Journal of Counseling Psychology
2003, Vol. 50, No. 2, 221–234

0022-0167/03/$12.00 DOI: 10.1037/0022-0167.50.2.221

The Scale of Ethnocultural Empathy: Development, Validation, and Reliability

Yu-Wei Wang, M. Meghan Davidson, Oksana F. Yakushko, Holly Bielstein Savoy, Jeffrey A. Tan, and Joseph K. Bleier
University of Missouri—Columbia

The authors present 3 studies on the initial validity and reliability of the Scale of Ethnocultural Empathy (SEE), a self-report instrument that measures empathy toward people of racial and ethnic backgrounds different from one's own. Specifically, the exploratory factor analysis yielded 4 factors: Empathic Feeling and Expression, Empathic Perspective Taking, Acceptance of Cultural Differences, and Empathic Awareness. A confirmatory factor analysis provided evidence for the stability and generalizability of this 4-factor solution. The SEE was correlated in the predicted directions with general empathy and attitudes toward people's similarities and differences. High internal consistency and test–retest reliability estimates were also found across the three studies. A discussion of how this scale can add to the research and practice regarding empathy and multicultural issues is provided.

In the past decade, multicultural and diversity issues have become increasingly discussed inside and outside of the field of counseling psychology. With respect to race and ethnicity, the population of the United States is becoming more diverse; in 2000, 29% of all Americans were racial and ethnic minority individuals (U.S. Census Bureau, 2001). Population projections now indicate that by the year 2015, racial and ethnic minorities will comprise one third of all Americans (U.S. Census Bureau, 2000) and that by the year 2050, that figure will increase to nearly half (47%) of the U.S. population (U.S. Census Bureau, 2000).

Given this diversification, as well as the reoccurring conflicts among racial and ethnic groups, from the beating of Rodney King to the hate crimes that resulted from the tragedies of September 11, 2001, more attention should be placed on people's attitudes toward various racial and ethnic groups. More than tolerance is needed to adapt and evolve with these growing diversity and societal changes. It has been suggested that what is needed is understanding, awareness, and acceptance of individuals from different racial and ethnic backgrounds (Sue & Sue, 1999). Developing and drawing on ethnocultural empathy (i.e., empathy directed toward people from racial and ethnic cultural groups who are different from one's own ethnocultural group) has been suggested by scholars as a promising way to promote the mutual understanding between various racial and ethnic groups, on both cognitive and affective levels (e.g., Batson et al., 1997; Carell, 1997; Cui, van den Berg, & Jiang, 1998). In this article, we briefly describe the existing literature regarding the constructs of general empathy and ethnocultural empathy and outline the development of a measure of ethnocultural empathy through three studies. Implications for the use of this new instrument are then discussed.

General Empathy

Empathy is defined as "feeling in oneself the feelings of others" (Strayer & Eisenberg, 1987, p. 391). From psychoanalysts such as Kohut to humanists such as Rogers, empathy has been a key concept in understanding why and how therapy works (e.g., Lichtenberg, Bornstein, & Silver, 1984; Strayer, 1987). Empathy has also been discussed in various subdisciplines of psychology and has been studied as a determinant of altruism, attribution, and social judgment (e.g., Batson, Batson, et al., 1995; Dovidio, Allen, & Schroeder, 1990; Houston, 1990; Omdahl, 1995; Unger & Thumuluri, 1997). Such a broad interest in empathy from various disciplines supports the claim that the ability to empathize with others is critical to all human relationships (Omdahl, 1995) and "an essential constituent" (Kohut, 1959, as cited in Duan & Hill, 1996, p. 262) of all psychological phenomena.

In their review of empathy research, Duan and Hill (1996) identified research regarding empathy as a situation-specific, cognitive–affective state. From this perspective, empathy is commonly defined as responding vicariously to a stimulus or a stimulus person (Scotland, 1969) or sensing another's private world as if it were one's own (Rogers, 1959; Traux & Carkhuff, 1967). Duan and Hill (1996) also identified several theories that refer to empathy as a personality trait or general ability. Such theories conceptualize empathy as a trait or an ability to "know another person's inner experience" (Buie, 1981, as cited in Duan & Hill, 1996, p. 262) or "feel (perceive) the feelings (emotions) of other

Yu-Wei Wang, M. Meghan Davidson, Oksana F. Yakushko, Holly Bielstein Savoy, Jeffrey A. Tan, and Joseph K. Bleier, Department of Educational and Counseling Psychology, University of Missouri—Columbia.

Results from Study 1 were previously presented in August 2000 at the 108th Annual Convention of the American Psychological Association, Washington, DC. Results from Study 2 were previously presented in March 2001 at the National Counseling Psychology Conference, Houston, Texas. We thank the following people for their support and contributions to the development of this article: Madonna Constantine, Mary Heppner, Puncky Heppner, Dong-gwi Lee, Brent Mallinckrodt, Helen Neville, and Roger Worthington.

Correspondence concerning this article should be addressed to Yu-Wei Wang, Department of Educational and Counseling Psychology, University of Missouri, 16 Hill Hall, Columbia, Missouri 65211. E-mail: yw89@yahoo.com

people" (Sawyer, 1975, as cited in Duan & Hill, 1996, p. 262). The core assumption is that, either by nature or through development, some individuals are more empathic than others.

Much of the current research on empathy has focused on the construct as being one that can be altered in individuals (e.g., Batson et al., 1997; Batson, Turk, Shaw, & Klein, 1995). Empathy has been viewed as an ability that can be manipulated so as to lead to increased valuing of other people's welfare as well as a source of attitude change toward people from groups that experience oppression (Batson et al., 1997; Batson, Turk, et al., 1995). Additionally, research has demonstrated that a lack of empathy is linked to both intergroup aggression (Struch & Schwartz, 1989) and social dominance orientation (Pratto, Sidanius, Stallworth, & Malle, 1994), whereas Hoffman (2000) suggested that empathy is one of the key antecedents and corollaries to prosocial behavior and justice orientation. Thus, it seems important to understand ways in which empathy can be measured via an instrument with adequate psychometric properties.

Culturally Specific Empathy

Although the literature on general empathy provides operationalization of the construct, in light of cultural differences between groups, it seems critical to expand this definition to include culture. The concept of ethnocultural empathy is relatively new in the psychological literature; thus, the terminology used to define this construct has not been solidified. Terms such as *cultural empathy* (Ivey, Ivey, & Simek-Downing, 1987; Ridley & Lingle, 1996), *empathetic multicultural awareness* (Junn, Morton, & Yee, 1995), *cultural role taking* (Scott & Borodovsky, 1990), *ethnic perspective taking* (Quintana, Ybarra, Gonzalez-Doupe, & Baessa, 2000), and *ethnotherapeutic empathy* (Parson, 1993) have been used interchangeably to speak to the concept of empathy in cross-cultural settings.

Ethnic perspective taking has been examined with Mexican American populations by Quintana (1994) and his colleagues (Quintana, Castaneda-English, & Ybarra, 1999). Quintana (1994) operationalized ethnic perspective taking using a theory of social perspective taking as well as ethnic cognition. Specifically, Quintana and his colleagues defined ethnic perspective taking as a cognitive–developmental ability that contains associated levels. Each level of ethnic perspective taking is reached as an individual proceeds through developmental life stages (Quintana, 1994; Quintana et al., 1999). For example, Quintana et al. (1999) summarized various developmental stages of ethnic perspective-taking ability as including "awareness of ethnic discrimination and prejudice," "awareness of perspectives, attitudes, experiences shared by ethnic group," and "enhanced ability to take the perspective of other ethnic groups" (p. 163).

Ridley and Lingle (1996) have developed the most complex and complete model of cultural empathy. These authors defined cultural empathy as a "learned ability" (Ridley & Lingle, 1996, p. 32) and added that it can be further characterized by its multidimensionality and interpersonal nature. Their model is composed of three subordinate processes: cognitive, affective, and communicative. Cognitive process can be understood as a cultural perspective-taking and cultural self–other differentiation. Affective process includes vicarious affect and the expressive concern. Probing for insight and conveying accurate understanding makes up the communicative process of cultural empathy in this model.

Operationalization of Ethnocultural Empathy

The construct of ethnocultural empathy measured by the Scale of Ethnocultural Empathy (SEE), developed in the current studies, is built on these theoretical discussions of general and culturally specific empathy. We posit that ethnocultural empathy, as a learned ability and a personal trait, can be assessed. Thus, the construct of ethnocultural empathy, as operationalized for this study, is composed of intellectual empathy, empathic emotions, and the communication of those two (Ridley & Lingle, 1996). Perhaps ethnocultural empathy includes more than three components; however, we base our operationalization on the existing literature, which discusses only these three dimensions.

Intellectual empathy is the ability to understand a racially or ethnically different person's thinking and/or feeling. It is also the ability to perceive the world as the other person does; that is, racial or ethnic perspective taking. The empathic emotions component of ethnocultural empathy is attention to the feeling of a person or persons from another ethnocultural group to the degree that one is able to feel the other's emotional condition from the point of view of that person's racial or ethnic culture. In addition, it refers to a person's emotional response to the emotional display of a person or persons from another ethnocultural group. The communicative empathy component is the expression of ethnocultural empathic thoughts (intellectual empathy) and feelings (empathic emotions) toward members of racial and ethnic groups different from one's own. This component can be expressed through words or actions. Further, we conceptualized ethnocultural empathy as a trait that can be developed over time. Such a conceptualization allows us to measure interpersonal differences and take into account the developmental model of ethnic perspective taking (Duan & Hill, 1996; Quintana et al., 1999).

