Spring 2016

An examination of the psychometric properties of and validity evidence for the Alliant Intercultural Competency scale

Elizabeth L. Smith
James Madison University

An Examination of the Psychometric Properties of and Validity Evidence for the Alliant Intercultural Competency Scale

Elizabeth L. Smith

A thesis submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Master of Arts

Department of Graduate Psychology

May 2016

FACULTY COMMITTEE:

Committee Chair: Dr. Tracy Zinn

Committee Members/ Readers:

Dr. Jeanne Horst

Dr. Matthew Lee

Acknowledgements

First and foremost, I would like to thank my wonderful advisor, Dr. Tracy Zinn. I have grown so much as both a student and a person because of you. You have encouraged me to step out of my comfort zone and pursue things that I did not think I think I was capable of. Because of that, I have learned so much about myself and even found a love for a field that I would have otherwise been unlikely to ever consider. Thank you for always believing in me and my abilities, and continuing to be such a strong sense of support and direction. Thank you for everything. You are amazing.

To my committee members, Dr. Matthew Lee and Dr. Jeanne Horst, I feel so fortunate to have had your support and guidance throughout this process. Matt, I have grown so much as a research assistant in your lab over the past four years. I would not have been nearly as equipped to handle a project like this without the opportunity to learn and pursue research topics related to social justice issues that I am passionate about. Thank you for continuing to provide guidance and mentorship to me throughout this process. I am so appreciative. And Jeanne, I am so grateful to have had the pleasure to meet you and work with you during my time in this program. Your teaching, motivation, kindness, and confidence in me has been unmatched. Your careful and thoughtful feedback has been so helpful to me throughout this process. You truly are an inspiring person, and I would not be the student I am today without your support. I feel like a better person for having known you.

To my supportive and patient friends, thank you for being there for me every step of the way. Thai Ong, you have been such an important part of my graduate experience. You are an amazing person. Thank you for always reminding me to keep things in

perspective. My best friends, Carly Morris and Katie Forehand, you are so important to me. Thank you for being such a strong support system and listening to me talk about research, even if you did not always find it as interesting as I do.

My amazing, supportive, fun-loving boyfriend, Sancho Sequeira, you bring so much happiness to my life. You have been the biggest support system of all this past year. Thank you for the many weekends you have traveled to see me, all of the hours on the phone, and for spoiling me with your amazing cooking and making everything else in my life easier throughout this process. I cannot thank you enough. You are so kind and hard-working. You motivate me and push me to be the best that I can, but also make sure I leave some time for myself. I am so lucky to have met you in this graduate program. Thank you, for everything.

To my wonderful mother, Nancy Smith, you have been there for me through all of the ups and downs. For 23 years, you have been nothing but supportive and loving, and I would not be who I am today without the guidance and opportunities you have provided me. You have sacrificed so much for me over the years, and I am so thankful. I am so proud to call you my mother. I love you to the moon, and back! And to my brother, Sandy, thank you for always being there for me. We are so different, but knowing that I can always count on you, no matter what, means the world to me. To my wonderful father who passed away too soon, I am so happy to have spent the past six years at James Madison University (and now, three more!), a school that you also loved. I hope that you would be proud.

Table of Contents

Acknowledgement..... ii
List of Tables.....v
List of Figures.....vii
Abstract.....viii
Introduction..... 1
Review of the Literature 2
Diversity and Globalization..... 2
Diversity in Higher Education..... 4
Diversity Curricula: Standalone vs. Infusion 7
Local vs. Global Diversity.....10
Intercultural Competency.....15
Existing Measures of Related Constructs.....16
The Alliant Intercultural Competency Scale (AICS) 26
Intercultural Competencies 27
Development of the AICS.....29
Current Study 32
Benson's Framework 32
Substantive Stage.....33
Structural Stage.....34
External Stage 34
Research Questions 35
Research Question 1 35
Research Question 2 35
Research Question 3 35
Research Question 4 35
Methods.....35
Participants 35
Participant Pool.....35
Assessment Day.....36
Mass Email/Social Media 36
Description of Sample.....36
Materials.....37
Alliant Intercultural Competency Scale, Revised.....37
External Measures.....37
Social Desirability Scale 38
Multicultural Experiences Questionnaire 38
Munroe Multicultural Attitudes Scale Questionnaire 38
Color-Blind Racial Attitudes Scale.....39
Demographics 39
Procedure.....40
Scale Development 40
Administration 40
Data Analysis.....40
Results.....41

Data Screening 41
Social Desirability 41
Exploratory Factor Analysis 42
Bivariate Correlation Analyses 46
Group Differences on AICS-R 48
Class Year 48
Major/Program of Study 49
Sociocultural Awareness Courses 51
Group Differences on External Measures 52
Class Year 52
Major/Program of Study 53
Sociocultural Awareness Courses 54
Discussion 54
Substantive Stage 55
Structural Stage 55
External Stage 57
Correlations with External Measures 57
Group Differences across Class Year 60
Group Differences across Majors/Programs of Study 61
Group Differences Based on Courses Taken 63
Implications for Higher Education and Assessment 66
Limitations and Suggestions for Future Research 70
Conclusions 75
Appendix 79
Tables and Figures 81
References 101

List of Tables

Table 1: Descriptive Statistics for AICS-R by Item (7-point Likert-type scale; 1 = Strongly disagree, 7 = Strongly agree) 81

Table 2: Inter-Item Correlation Matrix for AICS-R 83

Table 3: Pattern Coefficients of Item Loadings and Communalities.....86

Table 4: Structure Coefficients of Item Loadings.....89

Table 5: Factor Correlation Matrix.....92

Table 6: Means, Standard Deviations, and Pearson's r correlation coefficients for all measures 93

Table 7: Effect Sizes (R^2) for correlations among measures 94

Table 8: Means, Standard Deviations, and ANOVA results for comparisons across class year 95

Table 9: Means, Standard Deviations, and ANOVA results for comparisons across programs of study 96

Table 10: Means, Standard Deviations, and t-test results for comparisons across completed diversity courses among senior psychology majors.....98

List of Figures

Figure 1: Scree plot for AICS-R items…….. 99

Figure 2: Parallel analysis plot for AICS-R items……. 100

Abstract

The Alliant Intercultural Competency Scale was developed by Henderson et al. (2015) to measure intercultural competence in students in higher education. Henderson et al. outlined five domains representative of an interculturally competent professional: knowledge, communication, attitudes, professional practice, and negotiated space. In the current study, the AICS was revised (AICS-R) and then evaluated using Benson's (1998) framework for construct validity. Exploratory factor analysis results suggested a five-factor model strongly aligned with the five aforementioned domains; this provides support for the internal structure of the AICS-R. Scores from the AICS-R were correlated with external measures, and group differences in scores were examined to discern the validity evidence for the AICS-R. Overall, results from the current study provide support for the reliability and validity of the AICS-R as a measure of intercultural competence.

An Examination of the Psychometric Properties of and Validity Evidence for the Alliant Intercultural Competency Scale

In recent years, there has been an increase in diversity within the United States paired with an increase in international communication in the workplace. Due to this, higher education has focused on improving intercultural knowledge and communication skills in order to prepare graduates with the knowledge, skills, and abilities that will allow them to be interculturally competent professionals. In order to determine the preparedness for students to work with people from diverse cultural backgrounds, many quantitative measures have been developed to assess multicultural and intercultural competence (e.g., Munroe Multicultural Attitudes Scale Questionnaire; Munroe & Pearson, 2006; Global Awareness Profile; Corbitt, 1998). However, many of these instruments do not encompass the breadth of content and intricacies inherent to the construct of intercultural competence. Furthermore, there is a lack of consensus among experts regarding how intercultural competence should be defined.

Despite this lack of consensus, intercultural competence remains an important construct of interest among researchers. Henderson et al. (2015) performed an extensive literature review in order to aggregate competencies representative of a professional who embodies intercultural competence. The developers of the scale then constructed items to reflect these competencies. The preliminary analyses investigated the scale's psychometric properties, which were limited to reliability estimates, item intercorrelations, and descriptive statistics, due to the small and demographically limited sample. The current study will further evaluate the psychometric properties of the Alliant

Intercultural Competency Scale. In addition, validity evidence supporting the use of meaning of test scores will be evaluated.

Diversity and Globalization

Over the past several decades, the demographics of the United States have become increasingly ethnically and racially diverse (U.S. Census Bureau, 2011). Between 2000 and 2010, the population of the United States increased by 9.7%, with the fastest rate of growth in the Asian-identified population (43%). In terms of population numbers, the majority of the growth resulted from increases in those who identified their race as non-White and their ethnicity as Hispanic or Latino. All major race groups (defined by the Census Bureau as White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander) increased from 2000 to 2010. The White population was the only racial group to decrease in total proportion, accounting for 75% of the total population in 2000 and 72% in 2010. The majority of the growth in the White population between 2000 and 2010 was due to those who identified their race as White but indicated that they had Hispanic origins. This growth in racial and ethnic diversity is expected to increase in the coming years (U.S. Census Bureau, 2015). The United States is expected to become a "majority-minority" nation by 2044, in which the non-Hispanic White population will account for less than 50% of the total population. All other racial groups are expected to increase between 2014 and 2060.

Although the implications of increasing racial and ethnic diversity are imperative to consider, diversity is not limited to topics regarding solely racial and ethnic identity. Students and Americans in general have focused on race and ethnicity as the primary components when defining "diversity" (Chenoweth, 1998; Green, Callands, Radcliffe,

Luebbe, & Klonoff, 2009). In addition to race and ethnicity, the topic of diversity covers numerous other aspects of identity, including biological sex, gender identity, age, ability status, socioeconomic status, native language, accent, and religious beliefs (Hays, 1996, 2008). Each of these aspects of identity are related to varying degrees of privilege, or automatic societal advantages due to membership of a demographic group (e.g., Bonilla-Silva, 2006; Atewologun & Sealy, 2014; McIntosh, 1988). Additionally, aspects of identity representative of minority groups are often associated with different stereotypes, and may elicit implicit attitudes or biases (e.g., Devine, 1989; Dovidio, Kawakami, & Gaertner, 2002). These aspects of identity are not isolated, and can intersect in various ways to impact how a person both views the world and is viewed by others (e.g., Atewologun & Sealy, 2014).

Additionally, the interconnectedness and interdependence of countries and people, referred to as globalization, has increased over the past few years (World Health Organization, 2015). This is due in part to technological advancements (e.g., increased global communication via the Internet) and decreases in cost of the exchange of goods and capital across borders. Globalization has impacted economic activity as well as political, social, cultural, environmental, and technological issues (e.g., Arnett, 2002). Impacted by all of these realms, the workplace has become increasingly globalized. As a result, workers increasingly have to interact with others, both face-to-face and electronically, who have vastly different cultural backgrounds, values, or beliefs (Davidson, 2011).

According to the RAND Corporation, the three major trends that impact that 21st century workforce are changes in demographic patterns, economic globalization, and

technological advancements (RAND, 2004). In recent years, there has been an increase in women and minority workers who are entering the workforce. Furthermore, the labor force has been growing at a slower rate due to the end of the baby boomer generation and a pattern of earlier retirement for male workers. This has necessitated greater accommodations of families, workers with disabilities, and older workers. Technological advancements allow for these types of accommodations by enabling employees to have nonstandard working conditions, such as teleworking and/or flex time. Technology further facilitates greater communication and collaboration across borders for multinational corporations, which enable an increased ease and rate of transfer of ideas associated with economic globalization.

Diversity in Higher Education

Due to the aforementioned impact of globalization and shifting demographics, it is imperative for professionals to be able to interact and communicate competently with those from various backgrounds. Thus, it is important for students enrolled in institutions of higher education to develop these skills and abilities through their academic experiences. Institutions of higher education have placed increasing attention on incorporating diversity and sociocultural awareness student learning outcomes. Preparing students to enter an increasingly diverse and globalized workforce by promoting exposure to and interaction with those from other cultures can benefit students' personal growth and can also aid students in their professional growth. Many studies on data from college students show that heterogeneous learning environments are associated with positive learning outcomes (e.g., Antonio et al., 2004; Gurin, Dey, Hurtado, & Gurin, 2002; Narvaez & Hill, 2010).

Antonio et al. (2004) examined the effects of racially diverse groups on integrative complexity. Integrative complexity is the degree to which processes of thinking and reasoning involve variation and integration of different perspectives (Suedfeld, Tetlock, & Strufert, 1992). Those with low integrative complexity evaluate issues with a single dimension, whereas those with high levels of integrative complexity tend to integrate different perspectives and solutions. For example, consider the issue of whether the United States should increase the number of refugees accepted into the country. Those with low integrative complexity would view this situation as good or bad, without regard for nuance, differing perspectives, or alternative solutions. For those with high integrative complexity, the issue of increasing the number of refugees allowed in the United States may involve evaluating the issue from various perspectives (e.g., social, political, economic) and recognizing the trade-offs that would result from different decisions. Students with higher levels of integrative complexity tend to achieve higher grades in college (Gruenfeld & Hollingshead, 1993). Antonio et al. (2004) found that a Black student's inclusion in a classroom discussion, compared to a group composed only of White students, generally led to higher levels of integrative complexity. After participating in a discussion about a social issue, participants wrote an essay indicating their agreement or disagreement with a different social issue, and the essays were rated for integrative complexity. The presence of a minority opinion within classroom discussions was associated with stronger critical thinking skills and integrative complexity. Antonio et al. (2014) also examined the relationship between the self-reported number of racially diverse friends and scores on the integrative complexity measure. They found that this relationship was stronger than that found between the

classroom manipulation and integrative complexity, suggesting that prolonged interaction with racially diverse individuals tends to be associated with complex thinking.

These findings support research conducted by Gurin et al. (2002), who examined the relationship between informal interaction and diversity within the classroom among racially and ethnically diverse groups (i.e., African American, Asian American, Latino/a, and White students). They found that greater classroom diversity and informal interaction across diverse demographic groups of students were associated greater intellectual and motivational aspects of learning. These aspects of learning were measured through self-reported ratings of achievement motivation and intellectual self-efficacy, aspirations for graduate school, importance of original writing and creativity, and satisfaction with intellectual value and rigor of coursework. This suggests that informal interaction among different racial and ethnic groups during college groups may help cultivate students' academic and social growth.

Similarly, other data show that when college students report having broad multicultural experiences and interests, there are many positive outcomes. For example, diverse multicultural experiences and attitudes have been associated with growth mindset (Narvaez & Hill, 2010). Students who display attitudes corresponding to growth mindset tend to report that through personal endeavors and experiences, personal qualities and ability levels can be changed. Attitudes consistent with growth mindset have been associated with numerous positive outcomes of academic achievement (Dweck, 2007). This pattern of thinking allows individuals to engage in behaviors that will improve their levels of ability, which can lead to an enthusiasm for learning and an ability to confront challenges (Dweck, 2010). Additionally, those who hold a growth mindset tend to show

more resilience when faced with challenges and seek out opportunities to learn (Dweck, 2012). Moreover, experiences of and positive attitudes toward multiculturalism were predictive of attitudes consistent with growth mindset, both in intelligence and personality (Narvaez and Hill, 2010). Benefits of a growth mindset have been shown in various other realms, including the workplace, sports, and music (Dweck, 2007). Thus, fostering a growth mindset in higher education may help students be better prepared for experiences following their academic career.

Although these studies support the idea that experiences with people from diverse backgrounds are beneficial to student learning and growth throughout students' educational careers, many universities remain fairly homogenous in terms of race and ethnicity, in both student and faculty demographics. Because of such homogeneity, to engage in more diverse experiences often requires students to seek out opportunities that promote interaction with others from different cultural backgrounds. In higher education, such opportunities are often housed in curricula that incorporate diversity and sociocultural content covering various types of cultures, identities, and experiences.

Diversity Curricula: Standalone vs. Infusion

As previously noted, institutions of higher education have recognized the importance of developing students' knowledge of sociocultural content and intercultural competency skills, an awareness reflected in the increase in diversity courses offered to students (Zalaquett, Foley, Tillotson, Dinsmore, & Hof, 2008). Many of these courses are "standalone" courses in the sense that the course is developed specifically around diversity or sociocultural topics (Zalaquett et al.). These courses may include multicultural psychology courses, courses focused on issues related to women and

gender, Black psychology, or other topics focused broadly either on diversity issues or on a specific demographic group. However, some universities are taking an "infusion" approach by incorporating sociocultural and international content into multiple courses across curriculums and disciplines (Nelson Laird, 2014). This involves infusing sociocultural content into courses that are primarily focused on a different topic, such as industrial/organization psychology or developmental psychology.

The APA Guidelines for the Undergraduate Psychology Major were revised in 2013 to endorse the infusion approach. These guidelines promote recognizing, accepting, and appreciating sociocultural influences that can lead to differences among people. According to Carr (2007), the first researchers and professionals to emphasize the importance of incorporating a diversity course did so because of the push for incorporating sociocultural content in the curriculum on most campuses, with content often pertaining to power, privilege, oppression, and identity.

Although adding a diversity course to the curriculum is an important way to expose students to sociocultural content, Marchesani and Adams (1992) criticized the standalone diversity course approach and argued that course curricula should be reformed to infuse sociocultural content throughout the program of study. The Association of American Colleges and Universities (2005) argued for the infusion approach of incorporating content as the best method for achieving student-learning goals associated with diversity and sociocultural influences. When standalone diversity courses become an additional requirement for students during their program of study, the objectives can become marginalized. Students may view the course as just an additional class necessary for the completion of their degree. Thus, students may be resistant to the content, which

is often already challenging to students' views and beliefs. Infusing diversity content throughout the program of study promotes the linkage of sociocultural content to other objectives and content of the program (APA, 2013).

