no" response concerning completion of course/courses for credit during the graduate program were the independent variables. The interaction MANCOVA for gender and training was nonsignificant, $F(3, 144) = .80$, *ns*, as was the

overall MANCOVA for the gender main effect, $F(3, 144) = 1.29$, *ns*. However, the overall MANCOVA for the training main effect was significant, $F(3, 144) = 4.17$, $p < .01$, partial $\eta^2 = .08$. The univariate analyses of variance (ANOVAs) for the cultural skills and cultural knowledge subscales were significant. Specifically, individuals who had received cultural competence training ($M = 58.40$, $SD = 7.66$) self-reported significantly higher scores on the cultural skills subscale than individuals who had not received cultural competence training ($M = 54.09$, $SD = 8.20$), $F(1, 146) = 11.12$, $p < .01$, partial $\eta^2 = .07$. Additionally, individuals who had received cultural competence training ($M = 29.75$, $SD = 3.65$) self-reported significantly higher scores on the cultural knowledge subscale than individuals who had not received cultural competence training ($M = 27.95$, $SD = 4.47$), $F(1, 146) = 5.41$, $p < .05$, partial $\eta^2 = .04$.

Influence of Minority Status and Age on Self-Reported Cultural Competence (RQs 3 and 4)

To assess differences in self-reported levels of cultural competence based on age and minority status, a 5 × 2 factorial MANCOVA was conducted. The independent variables were age (30 years old or younger, 31–40 years old, 41–50 years old, 51–60 years old, over 60 years old) and minority status (minority, nonminority).[8] The MANCOVA for the interaction of age and minority status was nonsignificant, $F(12, 426) = .49$, *ns.* Additionally, the overall MANCOVA for age was nonsignificant, $F(12, 426) = .30$, *ns*, as was the overall MANCOVA for minority status, $F(3, 140) = .91$, *ns.*

Determining the Best Predictor of Cultural Competence (RQ 6)

To assess which demographic variable would best predict cultural competence, three standard multiple regressions were conducted. The predictor variables[9] included years of experience (less than 5 years as reference group), gender (males as reference group), age (30 years old or younger as reference group), minority status (nonminority as reference group), and receipt of formal cultural competence training (''no'' as reference group), and the criterion variables were each subscale of the CCPE (cultural skills, cultural knowledge, and cultural awareness). Initially, to test for the absence of multicollinearity among the predictor variables, Pearson *r* correlations were conducted. Age and years of evaluation experience correlated above |.6|, but this is to be expected based on the nature of these variables. Multicollinearity was not present for any other variables.

As shown in Table 5, the overall multiple regression for the cultural skills subscale was statistically significant, $F(5, 147) = 3.00$, $p < .05$, $R = .30$, Adj. $R^2 = .06$. Receipt of cultural competence training ($\beta = .26$, $sr_i^2 = .06$) was a significant predictor of this subscale, with individuals who had received cultural competence training obtaining higher scores on the cultural skills subscale.

A multiple regression analysis was performed for the cultural knowledge subscale (Table 5); the regression was nonsignificant, $F(5, 152) = 1.68$, *ns*, $R = .25$, Adj. $R^2 = .03$. However, receipt of cultural competence training ($\beta = .21$, $sr_i^2 = .04$) was a significant predictor of this subscale, with individuals who had received cultural competence training obtaining higher scores on the cultural knowledge subscale.

A final multiple regression analysis was performed for the cultural awareness subscale (Table 5); this regression was statistically significant, $F(5, 156) = 3.09$, $p < .01$, $R = .33$, Adj. $R^2 = .07$. Receipt of cultural competence training ($\beta = .15$, $sr_i^2 = .02$) was related to the score on this subscale. Individuals who had received cultural competence training obtained higher scores on the cultural awareness subscale.