To date, we have found no existing measures of cultural empathy or ethnocultural empathy. Therefore, the primary purpose of our investigation was to develop a quantitative measure of empathy directed toward members of racial and ethnic groups different from one's own. The topic of ethnocultural empathy is intricately tied to counseling psychology's current emphasis on diversity and multiculturalism as professional and scholarly issues. In this article, we present the methods and results of three studies, each representing a distinct stage of the scale development process, followed by an overall discussion of the instrument's implications and limitations. The creation of the scale, including item generation and validity as well as exploratory factor analysis, is the focus of Study 1. It was expected that ethnocultural empathy would be multidimensional, consisting of three components, or factors, including empathic emotions, intellectual empathy, and communicative empathy. It was further expected that these three factors would be moderately correlated with each other and with the total scale. Study 2 focuses on confirmatory factor analysis and validity estimates. For discriminant validity purposes, we hypothesized that scores on the instrument, both total and subscales, would have low correlations with a measure of social desirability. In terms of convergent validity, we hypothesized that scores on the scale would have moderate correlations with a multidimensional measure of general empathy. Additionally, we hypothesized that the

SEE would be positively associated with measures of attitudes and awareness of individual similarities and differences but that the SEE would focus on ethnocultural empathy rather than simply attitudes. In Study 3, we examine test–retest reliability of the scale over a 2-week time period.

Study 1: Instrument Development and Exploratory Factor Analysis

The purpose of Study 1 was to develop items for the SEE that reflected the construct of ethnocultural empathy and to investigate the factor structure of those items.

Method

Scale development and item generation procedure. A team of six counseling psychology doctoral students (four women and two men; one Taiwanese/Chinese, three Western European American, one Eastern European, and one Hispanic/Polynesian) composed the research team that constructed the SEE. A comprehensive literature search on empathy and ethnocultural theory served as the major theoretical base for the current project. Within this literature search, five existing scales were referenced to begin generating items for the SEE. These scales were the Multicultural Awareness–Knowledge–Skills Survey (D'Andrea, Daniels, & Heck, 1991), the Multicultural Counseling Awareness Scale (Ponterotto, Sanchez, & Magids, 1991), the Multicultural Counseling Inventory (Sodowsky, Taffe, Gutkin, & Wise, 1994), the Cross Cultural Counseling Inventory—Revised (LaFromboise, Coleman, & Hernandez, 1991) and the Interpersonal Reactivity Index (Davis, 1983).

From the findings of the literature review of empathy research, the research team brainstormed together and delineated three theoretically distinct domains of empathy: (a) intellectual empathy, (b) empathic emotions, and (c) communicative empathy. Two professors and one doctoral candidate, all of whom were in the counseling psychology field and possessed expertise in multicultural issues, reviewed the operational definition of ethnocultural empathy. Their feedback was incorporated into the final definition used to generate the item pool of the SEE.

Although we did not conceptualize intellectual empathy, empathic emotions, and communicative empathy as independent from each other, these three empathy constructs served as the tridimensional reference points from which item generation was conducted. Each SEE item needed to have a primary focus on one of these three domains of empathy. An original pool of 71 items was written; then, validity and reliability procedures, which are described below, were followed in an effort to revise and select only the strongest items for each subscale.

Validity and reliability procedures for item generation. Three judges (one PhD counseling psychologist, one counseling psychology doctoral candidate, and one master's level journalism student) determined the appropriateness level for each item by completing an item-to-construct assignment task (i.e., matching items to the three subscale construct definitions). The two individuals from psychology disciplines were selected to participate in this task because of their specific experience with multicultural issues and their familiarity with the literature; the individual from journalism was selected to participate to represent an interdisciplinary perspective. Each judge was given definitions describing the three subscales of intellectual empathy, empathic emotions, and communicative empathy and a complete set of items. The judges were asked to read each item and decide into which of the three category subscales the item should be placed. Also, each judge was asked to rate each item on a scale from 1 (*not appropriate to the construct measured*) to 6 (*most appropriate to the construct measured*). In addition, they were asked to provide comments indicating any sources of confusion they had experienced with the item.

A satisfactory level of interjudge agreement was obtained if two of the three judges agreed on the classification of the item to the subscale domain identified during item generation. Furthermore, item appropriateness was determined acceptable if the mean of the three judges' ratings was above a mean of 4 on the 6-point Likert-type scale. According to both of the above criteria, of the 71 items, 9 were deleted and 6 were revised. After these 9 items were deleted, the overall mean interjudge appropriateness rating was 5.27, with a standard deviation of 0.49. The revised items were reviewed again by the three judges for the item classification and appropriateness. All of the 6 revised items (mean ratings ranged from 4.5 to 6.0) met the above criteria and thus were included in the surveys administered to the participants.

Participants and procedure. After accounting for missing and invalid data ($n = 16$), we used a data set that included 323 undergraduate students who were enrolled in three Midwestern universities or colleges for the analysis. The sample included more women (66%, $n = 213$) than men (25%, $n = 81$), with 29 respondents not indicating their gender. Most of the participants were between the ages of 18 and 22 (97%, $M = 19.73$ years); most were single and had never been married (97%). Approximately 88% of the participants had been attending their current university or college for 2 or fewer years, 10% for 3 to 4 years, and 2% for more than 4 years. A majority of students in the sample described themselves as Caucasian (83%). African Americans constituted 6% of the sample, 5% were Asian American or Pacific Islander, 3% were biracial, 2% were Hispanic or Latino/Latina, and 1% were Native American. Less than 1% of the participants described themselves as international students or "other."

On average, participants indicated that they had approximately five friends of a different race or ethnicity, and 72% had no family members of a different race or ethnicity, in contrast to those who reported having one or more family member of a different race or ethnicity. A total of 51% described their neighborhood as not being diverse, and 67% had attended a high school in which they belonged to the majority group and were accepted by their peers. A total of 34% had attended a diverse high school that was composed of one majority group, and 40% had attended a high school with few students from diverse backgrounds.

Participants for this study were students recruited from psychology classes. Approval to conduct research involving human subjects was sought from and granted by the Institutional Review Board of all three universities or colleges. Participation was voluntary, and participants were guaranteed anonymity, confidentiality, and the right to know the results of their responses. Participants were informed that the purpose of the study involved collection of data for research.

Instruments. The questionnaire packet used in Study 1 consisted of three parts: a demographic questionnaire, the SEE, and the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984, 1991).

The demographic questionnaire requested participants' demographic and other relevant information (age, gender, academic level, race/ethnicity, racial/ethnic makeup of family members, high school classmates, friends, and neighborhood where they grew up, as well as how well they fit in with the racial/ethnic culture of their high schools).

The SEE used in this first study was composed of 62 items reflecting the three ethnocultural empathy dimensions. There were 25 intellectual empathy items, 21 empathic emotion items, and 16 communicative empathy items. All of the items were listed in random order and rated on a 6-point Likert-type scale (1 = *strongly disagree that it describes me* to 6 = *strongly agree that it describes me*). Items were phrased both positively (38 items) and negatively (24 items) to offset any potential response bias. Negatively phrased items were reverse-scored. In addition, 3 items were included to measure the participants' validity in truthfully completing the survey (e.g., "I am reading each of these items carefully" and "Please code a 9 for this item"). Scores for the SEE were obtained by summing the item scores. Higher scores indicated a higher level of ethnocultural empathy.

The BIDR (Paulhus, 1984, 1991), with a 7-point Likert-type scale (0 = *not true* to 6 = *very true*), was administered as a measure for assessing social desirability. The full scale contains 40 items that measure two constructs: (a) self-deception (the tendency to give self-reports that are

honest but positively biased) and (b) impression management (deliberate self-presentation to an audience; Paulhus, 1984). For the purposes of this study, the construct of concern was impression management, because it was suspected that some respondents may be worried about their responses associated with racism and discrimination. The Impression Management subscale was developed on the basis of the assumption that some people tend to underreport undesirable behaviors and overreport desirable behaviors in an attempt to manage their impression and receive social approval.

To assess truly exaggerated desirable responses, Paulhus (1991) recommended adding 1 point for each extreme response (6 or 7) after reversing the negatively keyed items. Thus, total possible scores on the 20-item Impression Management subscale range from 0 to 20. The Impression Management mean scores of a college student sample in a study by Paulhus (1988) revealed males scoring 4.3 ($SD = 3.1$) and females scoring 4.9 ($SD = 3.2$). In another study, an overall mean of 11.9 ($SD = 4.5$) was found among a college student population in a public disclosure condition (Paulhus, 1984). In a sample of religious adults, Quinn (1989, as cited in Paulhus, 1991) reported the Impression Management subscale scores for men at 7.3 ($SD = 3.1$) and for women at 8.9 ($SD = 3.2$). In addition, various validity estimates have been summarized and reported (Paulhus, 1991). The Impression Management scores are positively related to scores on lie scales and the Marlowe–Crowne Social Desirability Scale (for a review, see Paulhus, 1991). Internal consistency (coefficient alphas) for the Impression Management subscale scores ranged from .75 to .86 in the previous studies. The test–retest reliability correlation for the Impression Management subscale was found to be .65 by Paulhus (1984). The alpha coefficient for the Impression Management subscale in the current study was .73.