The APA noted that incorporating a standalone diversity course in the curriculum can be beneficial to student learning and a powerful way to address sociocultural content, and thus they do not advise against it as a requirement. However, infusing content throughout the curriculum can lead to a more enlightened and integrated understanding of the content (provided that faculty are actually skilled and/or trained in teaching about and infusing diversity content). Perhaps incorporating both approaches within the curriculum could be more beneficial than taking either approach alone. However, some scholars argue having both a standalone course dedicated to diversity and infusing content throughout the rest of the curriculum best helps to reinforce content so that diversity issues are not marginalized, but students are also provided a richer experience through in-depth topics that can be addressed in a standalone course. Dunn et al. (2008) discuss differences in infused and standalone courses more broadly, in terms of psychology major curriculum as a whole. Curricular structures that offer standalone content courses without providing substantial unification among content across courses put students at risk of viewing the field as fragmented and may make it difficult for students to draw connections across various domains and subfields within the same discipline. Standalone diversity courses are no exception to this risk, and may be even more at risk due to the sensitivity and complexity of the content. Dunn et al. (2008) argue that curricula should include standalone courses that address challenging and sensitive content, but also infuse content related to culture and diversity into other content courses. The ways in which

diversity content are addressed influence student learning and student experiences in the classroom. In turn, this could influence cultivation and development knowledge, skills, and abilities that would allow students to function effectively in an increasingly diverse society.

Local vs. Global Diversity

Although institutions of higher education have increased their emphasis on sociocultural content in students' educational experiences, there are three main challenges that universities face in the process (Henderson et al., 2015). The first challenge is the division between multicultural education and international training. In education there is typically a focus on either local or global dimensions of diversity and cultural competency; there is rarely a focus on both. Multicultural education tends to focus on local dimensions, or domestic diversity. The focus of multicultural education pertains largely to issues of diversity and cultural differences prevalent within the United States (or a person's country of origin). International training tends to focus on global dimensions, or cross-cultural trends and systems (Olson, Evans, & Shoenberg, 2007). This type of training focuses on relationships among different nations and cultures, particularly outside of or compared to the United States. The second challenge is the intricacy of the required knowledge, understanding, attitudes, and skillset that enables successful communication both within local cultures and across international cultures. This might involve applying knowledge of nation-state relationships and global systems to better understand how dynamics of power and oppression are different within a foreign country. Thus, teaching about and providing comprehensive training for students in intercultural competency is challenging due to the intricacy of the construct. The third

challenge is related to gaining support and attention from institutions to emphasize learning new competencies and assessing those competencies is difficult.

As previously, noted, most educational paradigms tend to focus on either local or global dimensions of diversity, rather than integrating them. This divide is a result of different educational paradigms, support of different offices, different emphases across disciplines, and different professional associations (Olson et al., 2007). Multicultural education emerged in the 1960s because of the social and political movements of that era, and focused on ethnicity, race, and social change. Drawing largely on ethnic studies, women's studies, history, and the social sciences, eventually, multicultural education grew to encompass more than race and ethnicity, and began to incorporate issues associated with gender, sexual orientation, and ability status. Finally, the field expanded even more to examine interrelationships among these different dimensions of identity, referred to as intersectionality (Crenshaw, 1989).

International education emerged from a focus on international relations following the World War II and Cold War era. Originally, it focused largely on Western and European culture, motivated by a need for understanding national relations, promoting peace, and developing strategic national plans. Internationalization grew in fields such as history, geography, language, international relations, and area studies. Recently international training has begun expanding its breadth from Western/European perspectives to a much more globally inclusive focus.

At an institutional level, there is a divide between international and multicultural agendas. For both multicultural and international education, initiatives are often the focus of student affairs or co-curricular offices, rather than academic affairs. Often there

is not a single office for either type of education, and initiatives are developed within individual programs in isolation from others. Although one office may focus on initiatives from a multicultural perspective, another may focus on an international perspective.

Faculty motivations also differ depending on whether they are aligned with multicultural education or internationalization. Faculty involved with multicultural education are often academics that are from cultural minority backgrounds, and they focus their efforts on societal change and social justice (Cornwell & Stoddard, 1999). In contrast, internationalization is often the focus of White academics interested in the comparison of the United States to other nations, and it often criticized for its lack of a social justice component. These differences can lead to tensions among scholars regarding how internationalization and multicultural education should be operationalized and addressed within higher education.

Student motivations for taking classes on diversity also differ (Hayward & Siaya, 2001). Multicultural education tends to attract students who embrace social justice and equity, and who may have been motivated by personal experiences of marginalization. International training tends to attract students who want to learn about a different culture, become proficient in another language, or broaden their career trajectories with a better understanding of global perspectives.

There is also often a lack of understanding across the two approaches (Olson et al., 2007). As described previously, multicultural education and internationalization are born out of separate movement/contexts, the focus of different offices, valued by different academic departments and disciplines (and thus, faculty), and attract students

with different interests and goals. This leads to a divide that has resulted in a lack of collaboration and understanding between the two paradigms, and resulted in tensions regarding what should be the focus of educational learning outcomes.

However, both internationalization and multicultural education share many similarities (Olson et al., 2007). Both paradigms strive to instill similar values in students, particularly the appreciation and awareness of diversity and improvement of interpersonal and intergroup communication among people of various backgrounds. Additionally, both approaches draw on research and constructs from various disciplines to inform theory and practice. They also necessitate collaboration among various departments, such as academic affairs (curriculum) student affairs (admissions, co-curricular organizations and activities), and business (financial and funding).

Multicultural education and international training are also faced with similar challenges (Olson et al., 2007). In both, there is a lack of consensus among those in the academic community regarding the definitions of terms and conceptualization of constructs. The goals and values of both paradigms tend to be seen by faculty and students as curricular "add-ons," rather than an integral aspect of education. They are often viewed in terms of an additional, lower-level course requirement. A similar view is held at the institutional level, where content promoted by both approaches is seen as a concern for only one or two offices that focus exclusively on diversity, rather than an a concern of the entire institution.

Whereas both approaches share similar perspectives in terms of learning outcomes, there remains a lack of consensus among academics and practitioners regarding the most important student learning outcomes. Olson et al. (2007) noted many

similarities across the outcomes most often presented by both fields. A roundtable led by the American Council on Education was asked to provide a list of student learning outcomes that are shared across the two approaches. These outcomes were categorized under knowledge (e.g., "Understand the various/different cultures and how culture is created;" "Recognize stereotypes and where they come from"), attitudes (e.g., "Appreciate difference; value and acknowledge other cultures as legitimate;" "Demonstrate open-mindedness and an understanding of complexity", and skills (e.g. "Develop and use intercultural communication skills;" "Internalize and apply cultural understandings and knowledge;" Olson et al., p. 13).

Although there is substantial overlap in the goals, values, and challenges of both multicultural education and international training, they remain largely distinct and tend to function separately within higher education. Henderson et al. (2015) raised the question of whether one can truly be culturally competent if they have learned about diversity from only either a local or global perspective. The criticism is that, if someone is trained only from local diversity perspective, they may not be able to function effectively when facing professional issues stemming from a more global aspect of diversity. Conversely, if someone is trained largely from a global perspective, will they be prepared to effectively deal with issues primarily focused on in areas of local diversity? Someone who is competent in both domains would be able to apply their understanding of multicultural issues (e.g., dynamics of power and privilege) to situations where cultural values may be different from their own.

Intercultural Competency

Henderson et al. (2015) proposed the merging of local (generally conceptualized as multicultural education) and global (generally conceptualized as international training) perspectives into a construct they refer to as intercultural competency. Henderson et al. described intercultural competence as a blend of the local and global perspectives that "bridges the paradigms of multicultural competency and international training" (p. 20). No measures to date have merged these two areas of the intercultural competency construct.

Deardoff (2006) addressed the complexity of defining the construct of intercultural competency as an explanation for why few universities identify the construct as a specific student learning outcome, and even fewer attempt to assess development in the domain. Deardoff (2006) utilized a Delphi approach to develop consensus among experts to provide a definition of intercultural competence and the components associated with it. The Delphi approach is a method of structured communication between subject experts to provide information about a complex topic (Linstone & Turoff, 1975). Subject experts offer their judgment on a topic in order to reach some consensus regarding definition or interpretation in this iterative process. The greatest consensus among expert scholars describes intercultural competency as, "the ability to communicate effectively and appropriately in intercultural situations based on one's intercultural knowledge, skills, and attitudes" (Deardoff, 2006, p. 194). Scholars agreed that there are many components, not just one, that encompass intercultural competency.

In the Delphi study, Deardoff (2006) also sought to identify the best methods of assessing intercultural competence, and found that different scholars suggested different

methods. The most commonly suggested methods were student interviews, student papers and presentations, student portfolios, observation of students by others/host culture, professor evaluations, and pretests and posttests. Furthermore, using both qualitative and quantitative assessment methods was preferred to just one type of measure.

Existing Measures of Related Constructs

Researchers often utilize publicly available scales when attempting to assess intercultural competency. Henderson et al. (2015) outlined the four main strategies used when assessing intercultural competency: "(a) multicultural knowledge, self-awareness, and skills for working across cultures; (b) intercultural skills in working across international borders (such as, flexibility, sensitivity, open-mindedness, perceptual acuity, personal autonomy, empathy, and respect); (c) behavioral assessments; and (d) vignette assessments" (p. 22). Scales focusing on multicultural dimensions largely reflect content related to local diversity, whereas scales that reflect intercultural skills tend to reflect content related to global diversity. Behavioral assessments attempt to assess more than just attitudes, beliefs, and knowledge. Rather, they focus on actual behaviors. They may involve self-reflection on ability to competently function and rating scales for individuals to evaluate their observations of others (Henderson, Horton, Saito, & Shorter-Gooden, 2014). Vignette assessments involve written, video, cartoon, or musical scenarios in which participants indicate either how they or a character in the vignette should respond (Barter & Reynold, 2000; Finch, 1987). This requires a level of participant engagement that self-report measures are usually unable to achieve (Finch, 1987; Wilks, 2004).

Although behavioral assessments and vignette assessments may offer some advantages over multicultural and intercultural measures that rely on self-report, self-report methods are often used due to their availability and general ease of administration. Furthermore, limitations that exist due to the nature of self-report methods of affective assessment. Self-report measures are subject to socially desirable responding, or the tendency of respondents to attempt to portray themselves in a favorable manner or align with what they believe is expected of their response (Edwards, 1953). Importantly, respondents may not have accurate self-perceptions, which could lead to biased responses. For example, self-report measures tend to reflect expectations for behavior rather than the actual behavior that would be carried out.

Constantine and Ladany (2000) found that White students tended to overestimate both their own multicultural skills and knowledge when asked to evaluate themselves. In this study, the researchers coded responses to evaluate the extent to which respondents incorporated racial or ethnic issues into multicultural case conceptualizations. Respondents also completed four measures of self-reported multicultural counseling competence. There was no significant relationship between any of the four self-report quantitative measures and the case conceptualizations. Furthermore, mean scores on the self-report measures suggested that respondents felt they were multiculturally competent, but ratings on the case conceptualizations fell in the low to middle ranges of competence. Thus, respondents appeared to have overestimated their level of multicultural competence.

There are further limitations specific to measures of sociocultural and multicultural awareness and competence that are important to address. The divide in

educational paradigms and training mentioned previously is reflected in assessment of intercultural competency as well. Many measures of multicultural competency are developed to assess those in the field of counseling psychology, and therefore may not be appropriate for those in other professional fields of psychology and in general.

Several measures of diversity, including multicultural and intercultural competencies, are discussed below. These measures are evaluated first on the content of the measure represented by the items and how well the measures cover the theoretical constructs. If the content addresses some of the aforementioned criticisms, then psychometric properties and validity evidence are discussed. And finally, accessibility to the instrument (publicly available or commercially available) is discussed. It is important to note that this list is not exhaustive, as there are hundreds of measures developed to measure various constructs associated with diversity and multicultural or intercultural competence. The purpose of this review is to provide examples of common limitations of existing measures.

Multicultural Attitude Questionnaire (Giles & Sherman, 1982). This survey was created to measure multicultural attitudes in those training to become teachers. The measure was developed to reflect attitudes regarding family, friends, social distance, opinions on group differences, and acceptance of diverse individuals. Although this measure covers a variety of different aspects of diversity, it measures only attitudes toward diverse individuals rather than competence of differences based on demographic groups.

Transcultural Self-Efficacy Tool (Jeffreys, 2000). The measure was developed to assess nursing students' perceived self-efficacy for working with clients from diverse

populations. This scale was designed to measure perceived competencies, rather than only attitudes. Although this scale covers a wide variety of cultures, it is limited in the scope of the measure, in that it does not cover various aspects of diversity such as gender, sexual orientation, age, etc. Additionally, the scale is intended for nursing students to assess their self-perceptions of efficacy in nursing situations. Thus, it is limited in its use for broader populations of students (i.e., non-nursing students).

Cross-Cultural Adaptability Inventory (CCAI; Kelley & Meyers, 1995). This measure was developed to assess students' ability to interact with members of different cultures and adapt to life in a new culture. The authors state that the measure is intended for any culture with a broad range of populations. The CCAI has been widely used; however, a recent examination of its psychometric properties indicated various concerns about its use (Davis & Finney, 2006). The original scale consisted of a hypothesized four-factor structure. However, confirmatory factor analyses indicated that this model was misspecified. Subsequent exploratory factor analyses were conducted to examine the underlying structure of the scale. Unfortunately, none of the analyses resulted in an interpretable solution. Thus, Davis and Finney (2006) suggested that the measure should not be used to assess cross-cultural adaptability until subsequent research is done on the scale.

Miville-Guzman University-Diversity Scale (Miville et al., 1999). This scale purports to measure attitudes and awareness of both differences and similarities among others. Examination of the reliability and validity of the scale provide potential support for its use. However, the scale is limited in scope as it includes items that address only issues pertaining solely to race and ethnicity.

Scale of Ethnocultural Empathy (SEE; Wang et al., 2003). This instrument was developed to measure empathy directed toward racial and ethnic groups that are different from one's own. Empathy was operationalized as feelings, expressions, perspective taking, acceptance of differences, and awareness. This measure was developed in light of the increased emphasis on diversity and multiculturalism within counseling psychology. Although there is some evidence to support reliable and valid interpretations of scores, the scale is limited in scope. The SEE focuses exclusively on empathy toward racial and ethnic groups.

Culturally Diverse Experiences and Comfort Questionnaire (Eliason & Raheim, 2000). This scale assesses how comfortable respondents feel about working with individuals from various cultural backgrounds. Participants are provided with 14 different cultural groups and asked to indicate the extent to which they would feel comfortable working with individuals from each. If they report that they are uncomfortable working with a particular group, they are asked to explain the reason for their discomfort. Although this scale covers a broad scope of cultural differences, it does not assess respondents' competency when working with individuals in specific situations or ability to apply knowledge of cultural differences. Rather, it focuses on their perceived comfort. Based on the content of the scale, it appears to be more of an external measure of prejudice rather than cultural competence. Furthermore, this scale was developed specifically for undergraduate pre-nursing students. Thus, it may not be applicable to students across different fields.

Cross-Cultural World Mindedness Scale (CCWMS; Der-Karabetian, 1992). Although the scale has been administered in 10 nations over several years, there is little

information publicly available about the psychometric properties of the scale or the validity of interpretations of scores. Alpha coefficients reveal adequate internal consistency across scores from samples from different countries (Der-Karabetian & Metzer, 1993). An examination of the relationship between political party and scores on the CCWMS indicated support for criterion validity in some students (sample taken from Southern California) but not others (sample taken from Texas). Further information regarding the validity of the scale is not available. The scale is described as measuring attitudes toward various components of diversity, including race, religion, immigration, patriotism, war, government, education, and economics. Example items indicate that the scale measures opinions and attitudes rather than competencies; however, the full scale is not publicly available.

Global Awareness Profile (GAP; Corbitt, 1998). The GAP measures global awareness and knowledge. The purpose of this instrument is to assess the extent to which an individual is knowledgeable about various aspects of the world. The test includes questions for various subcategories of geography, including Asia, Africa, North America, South America, Middle East, Europe, and Global. It covers contextual information regarding environment, culture, politics, geography, religion, socio-economic status, and global. Questions include content that would be obtained through visiting a country for more than one month through study, experience, conversation, or the news. Although the scale covers many global dimensions, it is only commercially available and may not be useful for assessing intercultural competence in students who have not had the opportunity to spend an extended amount of time in another country. Furthermore,

although it is stated that reliability and validity evidence have been examined, this information is not listed on the GAP website.

Global Perspective Inventory (GPI; Braskamp, Braskmap, Merrill, & Engberg, 2013). The GPI was developed to assess cognitive, intrapersonal, and interpersonal domains of intercultural competence. Each domain is assessed by two scales, one that reflects cultural development and another that reflects intercultural communication. The GPI was designed for individuals of any age or cultural group. Additionally, three forms of the GPI are available for specific groups: the general student form, new student form, and study abroad posttest form. The GPI appears to measure a breadth of content, is potentially usable by a broad range of populations, and its psychometric properties and validity evidence have been evaluated extensively. However, this instrument is limited by the fact that it is only commercially available. Not only do institutions have to pay for its use, but also results are only available in report form rather than as raw data. Thus, even if institutions or organizations are able to afford a commercially available assessment instrument, the results are not available in a form that enables researchers to compare data to other results or perform other analyses.

Intercultural Development Inventory (AIA; Hammer, Bennett, & Wiseman, 2003). This instrument was developed to measure intercultural competence and has been used in educational institutions and organizations across over 30 countries. Examinations of validity evidence suggest that the scale is not culturally biased and exhibits strong content and construct validity. The IDI has also shown strong predictive validity in both educational and corporate sectors. Cross-cultural generalizability has been shown both domestically and internationally. Although the test appears to address many of the

limitations noted in other scales, the IDI is only available commercially and individual items cannot be accessed or used by researchers.

Munroe Multicultural Attitudes Scale Questionnaire (MASQUE; Munroe & Pearson, 2006). The MASQUE was developed to assess multicultural attitudes based on Banks and Banks' (1995) domains of knowledge, empathy, and active experience. This instrument covers a broader range of aspects of diversity than some of the previously mentioned scales, including race, religion, sexual orientation, gender, and language. In addition, this instrument covers several aspects important for competency, including awareness of, caring about, and acting on issues related to multiculturalism. However, this scale focuses on issues of domestic diversity does not cover the global dimensions of intercultural competence.