Table 5. The Effect of Demographic Variables on Subscales of CCPE

Variable	*B*	β	sr_i^2
Cultural skills subscale			
Years of experience	.04	.01	.00
Gender	−.41	−.02	.00
Age	.72	.10	.01
Minority status	2.55	.13	.02
Training	4.33	.26**	.06
Cultural knowledge subscale			
Years of experience	.06	.02	.00
Gender	.88	.09	.01
Age	.01	.00	.00
Minority status	−.43	−.04	.00
Training	1.83	.21**	.04
Cultural awareness subscale			
Years of experience	−.22	−.11	.01
Gender	.03	.00	.00
Age	−.07	−.03	.00
Minority status	.52	.07	.00
Training	.92	.15*	.02

Note: Cultural skills subscale: $R = .30$ and Adj. $R^2 = .06$ ($N = 152$); Cultural knowledge subscale: $R = .25$ and Adj. $R^2 = .03$ ($N = 158$); Cultural awareness subscale: $R = .33$ and Adj. $R^2 = .07$ ($N = 162$).
*$p < .05$. **$p < .01$.

Discussion

The focus of this study was to develop a measure of self-reported cultural competence for use with program evaluators, as well as to examine possible differences in level of cultural competence based on various demographic factors. The overall goal was to construct a measure, the CCPE self-report scale, that would be reliable and valid. Several research questions were also posed concerning group differences based on demographics such as gender, minority status, age, years of evaluation experience, and receipt of cultural competence training.

The CCPE did exhibit moderate to high (>.70) reliability and validity. Reliability was assessed via PCA and internal consistency analyses, which reduced the original 52 Likert-type scale items to 26 items that accounted for approximately 53% of the variance. The final self-report measure consists of three subscales: cultural skills (14 items), cultural knowledge (7 items), and cultural awareness (5 items). Each subscale had an internal consistency of at least .70, and the overall self-report measure had an internal consistency of .88. Convergent validity was assessed via Pearson *r* correlations between scores on the CCPE and MCI. Because there are no other cultural competence self-report measures for use with program evaluators, we decided that the MCI was the best option. Two of the three CCPE subscales were significantly positively correlated with all subscales of the MCI, as well as with the overall MCI score. Although the cultural awareness subscale was not significantly positively correlated with all subscales of the MCI, it was significantly positively correlated with two of them. The relatively weak results for the awareness subscale may be due to its items not representing cultural awareness as fully as the items in the other two subscales capture cultural skills and knowledge. Overall, however, the total CCPE score was significantly positively correlated with the total score of the MCI, suggesting convergent validity of the new self-report measure.

The first research question examined whether individuals with more years of evaluation experience would have higher self-reported levels of cultural competence as measured by the CCPE. Data revealed no significant difference in level of cultural competence for any of the CCPE subscales or

the total CCPE score based on years of experience. This result supports other literature finding work experience per se does not necessarily provide cultural competence (Hansen et al., 2000).

Research questions 2–4 examined if there was a difference in self-reported level of cultural competence as measured by the CCPE based on gender, minority status, or age. Data revealed that none of these demographic variables were associated with differences in self-reported level of cultural competence. These results suggest that males and females, minorities and nonminorities, and people of all ages may have similar self-reported levels of such competence. However, due to the relatively small sample of minority participants, the results regarding the lack of minority-status differences in self-reported cultural competence should be interpreted with extreme caution.

Past research highlights the importance of cultural competence training for *all* evaluators, regardless of demographics, rather than just the majority group (Abernethy, 2005; Ladson, Lin, Flores, & Magrane, 2006). These preliminary results mirror past studies and indicate that the CCPE is tapping into perceived cultural competence rather than differences based on participants' innate characteristics.

Research question 5 examined whether individuals who had received formal cultural competence training would have higher self-reported levels of cultural competence as measured by the CCPE. Data showed that individuals who had received such training scored significantly higher on all three subscales of the CCPE. Although there are no current standards or consensus on the core objectives and competencies that should be achieved through cultural competence training, there is general agreement among experts that cultural awareness, knowledge, and skills are needed in order to deliver high-quality service to diverse populations (Betancourt, 2003; Hansen et al., 2000; Ladson et al., 2006; Roberson, Kulik, & Pepper, 2002). The need for such trainings for program evaluators has been reinforced by the results of this study.