Results

To adjust for Type I and Type II errors, we set a stringent alpha level of .01 for all of the analyses throughout the three studies.

Exploratory factor analysis. Prior to the analyses, we screened the data by using the 3 validity items in the 62-item scale. Any observations with one or more incorrectly answered validity items were removed from the sample; a total of 16 observations were excluded from the following analyses. Prior to analysis of the factor structure, 9 items were removed for having skewness or kurtosis greater than 2.0. The 323 observations in the final data set for this analysis provided a ratio of 5.2 cases per item, thus fulfilling a minimum requirement of 5 cases per item for factor analysis (Stevens, 1996). Also, as a general rule of thumb, at least 300 cases were recommended for factor analysis (Tabachnick & Fidell, 1996). Additionally, the Kaiser–Meyer–Olkin measure of sampling adequacy was calculated to be .90, further attesting to the factorability of the data set (Kaiser, 1970).

Because numerous items were moderately or highly correlated with each other statistically and theoretically, orthogonal rotation of factors was considered inappropriate (Stevens, 1996; Tabachnick & Fidell, 1996). A principal-components analysis with an oblique (promax) rotation was performed through SPSS 10.0 to estimate the number of components. The number of components to be extracted was then determined by (a) eigenvalues above 1.0 (Kaiser, 1958) and (b) Cattell's scree test (Cattell, 1966). However, Kaiser's criterion can yield too many factors; thus, the retention of factors was determined by the scree plot (Zwick & Velicer, 1986), which suggested retaining one to four components. On the basis of this information and the amount of total variance accounted for by one to four factors (12% to 49%), we performed principal-components analyses with an oblique (promax) rotation by specifying one-, two-, three-, and four-factor solutions.

The four-factor solution was finally chosen for two reasons: (a) It was the most conceptually interpretable, and (b) it resulted in the most sound factor structure with stronger item loadings and factor internal consistency. Variables with single-factor loadings less than .40 and variables with cross-loadings greater than .30 were eliminated. Thirty-one items were eliminated from the original 62 items on the basis of the above criteria. The first component (15 items) accounted for 29% of the total variance (eigenvalue = 8.99). The second (7 items), the third (5 items), and the fourth (4 items) components accounted for 7%, 6%, and 5% of the total variance, respectively (eigenvalues = 2.24, 1.81, and 1.62). These four factors, subsequently named Empathic Feeling and Expression (EFE), Empathic Perspective Taking (EP), Acceptance of Cultural Differences (AC), and Empathic Awareness (EA), as a set, accounted for 47% of the total variance. Factor loadings, communalities, item–total correlations, means, and standard deviations are presented in Table 1. Cronbach's alpha internal consistency estimates for the final 31-item SEE total scale and the four factors were .91, .90, .79, .71, and .74, respectively.

Normative information. The means and standard deviations for the SEE total and four factors are as follows: SEE total: $M = 4.3$, $SD = 0.71$; EFE: $M = 4.5$, $SD = 0.85$; EP: $M = 3.4$, $SD = 1.1$; AC: $M = 4.8$, $SD = 0.88$; EA: $M = 4.7$, $SD = 0.89$. In addition, skewness and kurtosis indices were computed for the SEE total and factors, ranging from –.56 to .28 and –.51 to .37, respectively. These low numbers reveal that the distributions of the SEE total and factor scores are approximate to a normal distribution (Tabachnick & Fidell, 1996).

Factor intercorrelations. The intercorrelations among the factors suggest that the four factors are somewhat interrelated but still represent distinct constructs (see Table 2).

Validity estimates. Correlation analyses were performed on each of the four scale factors and the total SEE scale score with the BIDR Impression Management subscale scores. Only one significant correlation was found, between the BIDR Impression Management score and AC ($r = .17$, $p < .01$; less than 4% of the total variance). There were no significant correlations found between BIDR Impression Management and either SEE total score or EFE, EP, or EA ($r = .08$, .10, .00, and –.03, respectively). These figures provide evidence for the discriminant validity of the SEE scale and its four factors.

Naming of the factors. The first scale, Empathic Feeling and Expression, includes items that pertain to concern about communication of discriminatory or prejudiced attitudes or beliefs as well as items that focus on emotional or affective responses to the emotions and/or experiences of people from racial or ethnic groups different from one's own. This expression of empathy can take the form of thoughts, feelings, words, or actions that relate directly to the discriminatory experiences. The second scale, Empathic Perspective Taking, is composed of items that indicate an effort to understand the experiences and emotions of people from different racial and ethnic backgrounds by trying to take their perspective in viewing the world. The third scale, Acceptance of Cultural Differences, includes items that center on the understanding, acceptance, and valuing of cultural traditions and customs of individuals from differing racial and ethnic groups. Finally, the fourth scale, Empathic Awareness, includes items that appear to focus on the awareness or knowledge that one has about the experiences of people from racial or ethnic groups different from one's own. This

Table 1
Item–Total Correlations, Factor Matrix, Communalities, and Item Means and Standard Deviations

Item	Item–total correlation	Factor loading	h^2	*M*	*SD*
Empathic Feeling and Expression					
30. When I hear people make racist jokes, I tell them I am offended even though they are not referring to my racial or ethnic group.	.54	.77	.40	3.53	1.63
21. I don't care if people make racist statements against other racial or ethnic groups. (R)	.61	.75	.48	5.08	1.21
16. I rarely think about the impact of a racist or ethnic joke on the feelings of people who are targeted. (R)	.56	.69	.45	4.99	1.18
23. When other people struggle with racial or ethnic oppression, I share their frustration.	.75	.64	.61	4.10	1.39
14. I feel supportive of people of other racial and ethnic groups, if I think they are being taken advantage of.	.60	.62	.37	4.70	1.13
12. I share the anger of those who face injustice because of their racial and ethnic backgrounds.	.68	.61	.51	4.24	1.41
26. I share the anger of people who are victims of hate crimes (e.g., intentional violence because of race or ethnicity).	.48	.59	.41	4.88	1.32
11. When I know my friends are treated unfairly because of their racial or ethnic backgrounds, I speak up for them.	.52	.57	.24	4.90	1.06
15. I get disturbed when other people experience misfortunes due to their racial or ethnic backgrounds.	.63	.57	.45	4.99	1.18
3. I am touched by movies or books about discrimination issues faced by racial or ethnic groups other than my own.	.50	.55	.29	4.87	1.29
22. When I see people who come from a different racial or ethnic background succeed in the public arena, I share their pride.	.60	.54	.50	4.47	1.31
17. I am not likely to participate in events that promote equal rights for people of all racial and ethnic backgrounds. (R)	.56	.53	.36	4.51	1.41
9. I seek opportunities to speak with individuals of other racial or ethnic backgrounds about their experiences.	.56	.51	.35	3.85	1.47
13. When I interact with people from other racial or ethnic backgrounds, I show my appreciation of their cultural norms.	.59	.49	.37	4.54	1.15
18. I express my concern about discrimination to people from other racial or ethnic groups.	.67	.46	.40	3.73	1.35
Empathic Perspective Taking					
19. It is easy for me to understand what it would feel like to be a person of another racial or ethnic background other than my own.	.58	.75	.54	2.94	1.55
31. It is difficult for me to relate to stories in which people talk about racial or ethnic discrimination they experience in their day to day lives. (R)	.51	.71	.45	3.55	1.39
28. It is difficult for me to put myself in the shoes of someone who is racially and/or ethnically different from me. (R)	.59	.71	.53	3.47	1.51
4. I know what it feels like to be the only person of a certain race or ethnicity in a group of people.	.46	.64	.38	3.04	2.09
6. I can relate to the frustration that some people feel about having fewer opportunities due to their racial or ethnic backgrounds.	.53	.48	.29	3.32	1.72
29. I feel uncomfortable when I am around a significant number of people who are racially/ethnically different than me. (R)	.48	.46	.37	3.84	1.46
2. I don't know a lot of information about important social and political events of racial and ethnic groups other than my own. (R)	.37	.44	.19	3.47	1.58
Acceptance of Cultural Differences					
10. I feel irritated when people of different racial or ethnic backgrounds speak their language around me. (R)	.39	.76	.53	4.75	1.40
1. I feel annoyed when people do not speak standard English. (R)	.40	.67	.42	4.00	1.59
5. I get impatient when communicating with people from other racial or ethnic backgrounds, regardless of how well they speak English. (R)	.39	.53	.29	5.18	1.09
27. I do not understand why people want to keep their indigenous racial or ethnic cultural traditions instead of trying to fit into the mainstream. (R)	.39	.41	.25	5.02	1.20
8. I don't understand why people of different racial or ethnic backgrounds enjoy wearing traditional clothing. (R)	.48	.40	.27	4.98	1.08
Empathic Awareness					
25. I am aware of how society differentially treats racial or ethnic groups other than my own.	.55	.70	.53	4.94	0.98
24. I recognize that the media often portrays people based on racial or ethnic stereotypes.	.44	.66	.40	4.95	1.11
20. I can see how other racial or ethnic groups are systematically oppressed in our society.	.55	.63	.47	4.52	1.17
7. I am aware of institutional barriers (e.g., restricted opportunities for job promotion) that discriminate against racial or ethnic groups other than my own.	.44	.58	.33	4.35	1.42

Note: There are 31 items; $\alpha = .91$. Reverse-scored items are indicated (R). $N = 323$ participants. The h^2 value is the communality of each item. All item–total correlations are significant at $p < .01$.