Color-Blind Racial Attitudes Scale (CoBRAS; Neville, Lilly, Duran, Lee, & Brown, 2000). The CoBRAS was designed to measure color-blind racial attitudes, or beliefs that race and ethnicity do not matter and are not noticed when interacting with others. There are three subscales: unawareness of racial privilege, institutional discrimination, and blatant race issues. Although there is strong evidence to support the validity and reliability of scores from this scale, it is limited in scope. All items assess attitudes regarding race and ethnicity, and thus the many other aspects of identity are not included.

Multiculturalism Experiences Questionnaire (MEQ; Narvaez & Hill, 2010) The MEQ was constructed to measure multicultural experiences with and attitudes toward individuals from diverse groups. In its initial testing, the MEQ demonstrated construct validity through a negative relationship with measures of closed-mindedness.

In a subsequent study, the scores on the MEQ were negatively correlated with measures of rigid thinking and predictive of growth mindset and higher levels of moral judgment. However, the MEQ is limited in scope as it only includes items that assess attitudes and experiences with individuals of different races, cultures, and viewpoints.

In addition, Hays (2008) provided a review of several multicultural competence instruments used primarily in counseling psychology. Each of the instruments included in her review were developed to assess either how prepared an individual is or how effective a training program was at preparing counselors to be culturally sensitive in their practice. Several limitations are acknowledged, including the lack of psychometric data and evaluation provided for the measures. Hays noted that each of the scales is limited in the scope of the content they cover. Many of these measures focus exclusively or primarily on competence associated with racial or ethnic minorities, and fall short in terms of other demographic conceptualizations, such as gender, sex, sexual orientation, age, religion, and ability status (Arredondo et al., 1996). This is due largely to the conceptualization of multicultural counseling that has focused on race and ethnicity (Hays, 2008). Furthermore, the instruments are designed explicitly for use in counseling trainees or practitioners, and may not be appropriate for a wider variety of respondents.

In sum, there are three main overarching criticisms that limit the utility of the existing measures in the literature. First, scales of multicultural competency often lack a broad range of content. They tend to focus on dynamics of race and ethnicity. When this is the construct of interest, such measures may be appropriate. However, often instruments are selected in higher education or organizations to measure a more encompassing definition of multicultural or intercultural competence. Second, most of

the measures reviewed above include content that solely represents local aspects of diversity. These measures lack a global component. Although this is reflective of the educational paradigm in which many of these measures were based, this definition of diversity is limited if the goal is to assess competencies that will allow students to function with intercultural competence. Although some measures do include a wide range of global content, these measures tend to be commercially available. For commercially available instruments, data are provided in reports or as an aggregate rather than as raw individual scores. Lastly, it is noteworthy that very few scales include content related to dynamics of power and oppression, which are often integral to the educational goals. Some exceptions are still limited in that they measure these dynamics in terms of a single aspect of identity (e.g., race/ethnicity). Or, some measures purport to measure solely awareness of power and oppression, and no other constructs. Such dynamics are paramount to intercultural competence, and should be included in measures that intend to measure this and other similar constructs.

Due to these aforementioned criticisms of the various instruments, many organizations and universities opt to create their own assessment measures designed to fit their specific interests and the context of their organization (Henderson et al., 2015). Because of this, many of these measures are not evaluated with adequate psychometric data or normed with respondents outside of the organization. Thus, there is a need to develop and critically evaluate a scale that could be useful to a broad range of organizations. Furthermore, it is critical to develop an instrument to measure student learning of intercultural competency so that institutions of higher education can assess students' intercultural competency as they prepare to enter the workforce.

The Alliant Intercultural Competency Scale (AICS)

Henderson et al. (2015) created a new instrument to reflect their definition of intercultural competency, illustrated through an outline of competencies that represent the knowledge, awareness, attitudes, and skills associated with the ideal professional who is both locally and globally competent. It is important to note that there is not consensus among researchers regarding the definition of intercultural competence (Deardoff, 2006). Furthermore, there are many terms used to refer to similar constructs that vary across fields (Fantini, 2009). However, in order to develop or use any assessment instrument, it is imperative to define the concept of interest before proceeding to any subsequent step in assessment (Fantini, 2009). Thus, it was important for Henderson et al. (2015) to define the construct of intercultural competency before beginning the process of item development.

A team of researchers first developed the competencies associated with the ideal locally and globally competent professional (Henderson et al., 2015). Postsecondary institutions tend to rely on faculty discussion of the term, rather than referring to definitions and discussions found in the literature (Deardoff, 2006; Hunter, White, & Godbeym 2006). Henderson et al. (2015) overcame this limitation by relying on both literature and content experts to develop the competencies later used to construct items. These competencies were developed through extensive review of literature on both multicultural and international competency, reviewing and consolidating both multicultural and international competencies outlined by professional business (CILT, 2008), education (Banks et al., 2005), and psychology (APA, 2003) associations that

were publicly available, and evaluation of items from multicultural and international competency scales.

Alliant faculty experts and alumni located internationally reviewed, discussed, and edited initial drafts of the competencies. The drafts were also compared to competencies associated with Alliant's graduate level business, education and psychology programs. The drafts were also reviewed against the initial literature review to make sure that important concepts in the literature had not been dropped from the competencies throughout the review process.

Intercultural Competencies. Henderson et al. (2015) identified five domains believed to be intrinsic to the proposed construct of intercultural competency. They are "the ability to (a) learn and embrace intercultural knowledge; (b) communicate and relate interculturally; (c) maintain inclusive attitudes; (d) conduct professional practice effectively within and across cultures; and (e) work effectively to create and maintain negotiated space" (Henderson et al, 2015, p. 22).

In order to be competent in each of the aforementioned domains, Henderson et al. (2015) argued for several specific competencies. This includes the acquisition and integration of knowledge and research of both multicultural and international diversity, as well as the ability to employ cognitive flexibility and critical thinking skills that allow for the facilitation of negotiated spaces in which meaningful conversations and collaborations across cultures and backgrounds can take place. Negotiated space refers to a problem-solving and decision-making process in which viewpoints and norms of those from underrepresented groups are weighted evenly with those from dominant groups. Within these conceptualized negotiated spaces, those involved actively work to share

ideas from all members, take perspectives of other members, and foster culturally respectful communication. Henderson et al. outlined several specific competencies that are necessary to address in order to achieve effective communication within negotiated spaces:

"The specific competencies…are to demonstrate, through an embrace of at least two cultures, understanding of, respect for, curiosity about, and advocacy toward:

- Cultures and their interconnectedness
- One's own culture within a global and comparative context
- The ever-changing nature of globalization and the ways in which economic, political, cultural and technological changes affect local, national and transnational cultures and communities (common problems, collective efforts and movements/global solutions)
- Cultural differences and similarities (examined in a comparative historical and political context)
- Social and political construction of identities (gender, ethnicity, sexual orientation, religion, class, ability, etc.) and issues associated with multiple identities
- The relationship between power, knowledge, language (discourse) and privilege in a global context (Henderson et al., 2015, p. 22)."

In addition, Henderson et al. (2015) explained the necessity of having the ability to work and

collaborate with people from diverse backgrounds and cultures in order to provide professional services that are culturally competent within the aforementioned negotiated spaces. The specific associated competencies include the ability to:

- Connect, integrate, and apply general multicultural and international knowledge and research and scholarship competencies in new modes of professional practice and problem solving
- Transfer insights from one's own group/region/nation-state/culture to another to create new knowledge and understanding in areas of professional practice
- Communicate with people/clients of diverse backgrounds
- Work in diverse/international teams
- Mediate/resolve intercultural conflict in diverse work environments (Henderson et al. 2015, p. 22)

Items on the AICS were written to address these specific competencies within each of the five domains.

Development of the AICS. Henderson et al. (2015) reviewed literature to determine how to assess competencies outlined above. The overarching research question foundational to the approach to assessment was, "How can the aforementioned intercultural vision and competencies be assessed in student learning?" (Henderson et al., 2015, p. 22). The purpose of the scale was to assess self-perceived intercultural competence in students involved in higher education. Results could serve as an indicator of intercultural competence in the workplace and other professional environments.

The five domains outlined above associated with the developed competencies were used as the foundation for the scale development procedure. Several of the authors wrote scale items to reflect the competencies within each of the domains. For many of the concepts, several items were written with different wording in order to determine the best wording for the specific concept. Scale responses used a 5-point Likert type scale and ranged from strongly disagree to strongly agree. Additionally, the response option, "I am unclear on the meaning of this item" was included in the initial testing of the scale. The scale items were then grouped by common stems within each respective domain.

Team members, colleagues, and alumni experienced in scale development were selected via convenience sample to review and revise the initial draft of the scale, resulting in an 80-item scale. The authors recruited respondents through an online posting call to all Alliant University students to participate in the online survey. Paper surveys were also administered to students in intercultural competency related courses. Alliant University is a predominantly graduate university in the western United States. Raffles for four $100 Visa gift cards were used as incentives. The survey was administered to 81 graduate students. The authors analyzed the 80 items, and evaluated and eliminated statistically and logically weak, unclear, and repetitive items in order to develop a more parsimonious scale.

Using the descriptive statistics (mean, standard deviation, minimum values, maximum values, skew values, and frequency counts), the researchers identified items exhibiting ceiling and floor effects, large numbers of missing values, and inadequate values. Descriptive statistics were also used to identify which items among several written to illustrate the same concept performed best. Items that participants reported as

being unclear were also eliminated. In total, 27 items were deleted: 4 for unclear wording, 4 for redundant wording and weaker statistical properties, 13 for unclear wording and weak inter-item correlations, and 6 for overlap with a different domain. Inter-item correlations were used to evaluate the reliability within each domain because the sample size was too small to support reliability and factor analyses on the whole scale (Field, 2005). The resulting scale consisted of 53 items and five domains.

Cronbach's alpha, a measure of internal consistency (Cronbach, 1951), was calculated for each of the subscales in order to investigate how strongly items in each domain were related. Cronbach's alpha for each of the domains ranged from .93 to .96: Knowledge (.93), Communication (.94), Attitudes (.95), Professional Practice (.93), Negotiated Space (.96). The ranges of inter-item correlations across items were also provided: Knowledge (.35 to .83), Communication (.45 to .82), Attitude (.46 to .80), Professional Practice (.44 to .82), Negotiated Space (.45 to .90). Kendall tau rank correlations were computed across domains to evaluate independence between domains, and ranged from .39 to .63. Thus, it appears that there is some overlap in addition to independence across the domains. However, due to the small sample size preventing factor analyses, there is no evidence that the items actually load onto their expected factor, or that the scale actually results in five factors.

The authors of the AICS suggest that future research focus on item wording and domain stem revisions in order to develop a more parsimonious scale that is still content and construct valid. Furthermore, a larger sample should be collected in order to further evaluate the psychometric properties of the scale. As aforementioned, the sample size in

the first administration of the scale was too small to support factor analyses or reliability analyses for the complete scale.

Additionally, the scale was only administered to a small number of students from a single university in the U.S. The authors noted the importance of administering the scale to students across different universities as well as internationally if possible. In order to generalize findings across demographic groups and geographic areas, it is important to evaluate properties of the scale with a sample that is representative of the population for which the scale is intended. Ideally, this scale would be a useful measure of intercultural competency across programs of study throughout the United States, and thus a representative sample should be obtained.

Of the 80 students, 49 (60.5%) responded with "strongly agree" or "somewhat agree" to the statement "I understood all survey items without assistance" and 27 (32.1%) responded with "strongly disagree" or "somewhat disagree." No significant differences on subscale scores were found between those indicated that they understood the survey items without assistance and those who did not for all except the negotiated space subscale. For the negotiated space subscale, participants who indicated that they understood the survey items scored significantly lower than those who indicated that they did not understand the survey items. Thus, it is imperative that item wording is evaluated, especially for the negotiated space subscale, in future research examining the scale.

Current Study

Benson's Framework. The current study sought to develop a strong program of evidence for construct validity regarding the AICS. Benson (1998) provides three stages

of the validity process that provide a foundation for a strong collection of evidence for construct validity: substantive, structural, and external. This thesis evaluates the AICS with regard to each element of Benson's framework in order to provide a thorough assessment of the construct validity of the measure.

Substantive stage. Benson (1998) emphasized the importance for the measure to be grounded in research and theory. This necessitates a thorough review of existing literature and research in order to define the construct of interest. Items should be written to address all aspects identified as defining the construct. The process by which the developers of the AICS identified the aforementioned domains and associated competencies intrinsic to intercultural competency provide a strong theoretical and empirical basis for the measure. As described previously, the scale developers first identified the differences between local and global diversity. This helped them to identify the scope that the term intercultural competency was believed to cover (e.g., merging of local and global diversity).

Next, the team of researchers thoroughly reviewed literature across various disciplines that referred to the construct of intercultural competency to develop an operational definition reflected by both theory and empirical research. Once articulated, the domains and competencies were reviewed and revised by subject matter experts. Items were then written to reflect each competency. Once revised, the items were reviewed against the previously defined competencies to assure that there were items to reflect each dimension.

Of importance is the lack of components of both local and global diversity discussed in the previous review of existing measures. The aforementioned measures

were not appropriate for measuring the construct defined by Henderson et al. (2015), as they were far too limited in the scope of the content they covered. The AICS is an important contribution to the literature because it was developed with the goal of combining aspects from both multicultural education and internationalization, which has been posited as the foundation of intercultural competence.

Structural stage. The structural stage refers to the internal consistency among items of the measure. The goal of this stage is to evaluate how strongly the observed responses to items covary among themselves as well with the internal structure of the measure. The developers of the AICS were able to examine internal consistency only within each of the five proposed domains, with Cronbach's alpha ranging from .93 - .96. This indicated a high level of internal consistency. However, due to the large number of items, internal consistency should be further examined. They also were able to examine intercorrelations using Kendall tau rank correlations, which suggested both some overlap and independence across domains. However, due to their small sample size, they were unable to examine the factor structure of the AICS. Thus, the current study seeks to further examine the internal structure of the scale by obtaining a large enough sample size to conduct an exploratory factor analysis. Using this procedure will allow the researchers to further evaluate the evidence that supports (or disconfirms) the internal structure of the scale (i.e., Do items written to assess each specific domain map on to their respective domain?).

External stage. Benson (1998) described the external stage of development as the most crucial stage. This stage focuses on the evaluation of the ways in which the construct covaries with other constructs or among subjects in expected ways. This stage

of the framework was not explored by the developers of the AICS. Thus, the current study will examine both relationships among other variables and differences between groups of test takers. Specifics of these investigations will be discussed below (see Method section).

Research questions. The following research questions will be evaluated:

Research Question 1. What is the factor structure of the AICS?

Research Question 2. How do scores on the AICS relate to scores on external measures?

Research Question 3. Is there a difference between undergraduate and graduate student scores the AICS?

Research Question 4. Is there a difference between students who have taken integrated and standalone diversity courses on the AICS or external measures?

Method

Participants

Data were collected primarily from undergraduate students at a mid-sized Southeastern university. A variety of sampling methods were used. These included using the university's participant pool, administering the survey to a group of students on the university's annual spring assessment day, and sharing the survey link publicly via email and social media.

Participant Pool. The survey was administered through the university's participant pool. Students in general education psychology courses are required to participate in three studies posted to the participant pool per semester; however, they are

also permitted to complete an alternative assignment instead. A total of 312 responses were collected through the participant pool.

Assessment Day. Additionally, the survey was administered to senior psychology majors during the university-wide assessment day in February, 2016. All graduating psychology majors are required to participate in assessment day, and seniors were randomly assigned to take either a paper-and-pencil psychology content test or a series of online assessments. A total of 126 students were assigned to the online portion of the assessment, during which they completed the survey for this study first.

Mass Email/Social Media. Lastly, a mass email was sent to all undergraduate and graduate students enrolled at the university requesting their completion of the survey. The link to the survey was shared via social media (i.e., Facebook). A total of 502 students who accessed the survey through email or social media responded.

Description of Sample. Participants with grossly incomplete data, defined as less than a 50% completion rate, were removed from the data set. As a result of the screening process, a total of 210 responses were removed from the dataset, leaving a final sample of 730.

Of the sample, 559 students identified as female, 145 were male, 6 preferred not to respond, 3 as non-binary, 1 identified as gender nonconformist, and 1 identified as gender-queer. There were 586 respondents who identified as White or Caucasian, 39 as Black or African American, 39 as biracial or "other," 22 as Asian or Pacific Islander, 15 as Hispanic or Latino/a, 7 as Middle Eastern, 7 who preferred not to respond, and 1 as Native American or American Indian. In terms of sexual orientation, 644 respondents identified as heterosexual, 31 as bisexual, 19 as homosexual, 12 as "other," and 10 did

not respond. Regarding class year, 228 were first-year students, 127 sophomores, 41 juniors, 200 seniors, 87 graduate students, and 1 adult learner. The mean age of the sample was 21 years old. Most respondents were born in the United States (N = 682), with 32 respondents from a foreign country. English was the primary language for 702 of the respondents, and 13 indicated a foreign language as their primary language.

Materials

Alliant Intercultural Competency Scale, Revised (AICS-R). The revised version of the AICS consisted of 45 items designed to measure intercultural competencies. Participants respond to items on a 7-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree). The scale consisted of five domains: knowledge, communication, attitudes, professional practice, and negotiated space. Previous research on the original version of the scale resulted in the following Cronbach's alpha values associated with each of the expected subscales: Knowledge (.93), Communication (.94), Attitudes (.95), Professional Practice (.93), Negotiated Space (.96). Cronbach's alpha for the scale as a whole was not calculated. The AICS-R was administered in this study, and resulted in the following Cronbach's alpha values associated with each of the suggested subscales: Knowledge (.91), Communication (.88), Attitudes (.90), Professional Practice (.90), Negotiated Space (.94). The results of this study support calculating means for each of the suggested subscales, although further research should evaluate whether an overall mean for all items is appropriate as well.