The final research question assessed the best demographic predictor/predictors of cultural competence. Characteristics such as years of evaluation experience, gender, age, minority status, and receipt of cultural competence training were included as predictors. Training was the only significant predictor for the cultural knowledge, cultural skills, and cultural awareness subscales of the CCPE, supporting claims that there is a need within the field for cultural competence training for all program evaluators. Training positively impacts self-reported level of cultural competence, which in turn has the potential to positively impact evaluations.

Implications

In April 2011, AEA released its Statement on Cultural Competence in Evaluation. The statement affirms AEA's stance on the importance of cultural competence and asserts that "cultural competence in evaluation theory and practice is critical for the profession" and is "integral to ethical, high-quality evaluation" (AEA, 2011, p. 1). If cultural competence is critical for the profession, then training for cultural competence is, as well. But how can we know if trainings are increasing cultural competence? The CCPE can help answer that question.

One's self-reported level of cultural competence, as measured by the CCPE, was predicted by training. The CCPE, which is primarily intended for use as a training tool, could be employed to assess if evaluators in fact exhibit increased awareness of self, increased reflection on their own cultural position, increased awareness of others' positions, and increased ability to interact genuinely and respectfully with others (AEA, 2011, p. 3) as a result of training.

Limitations and Suggestions for Future Research

Limitations of the study fall into two broad categories: the sample and the self-report measure. Limitations of the sample included size and lack of diversity. Specifically, the researchers had hoped to attain approximately 350 participants, but only 174 completed the online survey, which limits

generalizability. Several methods of recruitment were utilized that reached approximately 3,000 individuals. Thus, it is fairly safe to estimate a 5% response rate, which is disappointing. Additionally, this sample lacked diversity in ethnicity, gender, and nationality in that approximately 82% of participants were White, 75% were female, and 74% were from the United States.

The limited ethnic diversity of the sample also hindered the ability to effectively assess differences in scores on the CCPE based on ethnicity. Namely, due to the small number of minority participants that represented only 19% of the sample, the researchers had to collapse ethnicity into two categories for statistical analyses. Having a dichotomous ethnicity variable rather than one encompassing multiple ethnicities obviously limited the information garnered from the data. To avoid these limitations, future studies should attempt to obtain a larger and more diverse sample of program evaluators, ensuring that there are meaningful numbers of participants in various ethnic groups. These studies could also assess differences on the CCPE based on nationality, perhaps comparing participants from the United States to those not from the United States.

The limitations of utilizing a self-report measure are well known. Namely, self-report surveys are subject to participants' interpretations of items as well as to issues of social desirability (Colton & Covert, 2007). Participants that completed the CCPE could have chosen responses that made them look good to others or that reflected their own biased perceptions of themselves. In an effort to minimize this risk, the researchers included participants' scores on the Crowne-Marlowe Social Desirability Scale–short version as a covariate in all statistical analyses and found no significant relationships between scores on the social desirability measure and the dependent variables. However, recent studies suggest that the Crowne-Marlowe Social Desirability Scale is not a strong predictor of social bias (Johnson & Fendrich, 2002), so the results of this study should be interpreted with caution. Ways to assess cultural competence in a non-self-report manner are limited and could include conducting extensive observations of participants' work as evaluators in the field or asking evaluators to respond to case-study scenarios regarding how they would handle various culturally related evaluation tasks and challenges.

Conclusions

Despite its limitations, this study represents a significant step in establishing the concrete validity of the CCPE. The CCPE demonstrated appropriate psychometric properties, exhibiting both reliability and convergent validity, and it differentiated participants who had received cultural competence training from those who had not. The CCPE also fills a gap in research in that no such self-report measure currently exists in program evaluation. Of course, additional studies documenting the worth of the measure will need to be undertaken before its validity should be accepted by professionals. Cultural competence in program evaluation is vital, so evaluators' self-reported as well as directly measured level of such competence may have important implications for the quality and validity of their evaluation findings. This is an avenue that warrants further exploration.

Appendix A

Cultural Competence of Program Evaluators (CCPE) Self-Report Scale

Please circle the number that most accurately reflects your current understanding of the following terms.