Table 2
Factor Intercorrelations in Study 1

Factor	1	2	3	4
1. Empathic Feeling and Expression	—	.57	.49	.59
2. Empathic Perspective Taking		—	.29	.41
3. Acceptance of Cultural Differences			—	.39
4. Empathic Awareness				—

awareness of the emotions and experiences of others is particularly related to their experiences of discrimination or unequal treatment of different groups.

Study 2: Confirmatory Factor Analysis and Validity Estimates

The primary purpose of Study 2 was to examine the stability of the factor structure obtained in Study 1. A secondary purpose was to examine additional validity of the SEE. Specifically, we were interested in determining whether the SEE (a) was not strongly associated with social desirability, (b) was related to other measures of empathy, and (c) was associated with a measure of awareness of and attitudes toward people's similarities and differences. It was hypothesized that the SEE would be low to moderately related to a general multidimensional construct of empathy (Davis, 1983), yet that the construct of ethnocultural empathy measured by the present scale would be distinct from the general construct. Also, we predicted that the subscales from the SEE and general empathy inventory with underlying emotional components would have stronger correlation with each other than with the other subscales; the same direction was predicted for the subscales, with more focus on the perspective-taking abilities. We similarly hypothesized that the SEE would be associated with measures of attitudes and awareness of individual similarities and differences but also would be a scale that focuses on measurement of ethnocultural empathy rather than attitudes. Finally, possible associations were examined between the SEE and individual group memberships (i.e., gender, race, and ethnicity) as well as the diverse personal background.

Method

Participants and procedure. A total of 364 undergraduate students, who were enrolled in two large Midwestern universities, completed the survey; however, data were missing or invalid for 24 of these participants, and so the data set used for this analysis consisted of 340 observations. More women (63%, $n = 213$) than men (37%, $n = 127$) participated in this study. The mean age of the participants was 20.81 years, and 90% were between the ages of 18 and 22. Approximately 92% were single and had never been married. Nearly 32% of the participants had attended their respective universities for up to 2 years, whereas 53% had attended more than 2 but less than 4 years, and 8% had attended for more than 4 years. Most of the participants described themselves as European American (79%). African Americans comprised 14% of the sample, whereas 2% were Asian American or Pacific Islander, 2% were biracial, 2% were Hispanic or Latino/Latina, and less than 1% were Native American. Approximately 2% of the participants described themselves as international students, and less than 1% described themselves as "other."

On average, participants in Study 2 indicated that they had five friends of a different race or ethnicity, and 70% had no family members of a different race or ethnicity. Approximately 56% had attended a high school in which they belonged to the racial/ethnic majority and were accepted by their peers. A total of 40% had attended a very diverse high school with no single majority or one majority group, and 39% had attended a high school with few or no students from diverse backgrounds. A total of 56% described the neighborhood where they grew up as not being diverse, whereas 41% lived in a neighborhood ranging from somewhat to quite diverse.

Different from the first study, the participants for Study 2 were students recruited from various disciplines, including art, statistics, education, counseling, and mathematics classes. Approval to conduct research involving human subjects was sought from and granted by the Institutional Review Board of the two universities. Participation was voluntary, and participants were guaranteed anonymity, confidentiality, and the right to know the results of their responses. Participants were informed that the purpose of study involved collection of data for research.

Instruments. The measures used in Study 2 consisted of the same measures used in Study 1 and two additional measures for establishing convergent and discriminant validity: the Miville–Guzman Universality–Diversity Scale (M-GUDS; Miville et al., 1999) and the Davis Interpersonal Reactivity Index (IRI; Davis, 1983). A demographic questionnaire was included in the beginning of the survey and was described in Study 1.

The M-GUDS (Miville et al., 1999) is a 15-item scale designed to assess a *universal–diverse* orientation, defined as "an attitude of awareness and acceptance of both the similarities and differences that exist among people" (p. 291). It has three subscales: Diversity of Contact, Relativistic Appreciation, and Comfort With Differences, and each response is recorded on a 7-point Likert-type scale (1 = *strongly disagree* to 7 = *strongly agree*). The Diversity of Contact subscale assesses individual intent and previous behaviors that were involved in interpersonal contacts with people from different backgrounds. The Relativistic Appreciation subscale is related to the degree to which one is aware of and receptive to the differences as well as similarities among people. The Comfort With Differences subscale assesses one's level of comfort being with people from different personal backgrounds. Yet it was indicated that the total score, rather than the subscale scores, should be interpreted according to the unidimensional nature of the factor structure of the scale. Internal consistency and test–retest reliability ranged from .89 to .95 (Miville et al., 1999). Convergent and discriminate validity estimates were reported to support for the construct validity of the M-GUDS. For example, the M-GUDS was found to be significantly correlated with measures of empathy, racial identity, and feminism. In particular, the M-GUDS total score was significantly and positively correlated with the Perspective Taking and the Empathic Concern subscales of the IRI (Miville et al., 1999). The estimates of internal consistency for the M-GUDS in the current study were .75 (Diversity of Contact), .76 (Relativistic Appreciation), .75 (Comfort With Differences), and .85 (M-GUDS total).

The IRI (Davis, 1983) measures four separate aspects of empathy using a 5-point Likert-type scale (0 = *does not describe me well* to 4 = *describes me very well*). It has four subscales (Perspective Taking, Fantasy, Empathic Concern, and Personal Distress); these components reflect responsivity to others at the same time as representing distinct facets of general empathy. However, for the purposes of this study, only the Perspective Taking and Empathic Concern subscales were used, which resulted in a 14-item scale. The Perspective Taking subscale measures "the tendency to spontaneously adopt the psychological point of view of others," whereas the Empathic Concern subscale assesses "other-oriented feelings of sympathy or concern for unfortunate others" (Davis, 1983, pp. 113–114). In terms of construct validity, the Perspective Taking subscale was found to be associated with better interpersonal functioning, higher self-esteem, and less emotionality; the Empathic Concern subscale was reported to be clearly related to more emotional reactivity and selfless concern for others. Both subscales were correlated with the other measures of empathy (see Davis, 1983). All subscales of the IRI have demonstrated satisfactory internal consistencies (ranging from .71 to .77) and test–retest reliabilities (ranging from .62 to .71;

Davis, 1983). The coefficient alphas for the Perspective Taking and Empathic Concern subscales were .77 and .80, respectively, in the current study.

Results

Confirmatory factor analysis. In order to (a) avoid "estimating a large number of parameters (i.e., factor loadings and error terms) in fitting the model to the data" as well as (b) decrease the likelihood of obtaining distorted results due to "idiosyncratic characteristics of individual items" (Russell, Kahn, Spoth, & Altmaier, 1998, p. 22), we generated 9 item bundles (or parcels) instead of using the 31 individual items of the SEE prior to conducting the confirmatory factor analysis. Following the recommendation of Russell et al. (1998), we divided the 31 items from the four SEE subscales—EFE (15 items), EP (7 items), AC (5 items), and EA (4 items)—into 9 bundles (3 bundles for the EFE, 2 bundles for the EP, 2 bundles for the AC, and 2 bundles for the EA). More specifically, to create the item parcels, four exploratory factor analyses were conducted, and the items from each of the four subscales (EFE, EP, AC, and EA) were fit into four one-factor models. The items were rank-ordered according to the magnitude of their factor loadings and were subsequently assigned to bundles to equate the average loadings of each bundle on each factor. Using the EFE latent variable as an example, items ranked 1, 4, 7, 10, and 13 were assigned to Bundle 1 (EFE1); items ranked 2, 5, 8, 11, and 14, to Bundle 2 (EFE2); and items ranked 3, 6, 9, 12, and 15, to Bundle 3 (EFE3) (see Figure 1). Because the numbers of items in each bundle were unequal, we used the average scores (rather than the sum) of each item to form each bundle.

A confirmatory factor analysis was then conducted on the four-factor model of the SEE, using the SAS System's CALIS procedure (SAS Institute, 1992). The analysis was twofold: (a) examining the adequacy of the four-factor model and (b) comparing the four-factor model with a hierarchical factor model to assess which model provided a better fit to the data. Throughout the analysis, five indices were used to assess goodness of fit of the models: chi-square, chi-square/*df* ratio (Newcomb, 1994; best if less than 2.0), nonnormed fit index (NNFI; Bentler, 1995; best if .90 or greater), normed fit index (NFI; Bentler & Bonett, 1980; best if .90 or greater), and root-mean-square error of approximation (RMSEA; Browne & Cudeck, 1993; MacCallum, Browne, & Sugawara, 1996; best if .05 or less).

A maximum-likelihood method through CALIS in the SAS program was used to estimate goodness of fit of the four-factor model (see Figure 1). In confirmatory factor analysis, the variances of the factors were initially fixed to 1.0 (Hatcher, 1994). The estimation of the initial model suggested that the model was an excellent fit of the data, as indicated by the following indices: $\chi^2(21, N = 340) = 26.60$, $p = .18$; $\chi^2/df = 1.27$; NNFI = .99; NFI = .98; RMSEA = .03. In addition, all factor loadings were significant at the $p < .01$ level (ranging from .76 to .96), suggesting that the four factors were well constructed by the bundles. The four factors shared approximately 81% of the total variance.