External Measures. In addition to the AICS-R, a series of external measures were included in the survey. A measure of social desirability was included in order to determine the extent to which students were responding with a socially desirable response

set (i.e., responding in ways they believe are favorable or expected, rather than a true representation of their attitudes and beliefs). In addition, a measure of multicultural attitudes and experiences and a measure of racial color-blindness were included.

Social Desirability Scale (SDS). The short form of the Marlowe-Crowne Social Desirability Scale was included at the beginning of the survey (Reynolds, 1982). This scale is intended to assess whether participants are answering in ways that they think are favorable to others. The social desirability scale consists of 11 items scored in a true-false format with 1 for false and 2 for true (e.g., "There have been times when I was quite jealous of the good fortune of others"). There are five negatively worded items included in the scale. After reverse scoring the negatively worded items, a sum was computed for the 11 items. Reliability analyses performed by Reynolds indicated a Cronbach's alpha of .76. Cronbach's alpha for the present study was .338.

Multiculturalism Experiences Questionnaire (MEQ). The 15-item MEQ, developed by Narvaez and Hill (2010), was used to assess the level of exposure to and interest in multicultural experiences and activities. An example item from this scale reads, "I want to have friends from different cultural-racial-ethnic backgrounds." A sum was computed across all items. Narvaez and Hill reported a Cronbach's alpha of .75 for their study. Reliability analyses for the present study resulted in a Cronbach's alpha of .81.

Munroe Multicultural Attitudes Scale Questionnaire (MASQUE). The MASQUE was used to assess multicultural attitudes encompassing race, religion, sexual orientation, gender, and language (Munroe & Pearson, 2006). This measure assesses several dimensions important to intercultural competence. These include the extent to

which participants are aware of, care about, and act on issues related to each of the aforementioned aspects of diversity. An example item from the MASQUE reads, "I actively challenge gender inequalities." There are five negatively worded items included in the scale. After reverse scoring the negatively worded items, a sum was computed for the 18 items. Initial reliability testing resulted in an alpha of .80 as reported by Munroe and Pearson. Cronbach's alpha for the present study was .86.

Color-Blind Racial Attitudes Scale (CoBRAS). The CoBRAS consists of 20 items constructed to assess the degree of to which respondents consider themselves "blind" to color and privilege (Neville et al., 2000; $\alpha = .86$). The scale consists of three factors: racial privilege, institutional discrimination, and blatant racial issues. An example item from this scale reads, "White people in the U.S. have certain advantages because of the color of their skin." Sums can be calculated individual subscales or the total measure as a whole. There are six negatively worded items. After reverse coding the negatively worded items, a sum was calculated for the total measure. One item, "It is important for public schools to teach about the history and contributions of racial and ethnic minorities," was unintentionally omitted when uploading the CoBRAS items to the online survey. Therefore, the total score in the present study reflects only 19 of the original items. Neville et al. reported a Cronbach's alpha from their study of .86. Cronbach's alpha for the present study was .89.

Demographics. A brief demographics section was included at the end of the survey to collect information about the participants. Participants were asked to indicate the university they are attending, age, year in school, race, ethnicity, gender identity, and sexual orientation, country of origin, primary language, and major or program of study.

Procedure

Scale development. Prior to administering the scale, a team of researchers revised the items. A team of undergraduate research assistants responded to the scale while documenting their thought process as they read and answered each question. Feedback from the team was used to identify items necessitating revisions. Eight items were eliminated due to ambiguous wording. Additionally, the majority of the items were reworded to enhance the clarity and readability of the items. The resulting scale consisted of 45 items. Feedback was also used to revise the item stems. Qualifiers within the stems were eliminated in order to simplify interpretations of the wording. See the Appendix for the revised scale.

Administration. All data were collected via an online survey tool (Qualtrics). Data were downloaded from Qualtrics into IBM SPSS Version 23 software, which was subsequently used for all analyses. The revised version of the AICS (AICS-R) and a series of external measures were included in the survey. A measure of social desirability was presented first to respondents, then the AICS-R, followed by a series of three other measures presented in random order. The survey concluded with a brief demographics section. The total survey generally took between 5 and 20 minutes to complete.

Data Analysis

Data were analyzed using IBM SPSS Statistics Version 23. All analyses were performed using listwise deletion. First, an exploratory factor analysis (EFA) was performed to examine the factor structure of the AICS-R. Principal axis factoring (PAF) with promax rotation was used for all EFA analyses. Next, scores on the AICS-R, MASQUE, MEQ, and CoBRAS were correlated. The relationship between the AICS-R

and the other measures was evaluated as an aspect of validity evidence for the AICS-R. Last, ANOVAs were performed in order to examine group differences based on class year, type of diversity courses taken, and area of study. Group differences were also evaluated in terms of validity evidence for the AICS-R. Results and implications for the use of the AICS-R are discussed.

Results

Data Screening

Data obtained from the AICS-R were screened for univariate normality. For skewness and kurtosis, statistics less than the absolute value of two were considered normal (Bandalos & Finney, 2010). An examination of skewness values for all individual items indicated that all items were within acceptable ranges. An examination of kurtosis values indicated that eight items exhibited leptokurtic distributions with values greater than 2 (See Table 1). Item response distributions for the AICS-R were next examined for floor and ceiling effects. After examining the histograms of responses for each individual item, all items appeared to exhibit a ceiling effect to some extent. Almost all item means were above 5, with the exception of three items with means ranging between 4.49 and 4.99. Thus, on average, respondents appear to "slightly agree" to "agree" with the majority of the items. This means that for most items, students appear to agree that they have the skill or ability described.

Social Desirability

Affective items are susceptible to response styles such as socially desirable responding (Bandalos, in prep). Social desirability scores were correlated with scores on the external measures in order to evaluate the extent to which scores from other measures

may be influenced by socially desirable responding. Although social desirability scores were significantly correlated with scores from all other measures, this is likely due to the high sample size. The magnitudes of these correlations ranged from negligible to small (*r*s range from .017 to .163). This indicates that social desirability scores are not meaningfully related to scores on the other measures included in the survey, meaning that the items in each of the measures are likely not systematically capturing social desirable responding.

Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) was conducted to investigate the factor structure of scores from the AICS-R. Means, standard deviations, skewness values, and kurtosis values for each of the items can be found in Table 1. The inter-item correlation matrix is presented in Table 2. Observation of the inter-item correlation matrix revealed that all items were significantly positively correlated (*p*s < .001) and *r* values ranged from .181 to .581, suggesting that factor analysis was appropriate for all items. The Kaiser-Meyer-Olkin measure of sampling adequacy resulted in a value of .964, which was higher than the recommended value of .6 (Field, 2005). Additionally, Bartlett's test of sphericity was significant, $\chi^2(990) = 20601.251$, $p < .001$, suggesting that the correlation matrix is not an identity matrix and there are relationships among the variables. All item communalities were above .4 (see Table 3), suggesting that each item shared an adequate amount of common variance with the other items included in the scale. Thus, factor analysis was performed on scores from all 45 items on the AICS-R.

All factor analyses were conducted using principal axis factoring (PAF) with promax rotation (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Because

intercultural competence is a latent construct, PAF is preferable to PCA (Bandalos, in prep). To determine the degree to which items load onto each extracted factor, oblique rotation was used so that factors could correlate with each other. To determine the number of factors to retain, three criteria were examined. Eigenvalues were evaluated using the Kaiser (K1) criterion (Kaiser, 1960) in which factors with eigenvalues greater than one are retained; a scree plot was examined (Cattell, 1966); and Horn's parallel analysis was performed (1965). Initial examination of eigenvalues suggested a six-factor solution, with each factor explaining 41.82%, 7.286%, 4.958%, 3.512%, 3.355% of the common variance in AICS-R scores, respectively. Overall, the six factors explained 63.091% of the variance in AICS-R scores prior to extraction.

A visual inspection of the scree plot (see Figure 1) indicated a substantial drop in explained variance after the first factor and suggested support for retention of either five or six factors. Parallel analysis was performed in which 1000 randomly generated matrices that consisted of the same number of people and variables as the sample were simulated. The distribution of eigenvalues from these matrices were compared to the initial eigenvalues from the sample. The comparison of the plot of initial eigenvalues based on the sample and eigenvalues based on the randomly generated data suggested the retention arguably five or six factors (See Figure 2). The first five factors were clearly above of the cutoff line estimated using the simulated data, and the sixth factor was found to just overlap the cutoff line.

Although the initial factor analysis suggested a six-factor solution, three-, four-, and five-factor solutions were also examined for comparison. In order to compare the various factor solutions, both common variance explained by the model and the pattern

coefficient matrices were examined. A minimum pattern coefficient value of .4 was used to determine if items loaded meaningfully onto a factor (Stevens, 2008). In all of the models, two items (item 19; "I am able to confront stereotypes, prejudice, or racism in social situations;" and, item 25; "I am able to understand how my values, beliefs, and communication style may come across to people from other cultures/countries") consistently did not meet the criteria of a .4 or greater factor loading. In the three-factor solution, eight items did not load strongly (i.e., pattern coefficient value > .4) onto any factors and the common variance explained reduced to 54.069%. In the four-factor solution, five items did not load strongly onto any factors and the common variance explained was 57.581%. The five-factor solution resulted in only three items that failed to load strongly onto any factors, and two items that cross-loaded.

Due to the incremental change in the explained variance between the five and six factor models, and the relatively smaller number of items that did not load strongly onto any of the factors compared to the three and four factor models, the five and six factor solutions were deemed the most appropriate. To determine which factor solution to champion, several considerations were evaluated. The six-factor model explained just slightly more variance (about 3%) than the five-factor model. Both factor solutions resulted in only two cross-loading items. Furthermore, in both models, items loaded onto factors largely in the expected ways based on theoretical conceptualization of the subscales. However, when examining the item loadings for the sixth factor in the six-factor model, only two items loaded with one cross-loading onto another factor. Due to this and the theoretical support for the five-factor model, the five-factor model was championed.

For the five-factor model, each factor separately accounted for 41.82%, 7.29%, 4.96%, 3.51%, and 3.16% of the variance respectively. All items in this model had primary factor loadings above .4 (see Table 4 for pattern coefficients table 5 for structure coefficients). Both of the aforementioned cross-loading items were theoretically part of the Communication subscale, but they primarily loaded onto the theorized negotiated space subscale. The first item that cross-loaded stated, "I am skilled at discussing diversity issues related to nationality." The factor loading for the communication subscale was .443, and for the negotiated space subscale it was .511. The second cross-loading item read, "I am skilled at discussing diversity issues related to language." The factor loading for the communication subscale was .445, and for the negotiated space subscale it was .445. Due to the small discrepancies in factor loadings and the theoretical basis of the items, it was decided to maintain these items as part of the communication subscale. Three items did not load strongly onto any of the extracted factors. For theoretical reasons, these items were deemed important to the construct as a whole and it was decided to retain them in the measure until subsequent research can further evaluate their psychometric properties and relationships to the measure in another sample.

Correlations among the factors ranged from .505 to .670 (see Table 6 for factor correlation values). This suggests that the five subscales were related but also appeared to be measuring relatively distinct domains of the construct. Due to the moderate factor correlation values, computing a total score based on all items may not be warranted. Rather, individual subscale scores may be most appropriate for the AICS-R; however, until more information is collected from a CFA both subscale and total score are still plausible. Therefore, both subscale and total scores will be reported throughout this

study. The factor labels suggested by Henderson et al. (2015) corresponded with the five factors and were retained. Cronbach's alpha was calculated for each of the subscales. The alphas were high, suggesting internal consistency among items for each of the individual subscales: .91 for the 10 item knowledge subscale, .88 for the eight item communication subscale, .90 for the seven item attitudes subscale, .90 for the seven item professional practice subscale, and .94 for the 13 item negotiated space subscale. Cronbach's alpha was also calculated for all of the 45 items on the AICS-R and resulted in a value of .96, suggesting high internal consistency among the items.

Bivariate Correlation Analyses

Zero-order bivariate Pearson Product-Moment correlation coefficients were used to evaluate the strength of the relationship between scores on the AICS-R and scores on the other external measures included in the survey (i.e., MASQUE, MEQ, CoBRAS). Correlations between the AICS-R and the MEQ, MASQUE, and COBRAS were of primary interest. Correlation coefficients were evaluated for statistical significance and using Cohen's suggested benchmarks for *r* values (.1 to .3 considered small, .3 to .5 considered medium, and .5 considered large; Cohen, 1988). Considering the large sample size, *r-values* were expected to be statistically significant but moderate in magnitude. Some overlap was expected between the AICS-R and external measures. However, the AICS-R contributes to the existing instruments by measuring a more inclusive and comprehensive conceptualization of diversity and sociocultural awareness (i.e., intercultural competency). Thus, the magnitude of the correlations was expected to be within the moderate range.

Correlations were conducted following listwise deletion for missing data, resulting in a total *N* of 619. See Table 6 for correlation coefficients and Table 7 for associated R^2 effect sizes. As expected, all correlations among both the total and subscale scores and scores on the external measures were statistically significant (*p*s < .001). Total scores on the AICS-R correlated positively with scores on the MASQUE and scores on the MEQ, and these relationships were considered large in magnitude (*r*s = .639 and .658 respectively). Scores on the AICS-R and the CoBRAS were moderately negatively correlated (*r* = -.384). Scores on the communication subscale were moderately related to scores on the MASQUE and MEQ (*r*s = .467 and .479, respectively) in the positive direction and moderately related to scores on the CoBRAS in the negative direction (*r* = -.339). In terms of the attitudes subscale, scores were strongly related to the MASQUE and MEQ in the positive direction (*r*s = .626 and .624 respectively), and moderately related to scores on the CoBRAS in the negative direction (*r* = -.387). In terms of the professional practice subscale, scores were strongly related to the MASQUE and the MEQ in the positive direction (*r*s = 563 and .614 respectively) and moderately related to the CoBRAS in the negative direction (*r* = -.341). Scores on the negotiated space subscale were positively strongly related to scores on the MASQUE and MEQ (*r*s = .538 and .581 respectively), and weakly related to scores on the CoBRAS in the negative direction (*r* = -.240).

The relationships among the three external measures were also examined for exploratory purposes. The three correlations were statistically significant (*p*s < .001). Scores on the MASQUE were strongly correlated with scores on the MEQ in the positive direction (*r* = .546) and strongly correlated with scores on the CoBRAS in the negative

direction (r = -.582). Additionally, scores on the MEQ and CoBRAS were moderately correlated in the negative direction (r = -.362).

Group differences on AICS-R

Class Year. Students across class years (i.e., first-year, sophomore, junior, senior, graduate student) were compared using a one-way between-groups analysis of variance (ANOVA) on each of the five subscales and the total score of the AICS-R. Furthermore, in its initial administration, the AICS was administered solely to graduate students, so differences in responses across academic class years had not previously been examined.

Results for each of ANOVAs are presented in Table 8, which includes means, standard deviations, and associated test values. Results indicated that there was a significant difference across class year on the AICS-R total score, $F(4, 643) = 14.217$, $p < .001$. The size of this effect was between medium and large, with the differences between class years accounting for 8.1% of the variance in total AICS-R scores ($\eta^2 = .081$). Bonferroni post-hoc analyses revealed that senior students reported significantly higher scores on the AICS-R compared to all other class years (all ps < .01). Additionally, sophomore students scored significantly higher compared to first-year students ($p = .034$).

There was also a significant difference across class year for each of the five subscales of the AICS-R. The size of each of these effects fell between small and medium (η^2 ranged from .043 to .071). The largest effect among the subscales was for professional practice, whereas the smallest effect was for negotiated space. In order to examine which class years differed from each other for each of the subscales, Bonferroni

post-hoc analyses were performed. In the case of violations to the homogeneity of variance assumption (analyses that violated this assumption are noted in Table 8), Games-Howell post-hoc analyses are reported instead. In terms of the knowledge subscale, seniors scored significantly higher compared to first-year students and sophomores (all *p*s < .010). In terms of the communication subscale, seniors scored significantly higher than first-year students ($p < .001$). In terms of the attitudes subscale, seniors scored significantly higher than first-year ($p < .001$) and sophomore students ($p = .016$). In terms of the professional practice subscale, seniors scored significantly higher than first-years, sophomores, and graduate students (all *p*s < .047). Additionally, juniors and sophomores scored significantly higher than first-years (*p*s < .024), but not significantly different from each other ($p > .999$). In terms of the negotiated space subscale, seniors scored significantly higher than all other class years (all *p*s < .028).

Major/Program of Study. Due to data collection via assessment day, psychology majors made up over half of the sample of seniors. Because of this, it is possible that the aforementioned differences across class year reflect differences between students who are psychology majors and students who are enrolled in another plan of study. In order to investigate this, students' majors were coded into the following categories to represent different programs of study: (1) underclass psychology majors (freshmen – juniors; $N = 198$), (2) senior psychology majors ($N = 139$), (3) graduate students affiliated with Graduate Psychology ($N = 23$), (4) graduate students affiliated with a program other than psychology ($N = 24$), (5) other social science majors ($N = 22$), (6) science and math ($N = 76$), (7) education ($N = 22$), (8) business ($N = 65$), (9) health science and nursing ($N = 138$), (10) public and international affairs ($N = 31$), (11)

communication (N = 46), (12) arts and humanities (N = 36) and (13) undeclared (N = 25). Those who indicated that they had not yet declared a major were removed from the analysis.

These 13 groups were compared using a one-way ANOVA, and the primary question of interest was whether psychology seniors scored significantly differently than psychology underclassmen or students involved in other programs of study. There were significant differences across program of study for the AICS-R total score and each of the subscales; however, because the group sizes were go discrepant results may not be stable across estimates so it is important to evaluate effect sizes. The size of the effects ranged from small to medium, with η^2 values ranging from .043 to .081. Due to the applied nature of the research question, there may be some practical significance to the differences across class year. See Table 9 for means, standard deviations, and associated test values.