1. Ethnicity

Very limited				Very good
1	2	3	4	5

2. Racism

Very limited				Very good
1	2	3	4	5

3. Prejudice

Very limited				Very good
1	2	3	4	5

4. Ethnocentrism

Very limited				Very good
1	2	3	4	5

5. Discrimination

Very limited				Very good
1	2	3	4	5

6. Stereotype

Very limited				Very good
1	2	3	4	5

Please read the statements below and circle the number that most accurately reflects your perceptions or behavior. Answer to the best of your ability. Please keep in mind that there is no way to perform poorly.

7. At this time in your life, how would you rate yourself in terms of understanding how your cultural background has influenced the way you think and act?

Very limited				Very aware
1	2	3	4	5

8. At this point in your life, how would you rate your understanding of the impact of the way you think and act when interacting with persons of different cultural backgrounds?

Very limited				Very aware
1	2	3	4	5

9. At the present time, how would you generally rate yourself in terms of being able to accurately compare your own cultural perspective with that of a person from another culture?

Very limited				Very aware
1	2	3	4	5

10. Culture is not external but is within the person.

Strongly disagree				Strongly agree
1	2	3	4	5

11. Program evaluation as a whole has failed to meet the needs of racial/ethnic/cultural minorities.

Strongly disagree 1 2 3 4 5 Strongly agree

12. Ambiguity and stress often result from multicultural situations because people are not sure what to expect from each other.

Strongly disagree 1 2 3 4 5 Strongly agree

13. I think my beliefs and attitudes are influenced by my culture.

Strongly disagree 1 2 3 4 5 Strongly agree

14. I think my behaviors are influenced by my culture.

Strongly disagree 1 2 3 4 5 Strongly agree

15. I often reflect on how culture affects beliefs, attitudes, and behaviors.

Strongly disagree 1 2 3 4 5 Strongly agree

16. I believe program evaluators' own cultural beliefs influence their evaluation decisions.

Strongly disagree 1 2 3 4 5 Strongly agree

17. I can discuss my own ethnic/cultural heritage.

Strongly disagree 1 2 3 4 5 Strongly agree

18. I am able to discuss how my culture has influenced the way I think.

Strongly disagree 1 2 3 4 5 Strongly agree

19. I can recognize when my attitudes, beliefs, and values are interfering with providing the best services to those being evaluated.

Strongly disagree 1 2 3 4 5 Strongly agree

Please read the statements below and choose the number that most accurately reflects your perceived level of proficiency in performing the following tasks.

20. I feel comfortable discussing cultural issues.

Strongly disagree 1 2 3 4 5 Strongly agree

21. I can discuss within-group differences among ethnic groups (e.g., low SES Puerto Rican vs. high SES Puerto Rican).

Strongly disagree 1 2 3 4 5 Strongly agree

22. I can discuss program evaluation from a cultural/ethnic/racial perspective.

Strongly disagree 1 2 3 4 5 Strongly agree

23. How would you rate your ability to *conduct an effective evaluation* involving persons from a cultural background significantly different from your own?

Very limited 1 2 3 4 5 Very good

24. How well would you rate your ability to accurately identify culturally biased assumptions as they relate to your professional training?

Very limited 1 2 3 4 5 Very good

25. How well would you rate your ability to analyze a culture into its component parts?

Very limited 1 2 3 4 5 Very good

26. In general, how would you rate your skill level in terms of being able to provide appropriate evaluation services to culturally different individuals?

Very limited 1 2 3 4 5 Very good

Authors' Note

This research took place at Old Dominion University, Norfolk, VA.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Formal cultural competence training was described as completion of any graduate-level course/courses concerning cultural competence that were completed for credit toward a graduate degree.
2. Items from the CCSAQ were included in the survey that participants completed. However, none of these items were included in the final measure due to their being eliminated during principal components analyses.
3. Members from SEA were selected for participation due to accessibility of e-mail lists.
4. PCAs reduced the number of items from 52 to 26.
5. The qualitative questions, while part of the CCPE, do not contribute to the quantitative validation of items discussed in this article. The qualitative data will be discussed in a future publication.
6. Individuals wishing to use the CCPE self-report scale must obtain written permission from the primary author.
7. Social desirability, measured via the Marlowe-Crowne Social Desirability Scale–short version (Crowne & Marlowe, 1964), was included as a covariate in all inferential statistics. It was not statistically significant in any analysis.
8. The breakdown of ethnicities in the sample was 4 (2.3%) American Indian or Alaskan Native, 11 (6.3%) Black or African American, 1 (0.6%) Native Hawaiian or other Pacific Islander, 142 (81.6%) White, and 16 (9.2%) Other. Due to unequal sample sizes across ethnic groups, the ethnicity variable was collapsed into two categories (minority and nonminority).
9. Multiple regressions with interactions were also performed on all variables for each model. None of the interactions were significant predictors.