Theoretically and conceptually, there could also exist a higher order, more abstract construct (i.e., SEE: general ethnocultural empathy) that affects more specific ethnocultural empathy factors (i.e., the EFE, EP, AC, and EA). Thus, we performed another CALIS procedure with a maximum-likelihood method to assess the fit indices of this second-order hierarchical factor model (see Figure 2). The second-order factor (SEE) has direct effects on each first-order factor (EFE, EP, AC, and EA) and accounts for the intercorrelations among the four separate factors (Kline, 1998). Per Kline's recommendation, one loading on each first-order factor was fixed to 1.0. The analysis generated the following fit indices: $\chi^2(23, N = 340) = 27.56$, $p = .23$; NNFI = .99; NFI = .98;

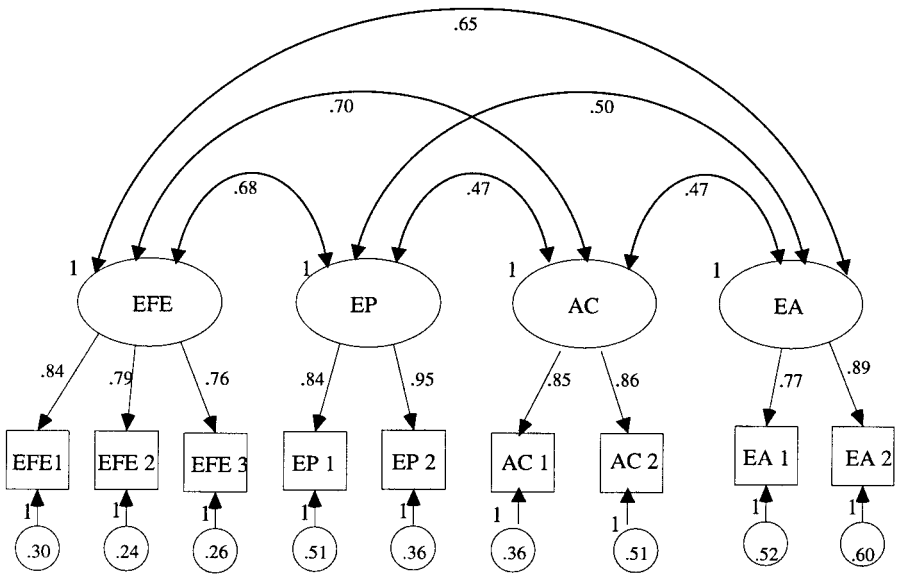

Figure 1. The four-oblique-factor model. The large circles designate the latent constructs, the small circles indicate the residual variances, and the rectangles are the measured variables. Factor loadings are standardized (all *p*s < .01). EFE = Empathic Feeling and Expression; EP = Empathic Perspective Taking; AC = Acceptance of Cultural Difference; EA = Empathic Awareness.

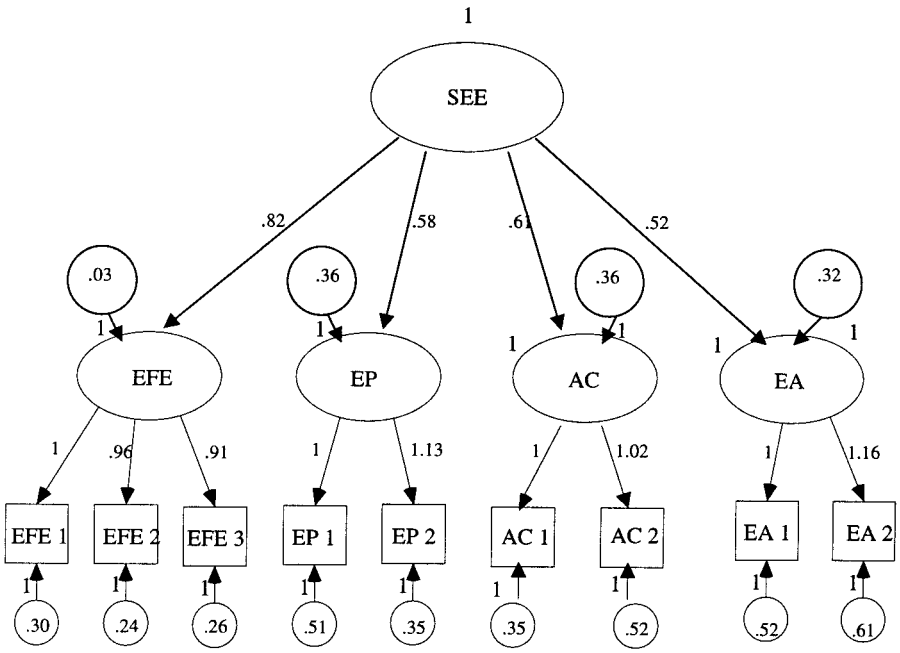

Figure 2. The second-order hierarchical model. The large circles designate the latent constructs, the small circles indicate the residual variances, and the rectangles are the measured variables. Factor loadings are standardized (all *p*s < .01). SEE = Scale of Ethnocultural Empathy total score; EFE = Empathic Feeling and Expression; EP = Empathic Perspective Taking; AC = Acceptance of Cultural Difference; EA = Empathic Awareness.

RMSEA = .02. All factor loadings were significant at the $p < .01$ level, suggesting satisfactory convergent validity for the first-order factors. The effects of the second-order factor on the first-order factor were strong, which leads to a relatively small magnitude of unexplained variance for the first-order factors.

Although both of the four-oblique-factor and the second-order hierarchical models generated excellent fit indices, there was a need to compare these two models in order to determine which model provided better fit to the data. Because these two models are not nested, it was impossible to conduct direct statistical comparison. Yet some procedures were recommended to be used for comparing nonnested models (Maruyama, 1998), such as Akaike's informational criteria (AIC; Akaike, 1987; Maruyama, 1998), consistent Akaile's informational criterion (CAIC; Bozdogan, 1987), expected cross-validation index (ECVI; Browne & Cudeck, 1993), and Schwarz's Bayesian criterion (SBC; Schwarz, 1978; Sclove, 1987). Lower values of AIC, CAIC, ECVI, and SBC are generally associated with better model fit in comparing nonnested models (Maruyama, 1998). The four-oblique-factor model yielded the following results: AIC = –15.40; CAIC = –116.87; ECVI = .22; SBC = –95.87. In contrast, the second-order hierarchical model resulted in the following indices: AIC = –18.44; CAIC = –129.58; ECVI = .21; SBC = –106.58. Therefore, despite the results' indication that both of the models well represented the SEE factor structure in the current study, the second-order hierarchical model is a better fit of the data according to all four indices. Furthermore, the second-order hierarchical model would be more consistent with the conceptualization of the SEE, which includes the four distinct factors as well as a broad ethnocultural empathy factor (as described earlier).

Normative information. The means and standard deviations for the SEE total and four factors are as follows: SEE total: $M = 4.2$, $SD = 0.75$; EFE: $M = 4.3$, $SD = 0.86$; EP: $M = 3.4$, $SD = 1.0$; AC: $M = 4.6$, $SD = 0.98$; EA: $M = 4.6$, $SD = 0.99$. These numbers are similar to the previous findings of the exploratory factor analysis conducted in Study 1. In addition, skewness and kurtosis indices were scrutinized for the SEE total and factors, ranging from –.67 to .24 and –.55 to .51, respectively. These low numbers reveal that the distributions of the SEE total and factor scores are approximate to a normal distribution (Tabachnick & Fidell, 1996).

Reliability. Estimates of internal consistency for the SEE total and each of the factors were measured by the alpha coefficients. Alphas of .91, .89, .75, .73, and .76 were obtained for the SEE total, EFE, EP, AC, and EA. These initial estimates of reliability were very similar to the ones found in Study 1, which suggests again that the SEE and the four factors have acceptable levels of internal consistency in the current sample.

Factor intercorrelations. The intercorrelations among the four factors are similar to the ones in Study 1, which again suggests that the factors are interrelated to a moderate level (see Table 3).

Discriminant validity. Correlation analyses were performed on each of the four scale factors and the total SEE scale score with the BIDR Impression Management subscale scores (see Table 3). Different from the finding in Study 1, three significant correlations were found, between the BIDR Impression Management score and EFE (7% of the total variance), AC (4% of the total variance), and the total scale (5% of the total variance). There were no significant correlations found between BIDR Impression Management and EP and EA. Although several of the correlations were statistically significant, they account for a minimal amount of the variance.

Table 3
Intercorrelations Between Subscales and Total Scores of SEE and Validity Instruments

Measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Empathic Feeling and Expression	—	.56*	.55*	.53*	.93*	.61*	.48*	.56*	.67*	.54*	.42*	.26*
2. Empathic Perspective Taking		—	.36*	.37*	.76*	.52*	.27*	.38*	.49*	.18*	.23*	.07
3. Acceptance of Cultural Differences			—	.34*	.68*	.47*	.39*	.59*	.59*	.34*	.30*	.21*
4. Empathic Awareness				—	.64*	.41*	.35*	.29*	.44*	.26*	.24*	.10
5. SEE Total					—	.65*	.48*	.58*	.70*	.48*	.42*	.23*
6. M-GUDS Diversity of Contact						—	.48*	.51*	.85*	.29*	.31*	.05
7. M-GUDS Relativistic Appreciation							—	.47*	.78*	.32*	.36*	.21*
8. M-GUDS Comfort with Differences								—	.81*	.40*	.37*	.21*
9. M-GUDS total									—	.41*	.42*	.20*
10. IRI Empathic Concern subscale										—	.41*	.38*
11. IRI Perspective Taking subscale											—	.29*
12. BIDR Impression Management subscale												—

Note. SEE = Scale of Ethnocultural Empathy; M-GUDS = Miville–Guzman Universality–Diversity Scale. IRI = Davis Interpersonal Reactivity Index; BIDR = Balanced Inventory of Desirable Responding.
* $p < .01$.