Assumptions of homogeneity of variance were violated for all of the subscales and the AICS-R total, so Games-Howell post-hoc comparisons were conducted. Across all AICS-R scores, psychology seniors scored significantly higher than students in science/math, business, health sciences/nursing, and communication (ps < .02). For the AICS-R Total and AICS-R professional practice subscale, senior psychology majors also scored significantly higher than psychology underclassmen and education students (ps < .011).

Students in public and international affairs also consistently scored significantly higher compared to other programs of study. For the AICS-R total and the knowledge, attitudes, and professional practice subscales, public/international affairs students scored

significantly higher than health sciences/nursing students (*p*s < .008). Additionally, public/international affairs students scored higher than science/math and business students on the AICS-R total and attitudes subscale (*p*s < .002). On the attitudes and professional practice subscales, public/international affairs students also scored higher than communication students (*p*s < .032). Lastly, graduate students in psychology scored significantly higher than undergraduates in health sciences/nursing on the AICS-R attitudes subscale ($p = .044$). Thus, most of the differences between groups were due to psychology seniors and public/international students scoring higher than peers from several other programs of study on the AICS-R total and subscales. Those in science/math, business, health sciences/nursing, and communication consistently scored lower than senior psychology majors and public/international affairs students.

Sociocultural Awareness Courses. An independent samples t-test was used to compare senior psychology students who took only integrated diversity courses (courses where sociocultural content is integrated with content from a subfield of psychology, e.g., Psychology of the Workplace with an emphasis on sociocultural perspectives) through the Psychology Department and those who took a standalone diversity through the Psychology Department course. Because curricular information was obtainable only for those students who took the survey during the university's Psychology Department Assessment Day, this analysis included only senior psychology majors at the university. The number of students who took a combination of both types of courses was too small to support statistical analysis as an independent group ($N = 8$). Additionally, students who only took study abroad courses ($N = 11$), only took practicum ($N = 1$), or took any combination of all of the sociocultural course options ($N = 28$, though *N*s for each

individual combination ranged from 1 - 8) were too small individually to support statistical analysis. Furthermore, there was no logical way to combine students who took various combinations of the courses into a single group. Thus, only those who took only one integrated or only one standalone diversity courses were included in the analysis. No significant differences were found between the two groups on any of the five subscales or the average of all the items. Additionally, the size of the effects for each difference ranged from negligible to small, with Cohen's d values ranging from .04 to .36. Results for each of t-tests are presented in Table 10, including means, standard deviations, and associated test values.

Group differences on External Measures

Class Year. ANOVAs were additionally conducted examining differences across class year for each of the external measures included in the survey (i.e., MASQUE, MEQ, CoBRAS). These ANOVAs were performed on the entire sample. See Table 8 for means, standard deviations, and associated test values. A significant difference was found across class year for scores on the MEQ. Bonferroni post-hoc analyses revealed that first-years scored significantly lower compared to juniors, seniors, and graduate students (all ps < .05). Furthermore, sophomores scored significantly lower than seniors and graduate students (ps < .01). Additionally, there were significant differences found across class year for scores on the MASQUE. Bonferroni post-hoc analyses revealed that first-years and sophomores both scored significantly lower on the MASQUE compared to seniors and graduate students (all ps < .009), but neither first-years compared to sophomores nor seniors compared to graduate students scored significantly different (ps > .999). Lastly, there were significant differences found across class year for scores on

the CoBRAS. Games-Howell post-hoc analyses revealed that both seniors and graduate students scored significantly lower on the CoBRAS compared to first-years, sophomores, and juniors (all *p*s < .005), but seniors and graduate students did not score significantly differently from each other ($p = .893$).

Major/Program of Study. Similar to the AICS-R, differences were consistently found between senior students and some of their underclass peers. As previously noted, psychology majors were overrepresented in the senior sample. This prompted the coding of students into different programs of study in order to compare to senior psychology majors and underclass psychology majors. One-way ANOVAs resulted in significant differences for each of the external measures. Because of the large discrepancies in group sizes, it is also important to consider effect sizes. Effect sizes evidenced by η^2 values ranged from small to medium, which may indicate some practical significance considering the applied nature of the research question. See Table 9 for means, standard deviations, and associated test values.

Bonferroni post-hoc analyses were conducted in order to determine which groups were significantly different from each other. On the MEQ, psychology seniors and public/international affairs students both scored significantly higher than science/math and health sciences/nursing students (*p*s < .005) did. Public/International affairs students also scored higher than psychology underclassmen and education, business, and communication students (*p*s < 008). Additionally, art/humanities students scored significantly higher than health sciences/nursing students ($p = .008$). On the MASQUE, psychology seniors, psychology graduate students, and public/international affairs students all scored significantly higher than business and health sciences/nursing students

(*p*s < .03). Additionally, senior psychology students scored higher than science/math students ($p = .023$). On the CoBRAS, both senior psychology and graduate psychology students scored significantly lower than science/math, business, and health science/nursing students (*p*s < .004). Graduate psychology students also scored lower than education, public/international affairs and communications students (*p*s < .049).

Sociocultural Awareness Courses. Additionally, a series of t-tests were conducted examining differences based sociocultural awareness courses completed (for senior psychology students) for each of the external measures included in the survey (i.e., MASQUE, MEQ, CoBRAS). There were no significant differences between those who completed only integrated courses and those who completed only standalone courses on the MEQ or the MASQUE. See Table 10 for means, standard deviations, and test values. Associated effect sizes were small for the MEQ (Cohen's $d = .30$) and negligible for the MASQUE (Cohen's $d = .06$). However, there was a significant difference between the two groups on the CoBRAS. Students who took only standalone diversity courses reported significantly lower scores on the CoBRAS compared to those who completed only integrated courses. The size of this effect was medium, as evidenced by a Cohen's *d* value of .73. Students who completed only integrated courses reported CoBRAS scores that were on average .73 standard deviations higher than students who completed only standalone courses. Thus, it appears that students who completed a standalone course reported less racially color-blind attitudes compared to those who took integrated courses.

Discussion

The purpose of this study was to evaluate the psychometric properties of and validity evidence supporting the use of the revised version of the Alliant Intercultural

Competency Scale using Benson's (1998) validity framework. This framework involves developing a measure that is firmly rooted in theory and research, is internally reliable, and is related measures of other constructs in theoretically expected ways. This study specifically focused on evaluating the internal structure of the AICS-R as well as how scores from the AICS-R relate to theoretically similar constructs. Evidence regarding the utility of the AICS-R is evaluated through Benson's framework below. Limitations and future directions are also discussed.

Substantive Stage

The original version of the AICS was developed by Henderson et al. (2015) after a thorough and rigorous review of literature regarding intercultural competencies. Henderson et al. outlined five overarching domains that they conceptualized as critical to intercultural competence. These domains included knowledge, attitudes, communication, professional practice, and negotiated space. Next, Henderson et al. wrote items to represent each of these domains. In the current study, a team of researchers reviewed the items. Although they felt the items addressed the content of the domains outlined by Henderson et al., results from a think-aloud type of procedure indicated that the item wording was too ambiguous and complex for the majority of the items. This feedback was used to revise all items in order to increase their readability and interpretability. The revised version of the AICS, referred to in this paper as the AICS-R, was then used in the next steps of Benson's program of establishing strong construct validity.

Structural Stage

The next step using Benson's (1998) framework was to evaluate the structural validity of the instrument. An EFA was performed on scores from all 45 items included

in the AICS-R to examine if the different domains conceptualized by Henderson et al. (2015) would manifest as separate factors of intercultural competence. Results from the EFA supported a five-factor model for the AICS-R. The factors strongly aligned with the theoretical basis for the measure conceptualized by Henderson et al. (2015). The five conceptualized domains were represented by the five extracted factors and associated item loadings, providing both statistical and conceptual support for the internal structure of the AICS-R. The five resulting subscales included Knowledge, Communication, Attitudes, Professional Practice, and Negotiated Space.

After evaluating results from the EFA, it was decided to retain all 45 of the original items. However, it is important to note that two items cross-loaded onto two factors and three items failed to load strongly onto any factor (item loadings < .4). These items were believed to be important to the construct and did not appear to overlap with content represented by any of the other items in great depth. Thus, they were retained in the measure, as removing the items could potentially eliminate important aspects of the construct. Additionally, when examining the item statistics from reliability analyses, coefficient alpha was not improved by removing any of these five items from either the total scale or their respective subscales. Coefficient alpha was above the acceptable cutoff for the total scale as well as the individual subscales (all alphas > .8; Kline, 1999). Furthermore, EFA is a procedure that may be influenced by idiosyncratic relationships present in the data that are not due differences among individuals in the actual construct. Although researchers commonly remove items that cross-load or have low coefficients, Bandalos (in prep) recommends waiting until results are replicated in an independent sample before deciding to remove items.

Bandalos's suggestion was heeded in the decision to retain all items, particularly because there were certain limitations regarding the EFA that could lead to idiosyncrasies in the data. In the survey, items were grouped by subscale. Similar wording and phrasing may influence items to factor together (Bandalos, in prep; Cattell & Tsujioka, 1964). This effect may be exaggerated by the fact that all items were grouped by subscale, and thus, common stem. So, it is possible that item loadings were an artifact of the item wording and order. Additionally, the demographics of the sample may not represent the population to which results are intended to be generalized. The sample was predominantly White and most respondents identified as female. Additionally, psychology students may be overrepresented in the sample, as over 100 of the responses were collected from senior psychology students during annual assessments.

External Stage

Correlations with External Measures. The next step in Benson's (1998) framework is to examine the extent to which the instrument of interest relates in expected ways to other instruments that are intended to measure similar constructs. Scores from each of the AICS-R subscales and the total score were correlated with scores from the MASQUE, MEQ, and CoBRAS. As expected, both total and subscale scores on the AICS-R were positively correlated with scores on both the MASQUE and MEQ. Most of these correlations were strong in magnitude, with the exception of the knowledge and communication subscales. Scores from the knowledge and communication subscales were moderately correlated with scores from the MEQ and MASQUE. Additionally, the MEQ and MASQUE were highly correlated, which would be expected considering the content of these scales (e.g., both focused on multicultural content).

Although a positive relationship between scores on the AICS-R (both total score and subscale scores) was expected, the magnitude of the relationship suggests that there may be more overlap between the MEQ, MASQUE, and the AICS-R than anticipated. Henderson et al. (2015) sought to develop a measure of intercultural competence that integrated key competencies from multicultural and international studies. These two realms overlap in content, values, and goals, yet have traditionally remained relatively divided and function separately within higher education (Olson et al., 2007). Because of this divide, most existing measures of sociocultural and intercultural competence (as well as other similar constructs) are developed primarily from either a multicultural or international perspective (Henderson et al., 2015). Henderson et al. attempted to bridge the gap between these two perspectives in their definition of intercultural competence, which includes being competent in issues related to both local and global diversity. The strong correlation between the AICS-R scores (both total and the three subscales) and both the MEQ and MASQUE suggest that the AICS-R may be more closely aligned with multicultural content than originally believed. Perhaps intercultural competence, as defined by the AICS-R, is not as distinct from multicultural content as initially proposed. Another plausible explanation for the strong correlation is that the AICS-R includes more items that are related to local diversity than items that are related to global diversity. However, without a measure focused on global diversity, this posited explanation is difficult to evaluate.

The correlation between the communication and knowledge subscales and the MEQ and MASQUE were moderate in nature. Compared to the strong correlations between scores from the other three subscales and the MEQ and MASQUE, this suggests

that the communication and knowledge subscales are related to the MEQ and MASQUE, but are also more distinct than the other subscales and the measure as a whole. These two subscales may include relatively more international content or less multicultural content than the other subscales. However, as previously noted, without a measure of global diversity this is difficult to evaluate.

Additionally, scores from the CoBRAS were negatively correlated with all other measures. This was expected, considering higher scores on the CoBRAS represent stronger racially color-blind attitudes. Thus, those who perceived themselves to be more interculturally competent, and those who reported greater awareness/ knowledge and more experiences of multicultural perspectives, reported less racially color-blind attitudes. The strength of the relationship between the CoBRAS and the total and subscale scores on the AICS-R was moderate. The only exception to this was the negotiated space subscale; scores from this subscale were weakly correlated with scores from the CoBRAS. The relatively weaker relationship between scores on the AICS-R and CoBRAS compared to scores on the AICS-R and the MEQ and MASQUE suggests that the CoBRAS and AICS-R are measuring relatively more distinct constructs than those measured by the MEQ and MASQUE. This is unsurprising, as the CoBRAS measure a very narrowly defined construct within the realm of local diversity. Those who are interculturally competent should be less likely to endorse racially color-blind attitudes; however, the AICS-R attempts to measure a much more widely defined construct of intercultural competence. Thus, both the direction and strength of the relationships between scores on the CoBRAS and AICS-R were aligned with what was expected.

Overall, the pattern of relationships between the AICS-R and external variables provide preliminary support for the theoretical basis of the AICS-R. The AICS-R is more strongly related to multicultural knowledge and attitudes than it is to racial color-blindness. The relationship between scores from the AICS-R and scores from the multicultural measures suggest the AICS-R may be more closely related to multicultural content than intended. Perhaps items on this scale are more representative of multicultural content than they are of international content, or perhaps there is more overlap between the multicultural and international perspectives than previously suggested.

Group differences across Class Year. As another aspect of the external stage of Benson's framework (1998), differences in scores on the AICS-R total scale and subscales were compared across class year. It was believed that graduate students would have been exposed to more intercultural content and experiences, and thus, they would rate themselves as more interculturally competent compared to undergraduate students. This hypothesis was not supported, as only senior students scored significantly higher than any of the other class years on the total scale and most subscales of the AICS-R. The only exceptions to this were sophomores scoring higher than first-year students on the total score, and juniors and sophomores scoring higher than first-year students on the professional practice subscale. In addition to statistical significance, the size of these effects all fell within the medium to large range, suggesting the differences may also be practically significant.

It is possible that the reasons for the hypothesized difference between undergraduate and graduate students applied instead to the difference between

underclassmen and senior students. Senior students would have been provided more opportunities to complete coursework related to intercultural competence through completion of their general education coursework and possibly their major curriculum. Additionally, senior students would have had more opportunities to participate in professional experiences (e.g., internships, jobs) or may have been exposed to more group work experiences in which they interacted with individuals from more diverse demographic backgrounds.

Interestingly, graduate students did not score higher on the AICS-R compared to first-year students or any other class year. This finding suggests that graduate students may not feel more efficacious when it comes to cultural competence than undergraduate students. Another potential explanation for this finding is the sample of senior students used for comparison in this study. Data were collected from half of the senior psychology majors at the university on the annual assessment day, psychology majors makeup over 50% of the seniors in the sample. Thus, the differences between seniors and the other class years may not be generalizable to other seniors at the university (or outside of the university). If senior psychology majors consider themselves more competent than their senior peers from other majors and fields of study, then this difference may instead be due to the overrepresentation of psychology majors in the senior sample.

Group differences across Majors/Programs of Study. To investigate whether the overrepresentation of senior psychology majors may be influencing the difference found across class year, students' scores on the AICS-R were compared across program of study. Students' majors and graduate programs were coded into the following groups

for comparison: psychology majors who were underclassmen (i.e., first-years, sophomores, and juniors), senior psychology majors, psychology graduate students, other graduate students, social sciences, science/math, education, business, health sciences/nursing, public/international affairs, communications, and arts/humanities.

Results from these analyses revealed that senior psychology majors scored higher on the AICS-R scores than many of their peers. Because these differences were not seen for psychology underclassmen the differences across class year may actually reflect differences between psychology seniors and students from different majors across class years. Interestingly, public/international affairs students also scored higher on the AICS-R compared to many of their peers. Public/International affairs students were likely exposed to intercultural information throughout their academic careers as well, but from an international or global perspective. Therefore, it would be expected that students in these programs of study would feel more interculturally competent compared to their peers who are less likely to have encountered this content in their academic careers. These findings provide support for the validity of the AICS-R, as students who you would expect to feel more interculturally competent based on their choice of major did indeed indicate this in their responses.

There were also differences on the external measures across programs of study. Similar to the AICS-R, senior psychology and public/international affairs students both scored higher than several of their peer groups on the MEQ and MASQUE. This suggests that seniors in psychology and students in public/international affairs may have more interest in and knowledge of multicultural perspectives and issues. Interestingly, arts/humanities students reported higher scores on the MEQ compared to health

sciences/nursing students. It seems that art and humanities students have had and are interested in having more multicultural experiences compared to health sciences and nursing students. Regarding the CoBRAS, overall, graduate psychology students reported less racially colorblind attitudes than undergraduate students in several programs of study. Additionally, psychology seniors reported less color-blind attitudes than several other undergraduate groups. Considering the content covered by the CoBRAS reflects a construct much more likely to be addressed courses developed from a multicultural perspective, these differences align with what would be expected based on the literature.

Consistently across measures, the senior psychology and public/international affairs students exhibited greater self-perceived intercultural competence, multicultural experiences, and multicultural knowledge and awareness. The differences across class year could be an artifact of this difference between senior psychology majors and their peers. Additionally, this suggests that students in psychology and public/international affairs have greater knowledge and awareness regarding intercultural and multicultural issues and perspectives.

Group Differences Based on Courses Taken. There were no significant differences found across any of the AICS-R scores (total or subscales), the MEQ, or the MASQUE depending on whether standalone or integrated diversity courses had been completed. However, students who took standalone courses reported significantly lower racially color-blind attitudes than students who took only integrated courses. This effect was medium in magnitude, indicating that this difference may also be practically important. Previous research suggested that an infusion-structured curriculum would better address student-learning goals related to diversity and cultural content than a

curriculum that utilizes a standalone curriculum approach (Marchesani & Adams, 1992; AACU, 2005, APA, 2013). Although this analysis does not compare across curriculum structures, it compares across courses that may represent these two approaches on an individual course level. However, courses vary widely across universities, departments, and even within departments and programs. Thus, these findings may not generalize to other courses and universities. Moreover, these findings may not generalize from the course level to curriculum level. With this in mind, these results may tentatively suggest standalone courses may more effectively address some aspects of sociocultural content than the integrated courses.