References

Abernethy, A. (2005). Increasing the cultural proficiency of clinical managers. *Journal of Multicultural Counseling and Development*, *33*, 81–93.

Abrums, M., & Leppa, C. (2001). Beyond cultural competence: Teaching about race, gender, class, and sexual orientation. *Journal of Nursing Education*, *40*, 270–275.

Alkon, A., Tschann, J., Ruane, S., Wolff, M., & Hittner, A. (2001). A violence-prevention and evaluation project with ethnically diverse populations. *American Journal of Preventive Medicine*, *20*, 48–55.

American Evaluation Association. (2004). *Guiding principles for evaluators*. Retrieved June 20, 2009, from http://www.eval.org/Publications/GuidingPrinciplesPrintable.asp

American Evaluation Association. (2008). *American Evaluation Association internal scan report to the membership*, by Goodman Research Group. Retrieved May 22, 2009, from www.eval.org

American Evaluation Association. (2011). *American Evaluation Association public statement on cultural competence in evaluation*. Fairhaven, MA: Author. Retrieved August 15, 2011, from www.eval.org

Barrett, K. H., & George, W. H. (2005). *Race, culture, psychology, and law*. Thousand Oaks, CA: Sage.

Beagan, B. (2003). Teaching social and cultural awareness to medical students: "It's all very nice to talk about it in theory, but ultimately it makes no difference." *Academic Medicine*, *78*, 605–614.

Benavente, V. G. (2004). Cultural and linguistic competency assessment. *Nurse Leader*, *2*, 24–30.

Betancourt, J. (2003). Cross-cultural medical education: Conceptual approaches and frameworks for evaluation. *Academic Medicine*, *78*, 560–569.

Campinha-Bacote, J. (2002). The process of cultural competence in the delivery of healthcare services: A model of care. *Journal of Transcultural Nursing*, *13*, 181–184.

Carter, R. T. (2003). Becoming racially and culturally competent: The racial-cultural counseling laboratory. *Journal of Multicultural Counseling and Development*, *31*, 20–30.

Colton, D., & Covert, R. W. (2007). *Designing and constructing instruments for social research and evaluation*. San Francisco, CA: Jossey-Bass.

Crowne, D. P., & Marlowe, D. (1964). *The approval motive*. New York, NY: Wiley.

D'Andrea, M., Daniels, J., & Heck, R. (1991). Evaluating the impact of multicultural counseling training. *Journal of Counseling and Development*, *70*, 143–150.

Doutrich, D., & Storey, M. (2004). Education and practice: Dynamic partners for improving cultural competence in public health. *Family Community Health*, *27*, 298–307.

Dumas, J. E., Rollock, D., Prinz, R. J., Hops, H., & Blechman, E. A. (1999). Cultural sensitivity: Problems and solutions in applied and preventive intervention. *Applied and Preventive Psychology*, *8*, 175–196.

Frechtling, J. (2002). *The 2002 user friendly handbook for project evaluation. The National Science Foundation, Division Of Research, Evaluation, And Communication*. Arlington, VA: Westat.

Greene, J. C. (1997). Evaluation as advocacy. *Evaluation Practice*, *18*, 25–35.

Guzman, B. L. (2003). Examining the role of cultural competency in program evaluation: Visions for new millennium evaluators. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problem: Visions for the new millennium* (pp. 167–181). Hillsdale, NJ: Lawrence Erlbaum.

Hansen, N. D., Pepitone-Arreola-Rockwell, F., & Greene, A. F. (2000). Multicultural competence: Criteria and case examples. *Professional Psychology: Research and Practice*, *31*, 652–660.