Thus, these data provide evidence for the discriminant validity of the SEE scale and its four factors. In addition, it should be noted that all of the other measures used in this study were significantly correlated with the Impression Management scale, with the exception of the Diversity of Contact subscale of the M-GUDS (see Table 3).

Concurrent validity. To further establish convergent validity, we performed correlation analyses on the four scale factors and the total SEE scale score with the IRI and the M-GUDS (see Table 3). For the purposes of this study, only the Perspective Taking and Empathic Concern subscales of the IRI were administered and analyzed. Significant correlations in the low to moderate range were found for each subscale of the IRI and each factor of the SEE and the total SEE score. As anticipated, the Empathic Concern of the IRI and the EFE of the SEE had the highest correlation coefficient, because of the "emotional reactivity" underlying both scales (Davis, 1983). Interestingly, the Perspective Taking subscale of the IRI also had the strongest correlation with the EFE—not the EP—of the SEE, which suggests that the EP subscale was tapping into a somewhat different perspective-taking ability (i.e., taking the perspectives of racial/ethnic groups different from one's own). Further analyses of the SEE and the M-GUDS yielded significant correlations between all subscales of both measures as well as the total scores for both measures. These results provide evidence for convergent validity of the SEE as a distinct measure of empathy.

Criterion-related validity. In an effort to ascertain the relationships between demographic variables and both the SEE total and the four factor scores, we performed two multivariate analyses of variance (MANOVAs), follow-up univariate analyses of variance (ANOVAs), and bivariate correlation procedures.

In terms of gender differences, an ANOVA revealed that women ($M = 4.33$, $SD = 0.69$) scored significantly higher than men ($M = 3.88$, $SD = 0.77$) on the SEE total, $F(1, 290) = 26.20$, $p < .01$. With the use of Wilks's lambda criterion, the MANOVA result suggested that the combined four SEE factors significantly differed by gender, $F(4, 287) = 10.51$, $p < .01$, Wilks's lambda = .87. Further univariate ANOVAs also indicated that women scored significantly higher than men on EFE, AC, and EA, yet the two groups did not differ in EP (see Table 4).

As for racial differences, because there was a relatively smaller number of participants from certain racial backgrounds (e.g., $n = 1$ for Native Americans, $n = 5$ for Latino Americans), the data from the non-White college students were combined in order to perform statistical comparisons. The non-White college students ($M = 4.75$, $SD = 0.56$) in our study scored significantly higher on SEE total than White college students ($M = 4.00$, $SD = 0.71$), $F(1, 290) = 62.54$, $p < .01$. To examine racial differences on SEE subscores, we performed another MANOVA, which revealed that the combined four SEE scores significantly differed by the participants' racial background, $F(4, 287) = 35.18$, $p < .01$, Wilks's

Table 4
Univariate Analyses of Variance on Gender Differences in SEE Subscale Scores

	Male		Female			
Measure	*M*	*SD*	*M*	*SD*	*F*	*df*s
Empathic Feeling and Expression	3.89	0.87	4.47	0.78	36.62*	1, 304
Empathic Perspective Taking	—	—	—	—	1.87	1, 331
Acceptance of Cultural Differences	4.30	1.04	4.78	0.89	19.85*	1, 329
Empathic Awareness	4.30	1.04	4.72	0.92	15.14*	1, 337

Note. $N = 340$ participants (male = 127, female = 213). SEE = Scale of Ethnocultural Empathy.
* $p < .01$.

lambda = .67. The results of the follow-up univariate ANOVAs indicated that the non-White college students in our study scored significantly higher than the White students on each of the four subscales (see Table 5).

The differences between SEE scores and diverse background variables were examined. The participants in this study who had more family members with different ethnic/racial backgrounds tended to score higher on SEE total and EFE and EP ($r = .23$, $p < .01$; $r = .23$, $p < .01$; and $r = .25$, $p < .01$, respectively). Similarly, the college students who had more friends with different racial/ethnic backgrounds tended to score higher on the SEE total, EFE, EP, and EA ($r = .30$, $p < .01$; $r = .25$, $p < .01$; $r = .32$, $p < .01$; and $r = .22$, $p < .01$, respectively). In addition, the participants who attended more diverse high schools and/or lived in more diverse neighborhoods tended to score significantly higher on SEE total ($r = –.15$, $p < .01$; $r = –.20$, $p < .01$, respectively) and EP ($r = –.22$, $p < .01$; $r = –.26$, $p < .01$, respectively).

Study 3: Examination of Test–Retest Reliability

The purpose of Study 3 was to provide additional reliability estimates, specifically test–retest reliability. We predicted that the SEE would be stable over time.

Method

A total of 51 undergraduate students (33 women and 18 men) enrolled in two large Midwestern universities completed the survey. Age of the participants ranged between 18 and 28 years ($M = 20.59$ years, $SD = 2.26$). Most of the participants described themselves as Caucasian (88%). African Americans constituted 2% of the sample, 2% were Asian American or Pacific Islander, and 2% were biracial. A total of 2% of participants described themselves as "other," and 4% described themselves as international students. No Latinos/Latinas or Native Americans completed the test–retest survey.

The participants for Study 3 were students recruited from classes in education, leadership, and psychology. Approval to conduct research involving human subjects was sought from and granted by the Institutional Review Board of the universities. Participation was voluntary, and participants were guaranteed anonymity, confidentiality, and the right to know the results of their responses. Participants were informed that the purpose of study involved collection of data for research. The measures used in Study 3 consisted of the SEE scale and a brief demographic questionnaire. Approximately 90% of participants completed the retest administration of the SEE 2 weeks after the first administration of the scale.

Results

Descriptive statistics. In general, the participants reported moderate levels of SEE total and subscores. Table 6 provides the means and standard deviations for test–retest data.

Reliability statistic. The 2-week test–retest reliability estimates for the SEE total and the subscales were as follows: SEE total ($r = .76$), EFE ($r = .76$), EP ($r = .75$), AC ($r = .86$), and EA ($r = .64$). These results suggest that both the SEE scale total and the subscales are acceptably stable over time.

Summary and Overall Discussion

The purpose of this study was to develop a scale to quantitatively and effectively measure the construct of ethnocultural empathy. On the basis of a review of the literature, ethnocultural empathy as a construct was derived from definitions of general empathy, multiculturalism, and cultural empathy. Our effort to create a three-factor scale was partially supported by the results of the factor analysis. Although there are parallels between the intended definition of the three factors and the final results, the factor analysis revealed the presence of a fourth factor. The factors contributed to the amount of variance accounted for by the scale in the following order, from strongest to weakest: Empathic Feeling and Expression, Empathic Perspective Taking, Acceptance of Cultural Differences, and Empathic Awareness. The four-factor solution suggests that there are four distinct components underlying scores on the SEE.

The four factors of the SEE closely replicate the theoretical constructs of empathy found in the literature and bring new understanding to the construct of ethnocultural empathy. Duan and Hill (1996) suggested that the concept of general empathy consists of two components: empathic emotions and intellectual empathy. The most comprehensive model of cross-cultural empathy, discussed by Ridley and Lingle (1996), includes communicative empathy in addition to the similar emotional and intellectual aspects. The four factors revealed in the present study not only support the existence of the emotional, intellectual, and communicative aspects of ethnocultural empathy but also suggested that the components of ethnocultural empathy may be more complicated than has been conceptualized in the existing literature (i.e., there may be more than three components). For example, the intellectual aspect of ethnocultural empathy may encompass one's perspective taking and awareness toward racial and ethnic differences.

Table 5
Univariate Analyses of Variance on Racial Differences in SEE Subscale Scores

	White		Non-White			
Measure	*M*	*SD*	*M*	*SD*	*F*	*df*s
Empathic Feeling and Expression	4.12	0.86	4.73	0.68	26.60*	1, 304
Empathic Perspective Taking	3.06	0.87	4.41	0.75	145.82*	1, 331
Acceptance of Cultural Differences	4.49	1.00	5.05	0.74	20.10*	1, 329
Empathic Awareness	4.43	0.99	5.05	0.82	24.58*	1, 337

Note. $N = 340$ participants (White = 268, non-White = 72). SEE = Scale of Ethnocultural Empathy.
* $p < .01$.

Table 6
Test–Retest Means and Standard Deviations for SEE Subscales and Total

Measure	*M*	*SD*
Test		
Empathic Feeling and Expression	4.57	0.58
Empathic Perspective Taking	3.25	0.92
Acceptance of Cultural Differences	4.67	0.86
Empathic Awareness	4.57	0.75
SEE total	4.29	0.53
Retest		
Empathic Feeling and Expression	4.70	0.69
Empathic Perspective Taking	3.58	0.96
Acceptance of Cultural Differences	4.80	0.89
Empathic Awareness	4.69	0.77
SEE total	4.45	0.63

Note. $N = 41$; retest data were collected 2 weeks after the initial testing. SEE = Scale of Ethnocultural Empathy.