It is important to note that this difference was only found for the CoBRAS, which as previously noted, includes content pertaining to the very narrowly defined construct of racial color-blindness. Although previous literature suggests that standalone diversity courses may marginalize the content resulting in resistance from students in embracing the content, the present findings suggest that standalone courses may actually be effective for addressing specific, important sociocultural issues. Courses that focus specifically on such constructs may actually be better addressed in a course focused on diversity that allows for more in-depth coverage of the content.

The AICS-R, MASQUE, and MEQ cover broader constructs than the CoBRAS. Thus, when it comes to multicultural awareness and experiences there appears to be no difference based on the type of sociocultural course taken. Moreover, self-perceived intercultural competence as measured by the AICS-R does not appear to differ based on the sociocultural content course taken. When it comes to broader conceptualizations of

multicultural and intercultural competence, both standalone and integrated courses may be comparably effective.

It is important to consider the sample used in these analyses. Not only is the range of courses compared across students limited, but the range of instructors is even further limited. It is possible that these differences are influenced by other systematic factors as well, such as the instructor of the course. Moreover, all students in the sample were psychology majors. Therefore, these differences may not be generalizable beyond the major or department from which the sample was obtained. Additionally, as there was no pre-test data collected for baseline comparison, it is not possible to attribute the differences in racial color-blindness to the type of course taken. It is possible that students who care more about sociocultural issues were already less racially color-blind than their peers who are less aware of or concerned with sociocultural issues. These students may have been more likely to enroll in courses that specifically address sociocultural content and diversity issues, whereas their peers may have preferred to enroll in a class that is focused on some other topic and infuses diversity content.

Overall, the evidence gathered to evaluate the external stage provides preliminary support for the validity of the use of the AICS-R to assess intercultural competence. Correlational evidence suggests that as awareness of and experience with multiculturalism increases, so does self-perceived intercultural competence. Although the AICS-R may cover relatively more multicultural content, scores on the AICS-R are more strongly related to the broader constructs of multicultural knowledge and awareness (measured by the MASQUE) and multicultural experiences (measured by the MEQ) than they are to the more narrowly defined construct of racial color-blindness.

Furthermore, although there were few differences between undergraduate and graduate students, it appears that undergraduates who are in higher class years rated themselves higher on the AICS-R compared to those who were in lower class years. This difference may be an artifact of the overrepresentation of senior psychology majors in the sample. If so, this would also provide support for the validity of the AICS-R, as psychology majors at the university sampled are required to complete a three credit hours of an approved sociocultural awareness course. These classes are likely to align more with the multicultural perspective than the international perspective, considering much of the literature regarding multicultural education has resulted from the social sciences, including psychology (Olson et al., 2007). These results suggest that students who are likely to have been exposed to more multicultural and intercultural content did rate themselves as more interculturally competent, as measured through responses to the AICS-R.

Implications for Higher Education and Assessment

This study provided preliminary support for the utility of the AICS-R to measure self-perceived intercultural competencies in higher education students. Results from this study also have implications for intercultural assessment. The importance of diversity in higher education has become increasingly emphasized in recent years. A goal related to this increased emphasis is to provide an education and experiences that develop competencies related to intercultural issues, perspectives, and communication in order to prepare students to enter an increasingly diverse and globalized workplace (Zalaquett et al., 2008). There are two main perspectives of diversity in higher education that remain largely distinct (Olson et al., 2007). These perspectives approach diversity from a local

(multicultural) and global (international) perspective, and programs, majors, and co-curricular organizations tend to be rooted in one of these perspectives. This divide is reflected in the types of quantitative measures used to assess awareness of issues, topics, content, and perspectives inherent to diversity, which tend to include content from either a multicultural or international perspective.

Henderson et al. (2015) noted the importance of both perspectives, rather than just one, in order to represent a truly intercultural perspective. Given the AICS-R was developed to address a gap in the literature and a need to develop a more comprehensive measure of intercultural competence for students in higher education, the AICS-R could be used to compare across courses and curriculum. Whether they are rooted in a local or global perspective, the AICS-R could be used to determine how effective a course or curriculum is at developing intercultural competencies. The AICS-R may be useful in in assessing how well different programs and universities prepare students to enter an increasingly diverse and globalized workforce and function as interculturally competent professionals. At the program level, faculty members involved in departmental assessment could use the measure to assess graduating seniors to evaluate their preparedness to work with diverse individuals. At the university-wide level, assessment specialists interested in measuring students' intercultural competence may find the AICS-R useful. Given the increased focus on accountability in higher education (Ewell, 2009), universities are increasingly faced with the need to demonstrate the skills and abilities students cultivate over the course of their education. Using the AICS-R as a pre- and post-test measure could help evaluate the effectiveness of the university in fostering the

development of competencies associated with intercultural knowledge, attitudes, and communication over the course of their education.

In addition to using the AICS-R to evaluate program and university-wide effectiveness for accountability purposes, the AICS-R could also be used in assessment for learning improvement purposes. Specifically, the AICS-R could be used as a tool to determine if curriculum reform is warranted. Students in psychology and public/international affairs scored consistently higher than students in other majors and programs of study, suggesting that students in other fields do not report feeling as interculturally competent as their peers. Faculty members involved with departmental assessment could develop program maps to determine where and how often students are being exposed to diversity content and evaluate opportunities in the curriculum to increase this exposure to content. Assessment specialists at the university-wide level could consider a similar process, where they determine where students are exposed to diversity content in their general education curriculum in addition to their program coursework. These evaluations and curriculum maps could be used to inform areas for change to curriculum. Changes could include incorporating more diversity-focused courses, infusing content-related diversity throughout the curriculum, expanding the coverage of content covered (especially if students appear more competent in topics related to either multicultural or international diversity), and incorporating novel and empirically-supported pedagogical practices effective for teaching about diversity.

This study also has implications for research regarding different approaches to addressing content in courses and curriculum. Research about the types of courses or curriculums that best introduce and address content related to diversity is limited, with

most literature providing speculation with minimal empirical evaluation. The AICS-R could be used to further evaluate the effectiveness of different approaches to incorporating diversity content into different types of curriculums. The AICS-R may be more appropriate than other existing measures for this purpose, as it represents competencies deemed important from both multicultural and international perspectives. Thus, the AICS-R can be used to assess the effectiveness of courses and curricula across a variety of majors and programs, regardless of the emphasis of the program.

Students' responses on the AICS-R could be used to inform whether a standalone, infusion, or combination of these approaches is most effective in developing intercultural competence and preparing students for the workforce. Based on results from this study, it is possible that standalone courses may be most effective for addressing specific, narrowly-defined constructs related to intercultural competence. However, when it comes to broader constructs like intercultural competence, either type of course may be effective. These implications are limited to the course-level. The AICS-R could be used to compare across curricula as well, to see if curricula structures are differentially effective at developing and sustaining perceptions of intercultural competencies. Because there did not appear to be different types of courses, perhaps a combination of approaches, such as that suggested by Dunn et al. (2008), may be better for developing such broad yet complex competencies. If this is the case, then a measure like the AICS-R that includes content from both a multicultural and international perspective may best address this question.

Limitations and Suggestions for Future Research

Although this study does provide support for use of the AICS-R in measuring self-perceptions of intercultural competencies, there are a number of limitations that should be acknowledged. In terms of gender and race/ethnicity, the sample was fairly representative of the university. However, both the sample and university are fairly homogeneous. Future research should attempt to attain a sample that is more representative of a variety of different demographic groups. In terms of class year, senior psychology majors were overrepresented in contrast to other majors. Thus, conclusions drawn from comparisons across class year may be influenced by this rather than due to actual differences.

On a similar note, differences across sociocultural courses were only examined for senior psychology majors. There was minimal variability in terms of the courses and instructors. Thus, these differences may be specific to the university's undergraduate psychology program and not generalizable across programs of study or universities. Moreover, because there were so many different combinations of sociocultural courses completed across the sample of senior psychology majors, the group sizes for each type of combination were too small to support statistical analysis. There was also no logical way to combine these smaller groups, so these students were excluded from analysis. Thus, other types and combinations of sociocultural courses were not compared. Future research should include students who complete both standalone and integrated courses, students who complete study abroad courses, students who complete independent research projects, and any that students may complete. Furthermore, future research

should also compare across standalone and infusion structured curriculums as a whole, in addition to standalone and integrated structured courses.

Of particular interest to future research examining differences in course effectiveness is evaluation of study abroad courses. This study was unable to examine differences between students who completed study abroad courses and other traditional types of courses. Along with these traditional courses offered on-site at universities, many students are now participating in study abroad programs. Ideally, these programs should provide the opportunity for participating students to develop various intercultural competences. Future research using the AICS-R should attempt to sample from enough students who have studied abroad to compare them to students who have not in order to evaluate the effectiveness of such programs.

In an attempt to examine construct validity evidence in terms of relationships with external variables, scores on the AICS-R were correlated with two measures based on multiculturalism. The AICS-R was developed in an attempt to incorporate content from both the local and global perspectives of diversity, in order to fully cover the construct of intercultural competencies. However, measures of global diversity developed from the international perspective tend to either be commercially available or not well-established. In contrast, measures of local diversity developed from the multicultural perspective tend to be publically available and encouraged for use in research, and are much more easily attainable. Thus, only measures from the multicultural perspective were included as external measures in this study. Future research should include a measure of international or global diversity as a component of the validity evidence gathered, and compare scores from this measure to scores from the AICS-R. Because of the strong

correlation between scores from the AICS-R and scores from the multicultural measures, it is important to see if scores from the AICS-R are also strongly correlated with an internationally focused measure, or if there is a weaker relationship. A stronger relationship would provide better support for the validity of the AICS-R in measuring intercultural competence. However, if there is a weaker relationship, this could suggest that the AICS-R is over representing the multicultural perspective. Furthermore, a measure that could help tease out the nuances of the five subscales would be useful.

Inherent to self-report measures are limitations related to response styles and bias (Bandalos, in prep). Self-report measures generally consist of affective items that require participants to indicate their own perceptions and attitudes. Systematic inaccuracies or response styles may introduce construct-irrelevant variance into what is measured by the items. Response styles include systematic responses to items that represent something other than the construct intended to be measured (Paulhaus, 1991). These responses occur over a range of items, rather than just one of a few items. Thus, overall scores indicate not only the intended construct but the systematic way of responding. A time of response style that is particularly common with items measuring sensitive content, such as views and attitudes toward underrepresented groups, is social desirable responding. When participants respond in socially desirable ways, responses are biased toward what is socially acceptable or right, rather than a true representation of their attitude or belief (Maccoby & Maccoby, 1954).

Social desirability scores were correlated with scores on the other measures in order to gain insight into the extent to which scores from other measures may be influenced by socially desirable responding. The relationship between social desirability

scores and scores on each of the other measures was negligible. However, conclusions about the extent to which socially desirable responding is present in the scores should be made cautiously. Internal consistency, of the social desirability scores as evidenced by Cronbach's alpha was low, suggesting that scores from this measure may not be reliable. Use of the Marlowe-Crowne Social Desirability scale and its various available abbreviated forms as a control for social desirability has been discouraged based on inconsistencies in model adequacy across independent samples (Barger, 2002). Thus, although scores from this measure may be used to evaluate the extent to which response bias is captured in the scores, this measure is not useful as a way to control for such bias. Additionally, the measure may not produce reliable scores, meaning that conclusions drawn from scores may not be accurate or trustworthy.

A more effective way to overcome limitations inherent to self-report measures may be to include compare scores from a direct measure to scores on the measure of interest. Scores from a direct measure of intercultural competence should be positively related scores from a self-report measure of intercultural competence. Specifically, respondents who exhibit greater intercultural competence as evidenced by a direct measure should also self-report they are high in intercultural competence; whereas as students who exhibit less intercultural competence on the direct measure should rate themselves as less interculturally competent. If a strong relationship were not found between the direct measure of intercultural competence and scores from the AICS-R, this would suggest that the AICS-R is measuring some other construct(s) in addition to intercultural competencies, such as socially desirable responding. If there is a strong

relationship, this would suggest that the AICS-R is not measuring response styles inherent to self-report measures, such as social desirability, to a great extent.

In terms of scale development, there are a few suggestions for future research that should be considered. As previously noted, items were grouped by common stem, which may affect the factor loadings. Similar wording and phrasing may influence items to factor together (Bandalos, in prep; Cattell & Tsujioka, 1964). This effect may be exaggerated by the fact that all items were grouped by subscale, and thus, common stem. In future administrations of the survey, items should be randomized so that they are not grouped together by their common stem/subscale. Factor analytic procedures that still support the five-factor structure based on Henderson's (2015) five domains of intercultural competencies would provide stronger support for the internal structure of the instrument.

Although the factors suggested by the EFA are aligned with the theoretical domains of intercultural competence as defined by Henderson et al. (2015), EFA is an exploratory technique. Thus, performing a confirmatory factor analysis (CFA) on an independent sample would aid in further evaluating the internal structure of the AICS-R. Additionally, although the correlations among factors suggested that scoring the items as separate subscales is most appropriate for the measure, CFA would help to further evaluate this suggestion. CFA would also help evaluate if including the two cross-loading and three low-loading items is appropriate, of if they should be removed from the instrument. CFA could also help to evaluate measure invariance across groups.

Lastly, future research should evaluate the content of the AICS-R to ensure coverage of both local and global diversity content. This is especially important

considering the external measures included in the current study were limited to local diversity measures. Furthermore, items were revised and some items were removed prior to the administration of the AICS-R. Future research using a CFA may suggest further item revisions or the removal of other items that perform poorly. In order to make sure the construct is still fully represented as intended by Henderson et al. (2015), items should be evaluated for content. Items could be mapped to a local or global perspective, to make sure both are adequately covered. Additionally, a backward translation in which items are mapped onto their suggested factors could help to ensure that the domains outlined by Henderson et al. (2015) are still sufficiently represented.

Research on the AICS-R is still in its preliminary stages. It is tempting to consider implications for curriculum structure, pedagogical practice, and assessment of intercultural competence and similar constructs, based on the results from this study. However, until further evidence is collected supporting the reliability and validity of the use of the AICS-R to assess intercultural competence in higher education students, conclusions drawn based on group differences should be interpreted with caution. Nevertheless, analyses on of the AICS-R in its current form suggest that the AICS-R exhibits sufficient reliability and validity to be used in the current study. Furthermore, results from this study provide a basis for future research regarding curriculum, pedagogy, and assessment.

Conclusions

The current study sought to investigate the psychometric properties of and validity evidence for the AICS-R. The AICS-R contributes to literature on intercultural competency by defining the construct in terms of two traditionally distinct perspectives of

diversity that tend to be conceptualized either locally (multicultural) or globally (international). Among scholars, there is a lack of consistency in terms of both the perspective generally taken when approaching diversity in higher education and the most effective ways to include content related to diversity and intercultural competence. Regardless of perspective, though, this topic is important to researchers, students, and employers alike. In recent years, the workplace has become increasingly diverse and globalized. Thus, those involved with higher education have increasingly focused on improving intercultural knowledge and communication skills in order to equip graduates to work with people from diverse cultural backgrounds. Although a variety of quantitative measures have been developed to assess constructs similar to Henderson's conceptualization of intercultural competence, these measures are limited in scope, perspective, empirical support, and availability. Henderson et al. attempted to address these limitations in constructing the original form of the AICS-R.

Benson's (1998) strong program for construct validity was used as a framework through which to evaluate evidence regarding the theory on which the instrument is based, the structural qualities of the instrument, and how the scores on the instrument relate to other constructs in expected ways. Prior to constructing items for the scale, Henderson et al. (2015) conducted an extensive literature review to identity domains representing the knowledge, skills, and abilities that an interculturally competent professional would embody. The developers of the scale next constructed items to represent all of these domains. In the initial testing of the scale, investigation of the internal structure of the scale were limited to reliability estimates for each of the domains, item intercorrelations, and descriptive statistics. The sample collected during the initial

testing was too small to support further analysis, such as EFA. After revising items on the scale, the current study collected a large enough sample size, including both undergraduate and graduate students, to further examine the internal structure of the AICS-R. Results from an EFA supported a five-factor model that strongly corresponds to the five hypothesized domains of the construct. Both the total AICS-R and the five proposed subscales appeared to be internally consistent. The five subscales seem to be related, but are also relatively distinct.

In terms of external evidence, scores from the AICS-R were compared to external measures and group differences across the total and scale scores on the AICS-R were evaluated. Scores from the AICS-R were positively related to scores from measures reflecting multicultural awareness, knowledge, and experience. Furthermore, scores on the AICS-R were negatively related to attitudes of racial color-blindness. These relationships are aligned with what literature regarding intercultural competence would suggest. Taken together, this provides preliminary evidence suggesting that the AICS-R is measuring intercultural competence. Group differences suggest that students who are likely to have had more multicultural and intercultural experiences scored higher on both the AICS-R total and each of the subscales.

Overall, the current study provides preliminary support for use of the AICS-R as a reliable and valid measure of intercultural competence in higher education. Future research examining both the internal structure and external relationships should be conducted on the AICS-R. This research would inform revisions to item wordings and stems and help determine if any items should be removed. Investigation of external relationships would help to evaluate whether intercultural competencies are being

measured as intended, particularly in regard to representing both multicultural and international perspectives of diversity adequately.

Appendix

Revised version of AICS (AICS-R) administered in the current study

Knowledge subscale

I can explain how…

1. Human diversity leads to many different cultural ways of life
2. Social, political, and historical events shape diverse cultural identities
3. Global events can affect local ways of life
4. How my culture and/or country of origin fits into a global context
5. Local political, cultural, and/or technological change in one culture affects cultures around the world
6. Someone from another culture may view my culture
7. Different aspects of identity combine to create different experiences (e.g. sexual identity and race combine to create different experiences)
8. People in others cultures perceive equality, opportunity, and human rights differently than I do
9. The social, political, and historical events in my country have affected my perceptions of diversity
10. The distribution of power and privilege varies across cultures and /or countries

Communication Subscale

I am skilled at discussing diversity issues related to….