Holcomb-McCoy, C. C. (2000). Multicultural counseling competencies: An exploratory factor analysis. *Journal of Multicultural Counseling & Development*, *28*, 83–97.

Hood, D. W., & Cassaro, D. A. (2002). Feminist evaluation and the inclusion of difference. *New Directions for Evaluation*, *96*, 27–40.

Howard, D. L. (2002). Culturally competent treatment of African American clients among a national sample of outpatient substance abuse treatment units. *Journal of Substance Abuse Treatment*, *24*, 89–102.

Johnson, T. P., & Fendrich, M. (2002). *A validation of the Crowne-Marlowe social desirability scale*. Retrieved August 16, 2011, from http://www.srl.uic.edu/publist/Conference/crownemarlowe.pdf

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

Ladson, G. M., Lin, J., Flores, A., & Magrane, D. (2006). An assessment of cultural competence of first- and second-year medical students at a historically diverse medical school. *American Journal of Obstetrics and Gynecology*, *195*, 1457–1462.

Mason, J. L. (1995). *Cultural competence self-assessment questionnaire: A manual for users*. Portland, OR: Portland State University, Research and Training Center on Family Support and Children's Mental Health.

McPhatter, A. R., & Ganaway, T. L. (2003). Beyond the rhetoric: Strategies for implementing culturally effective practice with children, families, and communities. *Child Welfare*, *82*, 103–124.

Mendias, E., & Guevara, E. (2001). Assessing culturally competent scholarship. *Journal of Professional Nursing*, *17*, 256–266.

Okazaki, S., & Sue, S. (1995). Methodological issues in assessment research with ethnic minorities. *Psychological Assessment*, *7*, 367–375.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Pope, R. L., & Reynolds, A. L. (1997). Student affairs core competencies: Integrating multicultural awareness, knowledge, and skills. *Journal of College Student Development*, *38*, 266–277.

Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research, and Practice*, *20*, 161–173.

Rew, L., Becker, H., Cookston, J., Khosropour, S., & Martinez, S. (2003). Measuring cultural awareness in nursing students. *Journal of Nursing Education*, *42*, 249–257.

Roberson, L., Kulik, C., & Pepper, M. (2002). Assessing instructor cultural competence in the classroom: An instrument and a development process. *Journal of Management Education*, *26*, 40–55.

Roysircar, G. (2004). Cultural self-awareness assessment: Practice examples from psychology training. *Professional Psychology: Research and Practice*, *35*, 658–666.

SenGupta, S., Hopson, R., & Thompson-Robinson, M. (2004). Cultural competence in evaluation: An overview. *New Directions for Evaluation*, *102*, 5–19.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.

Sodowsky, G. R., Taffe, R. C., Gutkin, T. B., & Wise, S. L. (1994). Development of the multicultural counseling inventory: A self-report measure of multicultural competencies. *Journal of Counseling Psychology*, *41*, 137–148.

Stanhope, V., Solomon, P., Pernell-Arnold, A., Sands, R. G., & Bourjolly, J. N. (2005). Evaluating cultural competence among behavioral health professionals. *Psychiatric Rehabilitation Journal*, *28*, 225–233.

Sue, D. W., Arredondo, P., & McDavis, R. J. (1992). Multicultural counseling competencies and standards: A call to the profession. *Journal of Counseling and Development*, *70*, 477–486.

Sue, D., Bingham, R., Porche-Burke, L., & Vasquez, M. (1999). The diversification of psychology: A multicultural revolution. *American Psychologist*, *54*, 1061–1069.

Symonette, H. (2004). Walking pathways towards becoming a culturally competent evaluator: Boundaries, borderlands and border-crossings. *New Directions in Evaluation*, *102*, 95–109.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Taket, A., & White, L. (1997). Working with heterogeneity: A pluralist strategy for evaluation. *Systems Research and Behavioral Science*, *14*, 101–111.

U.S. Census Bureau. (2007). *Question and answer center*. Retrieved February 16, 2008, from https://ask.census.gov

Wear, D. (2003). Insurgent multiculturalism: Rethinking how and why we teach culture in medical education. *Academic Medicine*, *78*, 549–554.