The SEE is the first empirical measure to provides support for the theoretical construct of empathy in multicultural settings. Philosophically, just as a general counseling competency measure (e.g., the Counseling Self-Estimate Inventory; Larson et al., 1992) would not capture individuals' specific counseling competencies regarding persons who are culturally different from themselves (e.g., the Multicultural Counseling Awareness Scale; Ponterotto et al., 1996), a general empathy measure would not adequately portray individuals' empathy toward persons who are culturally different from themselves. This is the unique contribution of the SEE to the multicultural and empathy literatures.

The evidence for convergent validity was provided by the moderate correlations between the SEE and the other scales that measure general empathy (the IRI) and attitudes toward interpersonal similarities and differences (the M-GUDS). The magnitude of the correlation coefficients between the SEE and these two scales, as well as the relatively low correlations with a social desirability scale (the BIDR), also supports the validity of the SEE as measuring a unique construct. Test–retest reliability analyses suggest that the construct measured by the SEE is stable over a short time frame.

Although the sample in this study was fairly homogeneous, there were a few significant correlations between demographic variables and levels of ethnocultural empathy. Women were found to be significantly more ethnoculturally empathic than men in terms of Empathic Feeling and Expression, Empathic Awareness, and Acceptance of Cultural Differences but not in terms of Empathic Perspective Taking. These findings are consistent with several previous studies of general empathy that found that women have higher levels of affective empathy than men, but that no gender differences exist in regard to the perspective-taking component of empathy (Davis, 1994). In addition, non-White individuals were found to have significantly higher levels of general and specific ethnocultural empathy than their White counterparts. Because of the relatively small number of non-White participants in the current study, this result needs to be interpreted with caution and should be further examined with a larger, more diverse sample. However, in general, these group memberships (i.e., gender and race) may have had a significant impact on the participants' personal experience, awareness, interpersonal functioning, and social attitudes, as indicated by previous research (e.g., Cross & Phagen-Smith, 1996; Helms, 1995; Miville et al., 1999). Congruent with previous research, this study found that female and non-White participants were more likely to express empathic emotions around the issue of justice or fairness, were more receptive of cultural differences, and were more aware of the experiences of people from different racial or ethnic backgrounds than were their male and White counterparts. Also, compared with the White students, non-White students in this investigation reported more understanding of the experience of other ethnocultural groups by taking their perspective or sharing in their emotional experiences.

When diversity backgrounds were examined, increased diversity within immediate families, groups of friends, high school settings, and neighborhoods had significant correlations with higher levels of ethnocultural empathy. These findings support Pettigrew's current work that builds on Allport's "contact hypothesis," which states that contact between two groups can promote tolerance and acceptance (see DeAngelis, 2001). Our results also support the findings of previous studies that have shown that intergroup contact and social influence are strongly associated with greater understanding of differences between groups (Blanchard, Crandall, Brigham, & Vaughn, 1994; Blanchard, Lilly, & Vaughn, 1991; Firebaugh & Davis, 1988). Similarly, Quintana (1994) concluded that the development of an individual's ethnic perspective-taking ability depends on the number of experiences one has with one's own ethnic group and with other ethnic groups. Furthermore, he stated that those who are constrained to monocultural experiences may be considerably delayed in their ethnic perspective-taking ability development. However, the recent U.S. Census report (Kong, 2001), popular sources (Dutton, 2001; Newsweek International, 2001; Schmitt, 2001), and scholarly works (Frey & Farley, 1996; Massey & Denton, 1987) indicate that racial segregation continues to exist in the United States on many social levels. More interactions among different racial/ethnic groups are needed, which may lead to the development of greater cultural empathy.

Several limitations should be noted. First, because a four-factor model rather than a three-factor model resulted from our factor analysis, a better model might emerge with a broader sampling that includes more than the three domains (emotional, intellectual, and communicative aspects) of ethnocultural empathy that we examined. Also, the homogeneity of the sample limits the generalizability of the reliability and factorial analyses. Specifically, ethnocultural empathy may vary with age, education, race, and ethnicity. With this sample, predominately White and traditional college-aged, we could not fully examine this potential variability. Similarly, the geographic area in which the study was conducted, the midwestern United States, may also limit the generalizability of the results. Another limitation to consider is the motivation of the participants to complete the survey. For participating in the study, the majority of respondents received required research credit or extra credit for a class. This incentive for participation may have introduced bias into the sample. Finally, other researchers have suggested that response formats (e.g., interviews, picture–story measures, paper-and-pencil self-report) may determine the aspects of empathy assessed (Batson, 1987). For example, behavioral observations may add to the breadth of these self-report findings.

Directions for future research are numerous. We hope that our more complex, four-factor model of ethnocultural empathy will serve as a stimulus for further research and theoretical conceptualization on ethnocultural empathy. Also, the use of a more diverse sample, particularly in terms of racial/ethnic background, education, and age, would allow for greater generalizability of the results. Additionally, the SEE could be examined in relation to measures of prejudice and racism. Further research may also expand the construct of ethnocultural empathy to other populations. For example, it may be feasible to use this scale to develop measures of empathy toward people of different sexual orientations, ability levels, and ages. Although several measures examining gender and sex exist (e.g., the Attitudes Toward Women Scale; Spence, Helmreich, & Stapp, 1973), the current line of research could be expanded by investigating empathy toward the opposite sex or gender. The possible variations of this scale could be further expanded to explore ethnocultural empathy toward specific groups or between specific groups, depending on the sample used (e.g., empathy toward African Americans or Asian Americans).

An assessment measure of ethnocultural empathy can make an essential contribution to the growing multicultural movement in the United States. Within psychology, appraisal of counselors' empathy toward racially and ethnically diverse clients, in addition to their knowledge, awareness, and skills, may add significantly to their multicultural therapeutic proficiency. Within the field of multicultural counseling, cognitive and affective empathy have been found to relate to counselors' self-reported multicultural counseling knowledge and awareness (Constantine, 2000). Thus, validating the SEE with a sample of counselors would be an additional area of exploration for future research.

We also foresee that the SEE can be used in multiple venues outside of the counseling setting. A growing number of education institutions have concentrated on issues of diversity. Such programs as Collegetown, by the National Conference for Community Justice, attempt to increase student and faculty awareness about multiculturalism and to assist them in their attitude change. Given that previous research (e.g., Batson et al., 1997) has demonstrated that the induction of empathy is effective in changing individuals' attitudes toward stigmatized groups, pre-and postmeasures of ethnocultural empathy may be an excellent tool for the evaluation of future programming.

Blanchard et al. (1994) suggested that "the elasticity of attitudes and opinions about racism may derive from the lack of interracial experience of Whites in America" (p. 996). There is hope that the efforts toward desegregation and improvement of interracial relationships within psychology and society as a whole will be reflected in positive changes in individual levels of ethnocultural empathy. Businesses, private and government based, are also attempting to increase their efforts to include multicultural awareness training (Bell, 1999). In fact, the boom in diversity training owes its existence to the Supreme Court and Equal Employment Opportunity Commission rulings that companies can legally protect themselves on discrimination charges through employee training on race, gender, and sexual orientation issues (Adelman, 2000). In light of these new developments, a measure of ethnocultural empathy may be a necessary tool in assessing workers' attitude change and the evaluation of diversity training outcomes.

The racial and ethnic diversity in our classrooms, neighborhoods, families, counseling rooms, and society in general will continue to be at the heart of research and training in psychology. The SEE, as a unique measure of cultural empathy toward individuals from racial and ethnic backgrounds other than one's own, can be a valuable tool to aid such efforts.

References

A nation of many millions of divisions. (2001, April 16). *Newsweek International,* p. 6.

Adelman, K. (2000, February 2). Shakespeare can make diversity training palatable. KRTBN Knight-Ridder Tribune Business News: Bridge News, New York, Section B.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332.

Batson, C. D. (1987). Self-report ratings of empathic emotion. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 356–360). New York: Cambridge University Press.

Batson, C. D., Batson, J. G., Todd, R. M., Brummett, B. H., Shaw, L. L., & Aldeguer, C. M. R. (1995). Empathy and the collective good: Caring for one of the others in a social dilemma. *Journal of Personality and Social Psychology, 68,* 619–631.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., et al. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology, 72,* 105–118.

Batson, C. D., Turk, C. L., Shaw, L. L., & Klein, T. R. (1995). Information function of empathic emotion: Learning that we value the other's welfare. *Journal of Personality and Social Psychology, 68,* 300–313.

Bell, J. (1999, December 10). Diversity training: Long-term commitment for companies. *Business First—Columbus,* p. 34.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Blanchard, F. A., Crandall, C. S., Brigham, J. C., & Vaughn, L. A. (1994). Condemning and condoning racism: A social context approach to interracial settings. *Journal of Applied Psychology, 79,* 993–997.

Blanchard, F. A., Lilly, T., & Vaughn, L. A. (1991). Reducing the expression of racial prejudice. *Psychological Science, 2,* 111–115.

Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika, 52,* 345–370.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equations models* (pp. 136–162). Newbury Park, CA: Sage.

Buie, D. H. (1981). Empathy: Its nature and limitations. *Journal of the American Psychoanalytic Association, 29,* 281–307.

Carell, L. J. (1997). Diversity in the communication curriculum: Impact on student empathy. *Communication Education, 46,* 234–244.

Cattell, J. L. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1,* 245–276.