11. Race and ethnicity
12. Nationality
13. Ability and disability
14. Social class
15. Gender identity
16. Religion
17. Sexual orientation
18. Language

Attitudes subscale

I…

19. Understand my values, beliefs, and communication style may come across to people from other cultures/countries.
20. Adapt my assumptions about diverse ways of life in different cultures/countries
21. Seek out and learn from intercultural opportunities as a way of life
22. Respond with openness when I encounter unfamiliar ways of life in different cultures/countries
23. Demonstrate an appreciation of unfamiliar ways of life in different cultures/countries
24. See my intercultural competency growth as a lifelong process
25. Confront stereotypes, prejudice, or racism in social situations

Professional practice subscale

I am skilled at …

26. Explaining human rights issues from the perspective of another culture/country
27. Using intercultural information to solve problems in my field
28. Accommodating how people from diverse cultures may think, analyze, and process information differently
29. Incorporating an atmosphere of intercultural cooperation to my work/school setting
30. Being flexible with how people from diverse cultures may approach their school/work (working pace, attendance, participation, solutions to problems)
31. Understanding how global events relate to local problems in my field
32. Understanding that people of diverse cultures may use language in different ways to express their ideas

Negotiated space subscale

I am skilled at …

33. Building trust and cooperation among team members of diverse cultures
34. Making decisions that people from different backgrounds can agree with
35. Helping resolve misunderstandings in diverse groups
36. Adapting my communication style and vocabulary for people from different cultures
37. Developing new professional relationships with people from diverse cultures
38. Explaining issues from team members' different cultural perspectives
39. Finding solutions that maintain a balance between team members' diverse worldviews
40. Fostering team member relationships
41. Tolerating periods of confusion as team members of diverse cultures work together to approach and solve problems
42. Understanding how my own cultural heritage may challenge group trust and cooperation
43. Accommodating language differences to enhance communication between team members
44. Explaining the dynamics of power and privilege that may be complicating the team process
45. Putting priority on team member relationships when working through difficult issues on a project

Tables and Figures

Table 1. *Descriptive Statistics for AICS-R by Item (7-point Likert-typle scale; 1 = Strongly disagree, 7 = Strongly agree)*

Item	Mean	SD	Skewness[a]	Kurtosis[b]
1	5.93	1.10	-1.62	3.56
2	5.93	1.12	-1.63	3.40
3	5.86	1.11	-1.49	2.61
4	5.61	1.14	-0.97	0.88
5	5.53	1.13	-1.00	1.16
6	5.70	1.14	-1.13	1.35
7	5.97	1.10	-1.42	2.66
8	5.67	1.20	-1.14	1.31
9	5.75	1.14	-1.18	1.75
10	5.94	1.10	-1.27	2.04
11	5.19	1.33	-0.72	0.26
12	4.98	1.34	-0.51	-0.14
13	5.15	1.36	-0.75	0.19
14	5.42	1.22	-0.78	0.47
15	5.17	1.48	-0.76	0.11
16	5.20	1.45	-0.81	0.12
17	5.16	1.47	-0.77	0.09
18	5.02	1.39	-0.61	-0.13
19	5.69	1.04	-1.02	1.58
20	5.70	1.03	-0.95	1.40
21	5.76	1.13	-1.00	1.22
22	5.96	0.99	-1.21	2.41
23	5.99	1.01	-1.24	2.24
24	5.99	1.05	-1.26	2.01
25	5.72	1.15	-0.91	0.80

26	4.49	1.44	-0.34	-0.58
27	4.96	1.33	-0.51	-0.10
28	5.2	1.24	-0.73	0.47
29	5.27	1.25	-0.70	0.33
30	5.48	1.16	-0.95	1.06
31	5.37	1.23	-0.82	0.54
32	5.73	1.08	-1.14	1.85
33	5.57	1.08	-0.78	0.63
34	5.45	1.05	-0.59	0.24
35	5.43	1.13	-0.74	0.57
36	5.38	1.19	-0.76	0.57
37	5.54	1.13	-0.82	0.81
38	5.21	1.18	-0.51	0.04
39	5.33	1.12	-0.57	0.17
40	5.69	1.05	-0.70	0.60
41	5.45	1.13	-0.72	0.49
42	5.45	1.19	-0.84	0.74
43	5.09	1.29	-0.50	-0.10
44	5.16	1.28	-0.59	0.16
45	5.41	1.17	-0.68	0.48

[a]*SE = .091*

[b]*SE = .181*

Table 2. *Inter-Item Correlation Matrix for AICS-R*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	-																						
2	0.73	-																					
3	0.67	0.74	-																				
4	0.55	0.53	0.60	-																			
5	0.54	0.58	0.61	0.64	-																		
6	0.42	0.40	0.38	0.47	0.46	-																	
7	0.52	0.50	0.48	0.45	0.52	0.56	-																
8	0.50	0.50	0.50	0.47	0.45	0.43	0.50	-															
9	0.51	0.49	0.53	0.51	0.50	0.41	0.53	0.50	-														
10	0.58	0.58	0.57	0.49	0.53	0.49	0.63	0.52	0.60	-													
11	0.39	0.36	0.33	0.38	0.36	0.32	0.42	0.29	0.37	0.38	-												
12	0.35	0.35	0.37	0.41	0.43	0.30	0.34	0.31	0.37	0.38	0.69	-											
13	0.27	0.21	0.25	0.26	0.31	0.21	0.31	0.20	0.23	0.24	0.44	0.47	-										
14	0.40	0.36	0.37	0.38	0.37	0.30	0.38	0.33	0.40	0.37	0.64	0.60	0.47	-									
15	0.26	0.22	0.25	0.26	0.25	0.25	0.44	0.24	0.32	0.27	0.51	0.39	0.50	0.53	-								
16	0.25	0.27	0.27	0.31	0.31	0.18	0.20	0.19	0.25	0.23	0.42	0.49	0.38	0.46	0.30	-							
17	0.28	0.24	0.27	0.28	0.30	0.30	0.42	0.30	0.34	0.33	0.53	0.44	0.49	0.53	0.75	0.39	-						
18	0.32	0.32	0.29	0.30	0.34	0.24	0.32	0.28	0.27	0.30	0.49	0.60	0.46	0.50	0.42	0.45	0.45	-					
19	0.43	0.37	0.33	0.37	0.36	0.35	0.37	0.36	0.35	0.39	0.41	0.40	0.34	0.37	0.22	0.34	0.28	0.34	-				
20	0.41	0.37	0.33	0.36	0.37	0.37	0.43	0.33	0.40	0.45	0.41	0.43	0.32	0.40	0.36	0.30	0.41	0.39	0.59	-			
21	0.41	0.40	0.39	0.34	0.37	0.33	0.44	0.31	0.40	0.45	0.45	0.44	0.34	0.39	0.37	0.32	0.35	0.37	0.49	0.66	-		
22	0.36	0.36	0.35	0.33	0.38	0.35	0.44	0.32	0.37	0.42	0.42	0.42	0.26	0.38	0.34	0.30	0.37	0.33	0.46	0.65	0.66	-	
23	0.42	0.40	0.40	0.33	0.37	0.38	0.51	0.36	0.41	0.48	0.45	0.39	0.29	0.35	0.35	0.23	0.36	0.31	0.43	0.59	0.65	0.75	-
24	0.43	0.40	0.38	0.37	0.39	0.35	0.52	0.35	0.43	0.53	0.45	0.40	0.29	0.40	0.35	0.27	0.34	0.33	0.48	0.62	0.64	0.61	0.68
25	0.38	0.37	0.35	0.32	0.33	0.34	0.45	0.33	0.35	0.46	0.45	0.43	0.30	0.42	0.38	0.29	0.45	0.35	0.44	0.49	0.49	0.55	0.56
26	0.35	0.33	0.32	0.39	0.41	0.36	0.28	0.32	0.36	0.35	0.43	0.48	0.31	0.41	0.30	0.32	0.38	0.43	0.43	0.39	0.40	0.37	0.32
27	0.38	0.34	0.36	0.37	0.44	0.35	0.36	0.32	0.39	0.37	0.45	0.52	0.39	0.47	0.35	0.33	0.42	0.48	0.42	0.46	0.50	0.45	0.40
28	0.38	0.39	0.36	0.38	0.39	0.37	0.39	0.34	0.39	0.39	0.45	0.49	0.34	0.48	0.39	0.32	0.41	0.47	0.44	0.49	0.56	0.50	0.50
29	0.42	0.37	0.35	0.38	0.41	0.35	0.46	0.34	0.40	0.43	0.46	0.50	0.37	0.48	0.37	0.30	0.41	0.46	0.42	0.50	0.56	0.53	0.52
30	0.36	0.34	0.32	0.34	0.38	0.31	0.36	0.30	0.33	0.39	0.40	0.43	0.35	0.36	0.28	0.26	0.33	0.36	0.43	0.48	0.52	0.51	0.52

31	0.41	0.37	0.40	0.40	0.48	0.31	0.40	0.32	0.41	0.43	0.44	0.48	0.34	0.47	0.28	0.31	0.29	0.39	0.42	0.44	0.51	0.47	0.44
32	0.43	0.38	0.38	0.36	0.41	0.39	0.43	0.31	0.42	0.47	0.45	0.42	0.32	0.43	0.30	0.27	0.32	0.37	0.41	0.47	0.54	0.49	0.48
33	0.31	0.29	0.27	0.29	0.31	0.27	0.29	0.24	0.24	0.30	0.30	0.29	0.31	0.31	0.26	0.17	0.27	0.26	0.44	0.41	0.41	0.43	0.43
34	0.32	0.26	0.26	0.31	0.33	0.32	0.32	0.21	0.23	0.27	0.35	0.36	0.33	0.35	0.27	0.29	0.33	0.34	0.46	0.42	0.43	0.48	0.44
35	0.31	0.28	0.24	0.26	0.31	0.30	0.34	0.22	0.19	0.30	0.36	0.38	0.32	0.38	0.30	0.28	0.32	0.38	0.43	0.47	0.46	0.47	0.44
36	0.31	0.30	0.27	0.30	0.36	0.26	0.33	0.21	0.26	0.30	0.29	0.34	0.27	0.28	0.23	0.25	0.21	0.36	0.39	0.44	0.45	0.43	0.43
37	0.33	0.31	0.28	0.30	0.35	0.27	0.39	0.22	0.29	0.34	0.40	0.39	0.33	0.36	0.29	0.27	0.29	0.42	0.41	0.47	0.50	0.44	0.49
38	0.33	0.31	0.31	0.36	0.36	0.30	0.39	0.30	0.33	0.36	0.38	0.42	0.30	0.36	0.30	0.26	0.31	0.43	0.42	0.46	0.48	0.41	0.44
39	0.34	0.27	0.31	0.30	0.35	0.29	0.39	0.27	0.33	0.34	0.40	0.43	0.33	0.39	0.33	0.29	0.36	0.39	0.41	0.47	0.51	0.47	0.50
40	0.34	0.25	0.27	0.29	0.29	0.26	0.37	0.23	0.28	0.32	0.32	0.32	0.33	0.39	0.30	0.25	0.32	0.28	0.36	0.39	0.40	0.39	0.45
41	0.33	0.26	0.26	0.30	0.36	0.29	0.37	0.27	0.30	0.33	0.37	0.36	0.31	0.37	0.30	0.26	0.32	0.28	0.36	0.43	0.45	0.48	0.48
42	0.34	0.29	0.26	0.34	0.31	0.30	0.37	0.26	0.38	0.33	0.37	0.39	0.32	0.32	0.32	0.21	0.32	0.29	0.38	0.46	0.49	0.45	0.39
43	0.24	0.24	0.23	0.30	0.36	0.28	0.31	0.23	0.29	0.27	0.34	0.42	0.31	0.29	0.26	0.31	0.26	0.45	0.33	0.37	0.43	0.35	0.35
44	0.29	0.27	0.27	0.32	0.39	0.28	0.38	0.27	0.35	0.35	0.44	0.43	0.38	0.42	0.38	0.27	0.37	0.36	0.41	0.44	0.49	0.42	0.41
45	0.27	0.22	0.25	0.29	0.29	0.24	0.35	0.24	0.25	0.30	0.35	0.31	0.32	0.34	0.32	0.22	0.32	0.29	0.37	0.43	0.44	0.43	0.44

Table 2 (continued). *Inter-Item Correlation Matrix for AICS-R*

Item	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
24	-																					
25	0.55	-																				
26	0.302	0.395	-																			
27	0.43	0.42	0.654	-																		
28	0.492	0.434	0.577	0.687	-																	
29	0.511	0.493	0.491	0.642	0.712	-																
30	0.464	0.405	0.406	0.479	0.584	0.656	-															
31	0.452	0.429	0.445	0.548	0.542	0.588	0.591	-														
32	0.506	0.393	0.366	0.484	0.557	0.551	0.634	0.616	-													
33	0.389	0.416	0.34	0.394	0.366	0.441	0.47	0.386	0.386	-												
34	0.33	0.425	0.389	0.425	0.444	0.462	0.495	0.411	0.432	0.69	-											
35	0.395	0.422	0.378	0.449	0.458	0.496	0.459	0.398	0.379	0.658	0.734	-										
36	0.381	0.393	0.38	0.413	0.422	0.487	0.462	0.426	0.398	0.479	0.528	0.58	-									
37	0.47	0.41	0.332	0.455	0.5	0.533	0.512	0.476	0.468	0.563	0.596	0.591	0.658	-								
38	0.437	0.471	0.435	0.47	0.507	0.497	0.477	0.458	0.416	0.538	0.569	0.607	0.587	0.645	-							
39	0.422	0.514	0.431	0.504	0.504	0.537	0.537	0.451	0.462	0.56	0.625	0.623	0.544	0.664	0.735	-						
40	0.408	0.421	0.297	0.384	0.38	0.395	0.397	0.364	0.38	0.581	0.573	0.557	0.42	0.566	0.517	0.605	-					
41	0.379	0.4	0.302	0.396	0.436	0.46	0.511	0.394	0.422	0.515	0.56	0.545	0.511	0.54	0.539	0.612	0.602	-				
42	0.419	0.396	0.369	0.442	0.484	0.492	0.474	0.427	0.442	0.452	0.503	0.506	0.487	0.498	0.533	0.524	0.469	0.562	-			
43	0.33	0.365	0.342	0.449	0.463	0.449	0.365	0.372	0.398	0.427	0.473	0.473	0.542	0.575	0.574	0.561	0.444	0.529	0.486	-		
44	0.454	0.487	0.462	0.502	0.512	0.488	0.459	0.472	0.432	0.5	0.514	0.547	0.492	0.561	0.605	0.619	0.521	0.539	0.545	0.576	-	
45	0.41	0.401	0.287	0.375	0.377	0.452	0.449	0.394	0.415	0.533	0.522	0.524	0.427	0.52	0.524	0.57	0.59	0.582	0.491	0.494	0.666	-

Determinant = 1.497E-14

Note: all correlations were significant at the $p < .001$ level.

Table 3. *Pattern Coefficients of Item Loadings and Communalities*

Item	1. Negotiated Space	2. Knowledge	3. Communication	4. Attitudes	5. Professional Practice	h^2
1. I can explain how human diversity leads to many different cultural ways of life	0.03	**0.78**	-0.04	0.01	0.00	0.62
2. I can explain how social, political, and historical events shape diverse cultural identities	-0.05	**0.85**	-0.09	-0.02	0.05	0.64
3. I can explain how global events can affect local ways of life	-0.02	**0.88**	-0.04	-0.09	0.02	0.65
4. I can explain how my culture and/ or country of origin fits into a global context	0.07	**0.73**	0.02	-0.19	0.11	0.55
5. I can explain how local, political, cultural, and/ or technological change in one culture affects cultures around the world	0.09	**0.71**	-0.02	-0.17	0.17	0.58
6. I can explain how someone from another culture may view my culture	0.04	**0.49**	0.02	0.11	-0.01	0.36
7. I can explain how different aspects of identity combine to create different experiences (e.g. sexual identity and race combine to create different experiences)	0.08	**0.55**	0.17	0.28	-0.25	0.58
8. I can explain how people in others cultures perceive equality, opportunity, and human rights differently than I do	-0.03	**0.65**	0.03	0.04	-0.04	0.44
9. I can explain how the social, political, and historical events in my country have affected my perceptions of diversity	-0.08	**0.60**	0.07	0.12	0.04	0.49
10. I can explain how the distribution of power and privilege varies across cultures and/ or countries	-0.06	**0.67**	-0.02	0.27	-0.07	0.61
11. I am skilled at discussing diversity issues related to race and ethnicity	-0.07	0.03	**0.56**	0.10	0.24	0.58
12. I am skilled at discussing diversity issues related to nationality	-0.06	0.03	**0.44**	-0.07	**0.51**	0.61
13. I am skilled at discussing diversity issues related to ability and disability	0.13	-0.04	**0.55**	-0.10	0.13	0.41

14. I am skilled at discussing diversity issues related to social class	-0.03	0.09	**0.59**	-0.05	0.24	0.58
15. I am skilled at discussing diversity issues related to gender identity	0.02	-0.05	**0.84**	0.12	-0.19	0.63
16. I am skilled at discussing diversity issues related to religion	-0.01	0.04	**0.41**	-0.13	0.29	0.32
17. I am skilled at discussing diversity issues related to sexual orientation	-0.01	-0.03	**0.84**	0.10	-0.10	0.66
18. I am skilled at discussing diversity issues related to language	0.03	-0.02	**0.44**	-0.15	**0.45**	0.49
19. I am able to understand how my values, beliefs, and communication style may come across to people from other cultures/ countries	0.16	0.14	0.00	0.25	0.20	0.40
20. I am able to adapt my assumptions about diverse ways of life in different cultures/countries	0.06	0.00	0.05	**0.61**	0.12	0.58
21. I am able to seek out and learn from intercultural opportunities as a way of life	0.04	-0.05	-0.02	**0.65**	0.23	0.64
22. I am able to respond with openness when I encounter unfamiliar ways of life in different cultures/ countries	0.03	-0.07	-0.02	**0.78**	0.10	0.66
23. I am able to demonstrate an appreciation of unfamiliar ways of life in different cultures/ countries	0.07	0.04	-0.02	**0.83**	-0.08	0.71
24. I am able to see my intercultural competency growth as a lifelong process	-0.04	0.08	0.02	**0.75**	0.03	0.64
25. I am able to confront stereotypes, prejudice, or racism in social situations	0.17	0.07	0.19	0.39	-0.01	0.47
26. I am skilled at explaining human rights issues from the perspective of another culture/ country	0.04	0.09	0.11	-0.09	**0.60**	0.48
27. I am skilled at using intercultural information to solve problems in my field	0.05	0.00	0.09	0.02	**0.68**	0.61
28. I am skilled at accommodating how people from diverse cultures may think, analyze, and process information differently	0.01	-0.05	0.03	0.22	**0.66**	0.64
29. I am skilled at incorporating an atmosphere of intercultural cooperation to my work/ school setting	0.09	-0.03	0.01	0.27	**0.54**	0.63
30. I am skilled at being flexible with how people from diverse cultures may approach their school/ work	0.20	-0.04	-0.11	0.32	**0.43**	0.54