Constantine, M. (2000). Social desirability attitudes, sex, and affective and cognitive empathy as predictors of self-reported multicultural counseling competence. *Counseling Psychologist, 28,* 857–872.

Cross, W., & Phagen-Smith, P. (1996). Nigrescence and ego identity development. In P. B. Pedersen, J. G. Draguns, W. J. Lonner, & J. E. Trimble (Eds.), *Counseling across cultures* (pp. 108–123). Thousand Oaks, CA: Sage.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology, 24,* 349–354.

Cui, J., van den Berg, S., & Jiang, Y. (1998). Cross-cultural adaptation and

ethnic communication: Two structural equation models. *Howard Journal of Communications, 9,* 69–85.

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

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126.

Davis, M. H. (1994). *Empathy: A social psychological approach.* Madison, WI: Brown & Benchmark.

DeAngelis, T. (2001). All you need is contact. *Monitor on Psychology, 32*(10), 61.

Dovidio, J. F., Allen, J. L., & Schroeder, D. A. (1990). Specificity of empathy-induced helping: Evidence for altruistic motivation. *Journal of Personality and Social Psychology, 59,* 249–260.

Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology, 43,* 261–274.

Dutton, T. A. (2001, June 18). "Violence" in Cincinnati. *The Nation, 24,* 6.

Firebaugh, G., & Davis, K. E. (1988). Trends in anti-Black prejudice, 1972–1984: Region and cohort effects. *American Journal of Sociology, 94,* 251–272.

Frey, W. H., & Farley, R. (1996). Latino, Asian, and Black segregation in the U.S. metropolitan areas: Are multi-ethnic metros different? *Demography, 33,* 35–50.

Hatcher, L. (1994). *Step-by-step approach to using the SAS system for factor analysis and structural equation modeling.* Cary, NC: SAS Institute.

Helms, J. E. (1995). An update of Helms' White and people of color racial identity models. In J. G. Ponterotto, J. M. Casas, L. A. Suzuki, & C. M. Alexander (Eds.), *Handbook of multicultural counseling* (pp. 191–198). Thousand Oaks, CA: Sage.

Hoffman, M. L. (2000). *Empathy and moral development.* New York: Cambridge University Press.

Houston, D. A. (1990). Empathy and the self: Cognitive and emotional influences on the evaluation of negative affect in others. *Journal of Personality and Social Psychology, 59,* 859–868.

Ivey, A. E., Ivey, M. B., & Simek-Downing, L. (1987). Individual and cultural empathy. In E. Ivey, M. B. Ivey, & L. Simek-Morgan (Eds.), *Counseling and psychotherapy: Integrating skills, theory, and practice* (2nd ed., pp. 91–118). Englewood Cliffs, NJ: Prentice Hall.

Junn, E. N., Morton, K. R., & Yee, I. (1995). The "Gibberish" exercise: Facilitating empathetic multicultural awareness. *Journal of Instructional Psychology, 22,* 324–329.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23,* 187–200.

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35,* 401–416.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Kong, D. (2001, December 3). U.S. Census shows segregation of races. *Columbia Missourian,* p. 4A.

LaFromboise, T. D., Coleman, H. L., & Hernandez, A. (1991). Development and factor structure of the Cross-Cultural Counseling Inventory—Revised. *Professional Psychology: Research and Practice, 22,* 380–388.

Larson, L. M., Suzuki, L. A., Gillespie, K. N., Potenza, M. T., Bechtel, M. A., & Toulouse, A. L. (1992). Development and validation of the counseling self-estimate inventory. *Journal of Counseling Psychology, 39,* 105–120.

Lichtenberg, J., Bornstein, M., & Silver, D. (1984). *Empathy.* Hillsdale, NJ: Erlbaum.

MacCallum, R. C., Browne, M., & Sugawara, H. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149.

Maruyama, G. M. (1998). *Basics of structural equation modeling.* Thousand Oaks, CA: Sage.

Massey, D. S., & Denton, N. A. (1987). Trends in the residential segregation of Blacks, Hispanics, and Asians: 1970–1980. *American Sociological Review, 52,* 802–825.

Miville, M. L., Gelso, C. J., Pannu, R., Liu, W., Touradji, P., Holloway, P., & Fuertes, J. (1999). Appreciating similarities and valuing differences: The Miville–Guzman Universality–Diversity Scale. *Journal of Counseling Psychology, 46,* 291–307.

Newcomb, M. D. (1994). Drug use and intimate relationships among women and men: Separating specific from general effects in prospective data using structural equation models. *Journal of Consulting and Clinical Psychology, 62,* 463–473.

Omdahl, B. L. (1995). *Cognitive appraisal, emotion, and empathy.* Mahwah, NJ: Erlbaum.

Parson, E. R. (1993). Ethnotherapeutic empathy (EthE): II. Techniques in interpersonal cognition and vicarious experience across cultures. *Journal of Contemporary Psychotherapy, 23,* 171–182.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Paulhus, D. L. (1988). Assessing self-deception and impression management in self-reports: The Balanced Inventory of Desirable Responding. (Manual available from the author at the Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Y7)

Paulhus, D. L. (1991). Balanced Inventory of Desirable Responding. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 37–41). San Diego, CA: Academic Press.

Ponterotto, J. G., Rieger, B. P., Barrett, A., Sparks, R., Sanchez, C. M., & Magids, D. (1996). Development and initial validation of the multicultural counseling awareness scale (pp. 247–282). In G. R. Sodowsky & J. C. Impara (Eds.), *Multicultural assessment in counseling and clinical psychology.* Lincoln, NE: Buros Institute of Mental Measurements.

Ponterotto, J. G., Sanchez, C. M., & Magids, D. M. (1991, August). *Initial development and validation of the Multicultural Counseling Awareness Scale.* Paper presented at the 99th Annual Convention of the American Psychological Association, San Francisco, CA.

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67,* 741–763.

Quintana, S. M. (1994). A model of ethnic perspective-taking ability applied to Mexican-American children and youth. *International Journal of Intercultural Relations, 18,* 419–448.

Quintana, S. M., Castaneda-English, P., & Ybarra, V. C. (1999). Role of perspective-taking abilities and ethnic socialization in development of adolescent ethnic identity. *Journal of Research on Adolescence, 9,* 161–184.

Quintana, S. M., Ybarra, V. C., Gonzalez-Doupe, P., & Baessa, Y. D. (2000). Cross-cultural evaluation of ethnic perspective-taking ability: An exploratory investigation with U.S. Latino and Guatemalan Latino Children. *Cultural Diversity & Ethnic Minority Psychology, 6,* 334–351.

Ridley, C. R., & Lingle, D. W. (1996). Cultural empathy in multicultural counseling: A multidimensional process model. In P. B. Pedersen & J. G. Draguns (Eds.), *Counseling across cultures* (4th ed., pp. 21–46). Thousand Oaks, CA: Sage.

Rogers, C. R. (1959). A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science. Study 1. Conceptual and systematic: Vol. 3. Formulations of the person and the social context* (pp. 184–256). New York: McGraw-Hill.

Russell, D. W., Kahn, J. H., Spoth, R., & Altmaier, E. M. (1998). Analyzing data from experimental studies: A latent variable structural equation modeling approach. *Journal of Counseling Psychology, 45,* 18–29.

SAS Institute. (1992). *SAS/STAT User's Guide* (Vols. 1–2). Cary, NC: SAS Institute.

Schmitt, E. (2001, May 6). Segregation growing among U.S. children in metropolitan areas. *New York Times,* pp. N20, L28.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52,* 333–343.

Scotland, E. (1969). Exploratory studies of empathy. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 271–314). New York: Academic Press.

Scott, N. E., & Borodovsky, L. G. (1990). Effective use of cultural role taking. *Professional Psychology: Research & Practice, 21,* 167–170.

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.

Spence, J. T., Helmreich, R., & Stapp, J. (1973). A short version of the Attitudes Toward Women Scale. *Bulletin of the Psychonomic Society, 2,* 219–220.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Erlbaum.

Strayer, J. (1987). Affective and cognitive perspectives on empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 218–244). New York: Cambridge University Press.

Strayer, J., & Eisenberg, N. (1987). Empathy viewed in context. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 389–398). New York: Cambridge University Press.

Struch, N., & Schwartz, S. H. (1989). Intergroup aggression: Its predictors and distinctness from in-group bias. *Journal of Personality and Social Psychology, 56,* 364–373.

Sue, D. W., & Sue D. (1999). *Counseling the culturally different: Theory and practice* (3rd ed.). New York: Wiley.

Tabachnick, B., & Fidell, L. C. (1996). *Using multivariate statistics.* New York: HarperCollins College.

Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice.* Chicago: Aldine.

Unger, L. S., & Thumuluri, L. K. (1997). Trait empathy and continuous helping: The case of volunteerism. *Journal of Social Behavior and Personality, 12,* 785–800.

U.S. Census Bureau. (2000, January). *National population projections by race, Hispanic origin, and nativity.* Retrieved September 1, 2001, from *http://www.census.gov/population/www/projections/natsum-T5.html*

U.S. Census Bureau. (2001, January). *Resident population estimates of the United States by sex, race, and Hispanic origin: April 1, 1990 to July 1, 1999, with short-term projection to November 1, 2000.* Retrieved September 1, 2001, from http://www.census.gov/population/estimates/nation/intfile3-1.txt

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442.

Received April 23, 2002
Revision received September 13, 2002
Accepted September 16, 2002 ■