(working pace, attendance, participation, solutions to problems)						
31. I am skilled at understanding how global events relate to local problems in my field	0.05	0.11	-0.06	0.16	**0.54**	0.53
32. I am skilled at understanding that people of diverse cultures may use language in different ways to express their ideas	0.06	0.10	-0.05	0.32	0.38	0.50
33. I am skilled at building trust and cooperation among team members of diverse cultures	**0.81**	0.05	-0.07	0.02	-0.11	0.56
34. I am skilled at making decisions that people from different backgrounds can agree with	**0.83**	-0.02	-0.01	-0.05	0.01	0.62
35. I am skilled at helping resolve misunderstandings in diverse groups	**0.81**	-0.07	0.02	-0.01	0.02	0.62
36. I am skilled at adapting my communication style and vocabulary for people from different cultures	**0.62**	0.01	-0.15	0.02	0.20	0.51
37. I am skilled at developing new professional relationships with people from diverse cultures	**0.71**	-0.04	-0.06	0.05	0.13	0.62
38. I am skilled at explaining issues from team members' different cultural perspectives	**0.72**	0.05	-0.03	-0.06	0.14	0.62
39. I am skilled at finding solutions that maintain a balance between team members' diverse world views	**0.76**	-0.03	0.02	0.02	0.07	0.67
40. I am skilled at fostering team member relationships	**0.80**	0.06	0.10	0.00	-0.24	0.56
41. I am skilled at tolerating periods of confusion as team members of diverse cultures work together to approach and solve problems	**0.73**	0.03	0.02	0.07	-0.09	0.56
42. I am skilled at understanding how my own cultural heritage may challenge group trust cooperation	**0.51**	0.03	0.00	0.11	0.11	0.46
43. I am skilled at accommodating language differences to enhance communication between team members	**0.62**	-0.01	0.03	-0.16	0.21	0.49
44. I am skilled at explaining the dynamics of power and privilege that may be complicating the team process	**0.62**	-0.02	0.13	-0.02	0.10	0.57
45. I am skilled at putting priority on team member relationships when working through difficult issues on a project	**0.74**	-0.03	0.08	0.09	-0.15	0.54

Note. Pattern coefficients > .4 are in bold. h^2 = communalities.

Table 4. *Structure Coefficients of Item Loadings*

Item	1. Negotiated Space	2. Knowledge	3. Communication	4. Attitudes	5. Professional Practice
1. I can explain how human diversity leads to many different cultural ways of life	0.41	0.78	0.39	0.49	0.46
2. I can explain how social, political, and historical events shape diverse cultural identities	0.35	0.80	0.34	0.45	0.45
3. I can explain how global events can affect local ways of life	0.35	0.80	0.36	0.42	0.44
4. I can explain how my culture and/ or country of origin fits into a global context	0.40	0.73	0.39	0.39	0.49
5. I can explain how local, political, cultural, and/ or technological change in one culture affects cultures around the world	0.44	0.74	0.40	0.42	0.53
6. I can explain how someone from another culture may view my culture	0.37	0.59	0.35	0.45	0.39
7. I can explain how different aspects of identity combine to create different experiences (e.g. sexual identity and race combine to create different experiences)	0.46	0.70	0.50	0.61	0.39
8. I can explain how people in others cultures perceive equality, opportunity, and human rights differently than I do	0.32	0.66	0.36	0.42	0.37
9. I can explain how the social, political, and historical events in my country have affected my perceptions of diversity	0.36	0.69	0.42	0.50	0.45
10. I can explain how the distribution of power and privilege varies across cultures and/ or countries	0.41	0.76	0.41	0.60	0.44
11. I am skilled at discussing diversity issues related to race and ethnicity	0.46	0.48	0.72	0.51	0.58
12. I am skilled at discussing diversity issues related to nationality	0.48	0.48	0.68	0.45	0.69
13. I am skilled at discussing diversity issues related to ability and disability	0.42	0.33	0.62	0.34	0.45

14. I am skilled at discussing diversity issues related to social class	0.45	0.49	0.73	0.45	0.57
15. I am skilled at discussing diversity issues related to gender identity	0.39	0.35	0.78	0.44	0.34
16. I am skilled at discussing diversity issues related to religion	0.33	0.34	0.52	0.29	0.46
17. I am skilled at discussing diversity issues related to sexual orientation	0.41	0.39	0.81	0.46	0.40
18. I am skilled at discussing diversity issues related to language	0.45	0.39	0.61	0.37	0.61
19. I am able to understand how my values, beliefs, and communication style may come across to people from other cultures/ countries	0.53	0.49	0.40	0.56	0.54
20. I am able to adapt my assumptions about diverse ways of life in different cultures/countries	0.57	0.50	0.48	0.75	0.55
21. I am able to seek out and learn from intercultural opportunities as a way of life	0.60	0.50	0.46	0.78	0.61
22. I am able to respond with openness when I encounter unfamiliar ways of life in different cultures/ countries	0.57	0.48	0.44	0.81	0.54
23. I am able to demonstrate an appreciation of unfamiliar ways of life in different cultures/ countries	0.58	0.53	0.44	0.84	0.48
24. I am able to see my intercultural competency growth as a lifelong process	0.53	0.55	0.45	0.80	0.50
25. I am able to confront stereotypes, prejudice, or racism in social situations	0.56	0.49	0.52	0.64	0.49
26. I am skilled at explaining human rights issues from the perspective of another culture/ country	0.47	0.46	0.46	0.40	0.68
27. I am skilled at using intercultural information to solve problems in my field	0.57	0.48	0.51	0.51	0.78
28. I am skilled at accommodating how people from diverse cultures may think, analyze, and process information differently	0.59	0.49	0.50	0.61	0.78
29. I am skilled at incorporating an atmosphere of intercultural cooperation to my work/ school setting	0.62	0.51	0.49	0.65	0.75
30. I am skilled at being flexible with how people from diverse cultures may approach their school/ work (working pace, attendance, participation, solutions to problems)	0.62	0.45	0.39	0.62	0.67

31. I am skilled at understanding how global events relate to local problems in my field	0.55	0.52	0.42	0.56	0.70
32. I am skilled at understanding that people of diverse cultures may use language in different ways to express their ideas	0.55	0.52	0.41	0.62	0.64
33. I am skilled at building trust and cooperation among team members of diverse cultures	0.74	0.37	0.33	0.49	0.44
34. I am skilled at making decisions that people from different backgrounds can agree with	0.79	0.37	0.39	0.49	0.52
35. I am skilled at helping resolve misunderstandings in diverse groups	0.79	0.36	0.41	0.51	0.52
36. I am skilled at adapting my communication style and vocabulary for people from different cultures	0.70	0.38	0.30	0.48	0.55
37. I am skilled at developing new professional relationships with people from diverse cultures	0.78	0.40	0.40	0.55	0.58
38. I am skilled at explaining issues from team members' different cultural perspectives	0.78	0.44	0.41	0.51	0.59
39. I am skilled at finding solutions that maintain a balance between team members' diverse world views	0.82	0.42	0.45	0.56	0.58
40. I am skilled at fostering team member relationships	0.73	0.38	0.42	0.49	0.40
41. I am skilled at tolerating periods of confusion as team members of diverse cultures work together to approach and solve problems	0.74	0.40	0.41	0.53	0.47
42. I am skilled at understanding how my own cultural heritage may challenge group trust cooperation	0.67	0.41	0.40	0.53	0.53
43. I am skilled at accommodating language differences to enhance communication between team members	0.68	0.35	0.39	0.40	0.55
44. I am skilled at explaining the dynamics of power and privilege that may be complicating the team process	0.74	0.42	0.50	0.52	0.57
45. I am skilled at putting priority on team member relationships when working through difficult issues on a project	0.72	0.35	0.42	0.52	0.43

Table 5. *Factor Correlation Matrix*

Factor	1	2	3	4	5
1	-				
2	.505	-			
3	.523	.514	-		
4	.670	.618	.539	-	
5	.669	.586	.561	.601	-

Table 6. *Means, Standard Deviations, and Pearson's r correlation coefficients for all measures*

Measure	Mean(*SD*)	1	2	3	4	5	6	7	8	9
1. AICS-R Total	5.50(0.74)	---								
2. AICS-R Knowledge	5.82(0.82)	.784	---							
3. AICS-R Communication	5.17(1.01)	.794	.536	---						
4. AICS-R Attitudes	5.85(0.83)	.853	.625	.596	---					
5. AICS-R Professional Practice	5.24(0.97)	.880	.636	.646	.724	---				
6. AICS-R Negotiated Space	5.41(0.84)	.854	.500	.546	.692	.706	---			
7. MASQUE	87.29(10.85)	.639	.493	.467	.626	.563	.538	---		
8. MEQ	47.78(7.62)	.658	.467	.479	.614	.614	.581	.546	---	
9. CoBRAS	55.48(15.20)	-.384	-.338	-.339	-.387	-.341	-.240	-.582	-.362	---

Note: All correlation coefficients were significant at the $p < .001$ level.

Table 7. *Effect Sizes (R^2) for correlations among measures*

Measure	1	2	3	4	5	6	7	8	9
1. AICS-R Total	---								
2. AICS-R Knowledge	.615	---							
3. AICS-R Communication	.630	.287	---						
4. AICS-R Attitudes	.728	.391	.355	---					
5. AICS-R Professional Practice	.774	.404	.417	.524	---				
6. AICS-R Negotiated Space	.729	.250	.298	.479	.498	---			
7. MASQUE	.408	.243	.218	.392	.317	.289	---		
8. MEQ	.433	.218	.229	.377	.377	.338	.298	---	
9. CoBRAS	.147	.114	.115	.150	.116	.058	.338	.131	---

Table 8. *Means, Standard Deviations, and ANOVA results for comparisons across class year*

Dependent Variable	Mean (*SD*) First-Year	Mean (*SD*) Sophomores	Mean (*SD*) Juniors	Mean (*SD*) Seniors	Mean (*SD*) Graduate Students	*F*	*df*	*p*	η^2
AICS-R Total	5.22(0.80)	5.48(0.69)	5.42(0.66)	5.78(0.68)	5.43(0.78)	14.22	4, 643	< .001	0.081
AICS-R *Knowledge	5.53(0.97)	5.77(0.83)	5.82(0.73)	6.07(0.68)	5.87(0.46)	11.21	4, 261.49	<.001	0.061
AICS-R Communication	4.86(1.07)	5.16(0.95)	5.07(0.97)	5.47(0.99)	5.19(0.93)	9.67	4, 678	< .001	0.054
AICS-R Attitudes	5.60(0.98)	5.80(0.79)	5.82(0.82)	6.09(0.72)	5.82(0.83)	9.41	4, 704	< .001	0.051
AICS-R Professional Practice	4.90(1.01)	5.22(0.88)	5.27(0.95)	5.57(0.86)	5.21(1.08)	13.35	4, 702	< .001	0.071
AICS-R Negotiated Space	5.27(0.91)	5.37(0.84)	5.30(0.87)	5.67(0.77)	5.20(0.93)	7.64	4, 687	< .001	0.043
MASQUE	45.54(7.64)	45.67(7.70)	48.46(8.35)	50.0(6.76)	49.26(7.73)	12.23	4, 673	< .001	0.068
MEQ	84.70(10.94)	84.94(10.73)	87.05(12.08)	90.11(10.69)	89.35(10.47)	8.71	4, 662	<.001	0.050
CoBRAS	60.40(13.25)	59.04(15.3)	58.35(14.33)	51.03(114.22)	49.05(17.78)	16.78	4, 674	<.001	0.091

**Assumption of homogeneity of variance violated; adjusted values based on Welch corrections reported.*

Table 9. *Means, Standard Deviations, and ANOVA results for comparisons across programs of study*

Means and Standard Deviations											
Dependent Variable	Mean (*SD*) Psychology Underclassmen	Mean (*SD*) Psychology Seniors	Mean (*SD*) Graduate Psychology Students	Mean (*SD*) Other Graduate Students	Mean (*SD*) Science/Math	Mean (*SD*) Education	Mean (*SD*) Business	Mean (*SD*) Health Sciences/Nursing	Mean (*SD*) Public/International Affairs	Mean (*SD*) Communications	Mean (*SD*) Arts/Humanities
AICS-R Total	5.48 (0.79)	5.93 (0.55)	5.61 (0.67)	5.52 (0.76)	5.27 (0.70)	5.27 (0.65)	5.27 (0.67)	5.16 (0.85)	5.85 (0.66)	5.30 (0.65)	5.58 (0.90)
AICS-R Knowledge	5.73 (0.85)	6.14 (0.64)	6.00 (0.74)	5.99 (0.68)	5.72 (0.76)	5.76 (0.75)	5.65 (0.84)	5.51 (1.04)	6.14 (0.75)	5.65 (0.77)	6.01 (0.85)
AICS-R Communication	5.23 (1.12)	5.70 (0.78)	5.32 (0.89)	5.22 (1.00)	4.85 (1.11)	4.92 (1.09)	4.82 (0.86)	4.83 (1.12)	5.43 (0.86)	5.01 (0.75)	5.27 (0.99)
AICS-R Attitudes	5.87 (0.82)	6.21 (0.62)	6.04 (0.70)	5.96 (0.62)	5.62 (0.83)	5.83 (0.85)	5.58 (0.87)	5.47 (0.93)	6.35 (0.72)	5.65 (0.62)	6.01 (0.88)
AICS-R Professional Practice	5.16 (1.07)	5.78 (0.70)	5.47 (0.93)	5.27 (0.98)	4.99 (0.98)	5.02 (0.77)	5.08 (0.89)	4.86 (1.03)	5.62 (0.84)	4.87 (0.87)	5.37 (1.21)
AICS-R Negotiated Space	5.43 (0.88)	5.84 (0.66)	5.32 (0.78)	5.32 (1.05)	5.28 (0.78)	5.15 (0.98)	5.28 (0.83)	5.09 (0.92)	5.63 (0.87)	5.29 (0.81)	5.41 (1.07)
MEQ	46.10 (7.27)	50.04 (6.27)	50.33 (7.49)	51.10 (8.93)	45.75 (7.70)	45.14 (9.19)	46.42 (7.85)	44.64 (7.30)	53.32 (8.21)	46.73 (7.59)	50.00 (7.22)
MASQUE	87.18 (11.34)	91.36 (9.02)	92.95 (8.08)	90.04 (10.51)	85.99 (10.67)	85.14 (10.68)	83.02 (11.13)	82.69 (11.17)	91.40 (10.71)	85.71 (9.56)	89.15 (10.60)
CoBRAS	57.50 (14.44)	49.95 (13.43)	41.57 (12.15)	51.57 (20.07)	58.53 (15.87)	61.20 (13.06)	61.44 (14.93)	60.35 (14.25)	55.20 (15.93)	56.93 (13.96)	54.36 (17.17)

Table 9 (continued). *Means, Standard Deviations, and ANOVA results for comparisons across programs of study*

	ANOVA results			
Dependent Variable	*F*	*df*	*p*	η^2
*AICS-R Total	12.30	10, 144.83	<.001	.150
*AICS-R Knowledge	5.79	10, 158.31	<.001	.080
*AICS-R Communication	9.26	10, 151.06	<.001	.113
*AICS-R Attitudes	10.02	10, 159.98	<.001	.127
*AICS-R Professional Practice	11.8	10, 158.167	<.001	.127
*AICS-R Negotiated Space	7.78	10, 154.722	<.001	.092
MEQ	7.67	10, 612	<.001	.111
MASQUE	7.29	10, 603	<.001	.108
CoBRAS	7.29	10, 608	<.001	.107

**Assumption of homogeneity of variance violated; adjusted values based on Welch corrections reported.*

Table 10. *Means, Standard Deviations, and t-test results for comparisons across completed diversity courses among senior psychology majors*

Dependent Variable	Mean (*SD*) Standalone	Mean (*SD*) Integrated	*t*	*df*	*p*	*d*	95 CI of Mean Difference
AICS-R Total	6.03(0.54)	5.95(0.60)	0.46	50	.65	0.14	[-.27, .43]
AICS-R Knowledge	6.38(0.50)	6.20(0.55)	1.18	50	.25	0.36	[-.13, .50]
AICS-R Communication	5.61(0.63)	5.63(0.90)	0.78	50	.44	0.24	[-.30, .68]
AICS-R Attitudes	6.09(0.68)	6.15(0.65)	-0.28	50	.78	0.08	[-.46, .35]
AICS-R Professional Practice	5.88(0.57)	5.85(0.65)	0.16	50	.88	0.05	[-.34, .40]
AICS-R Negotiated Space	5.85(0.70)	5.82(0.72)	0.14	50	.89	0.04	[-.39, .45]
MASQUE	51.82(7.65)	49.80(6.34)	1.01	50	.32	0.30	[-2.01, 6.05]
MEQ	90.65(7.75)	91.23(10.30)	-0.21	50	.84	0.06	[-6.26, 5.10]
CoBRAS	44.59(10.47)	53.26(12.81)	-2.42	50	.02	0.73	[-15.86, -1.48]

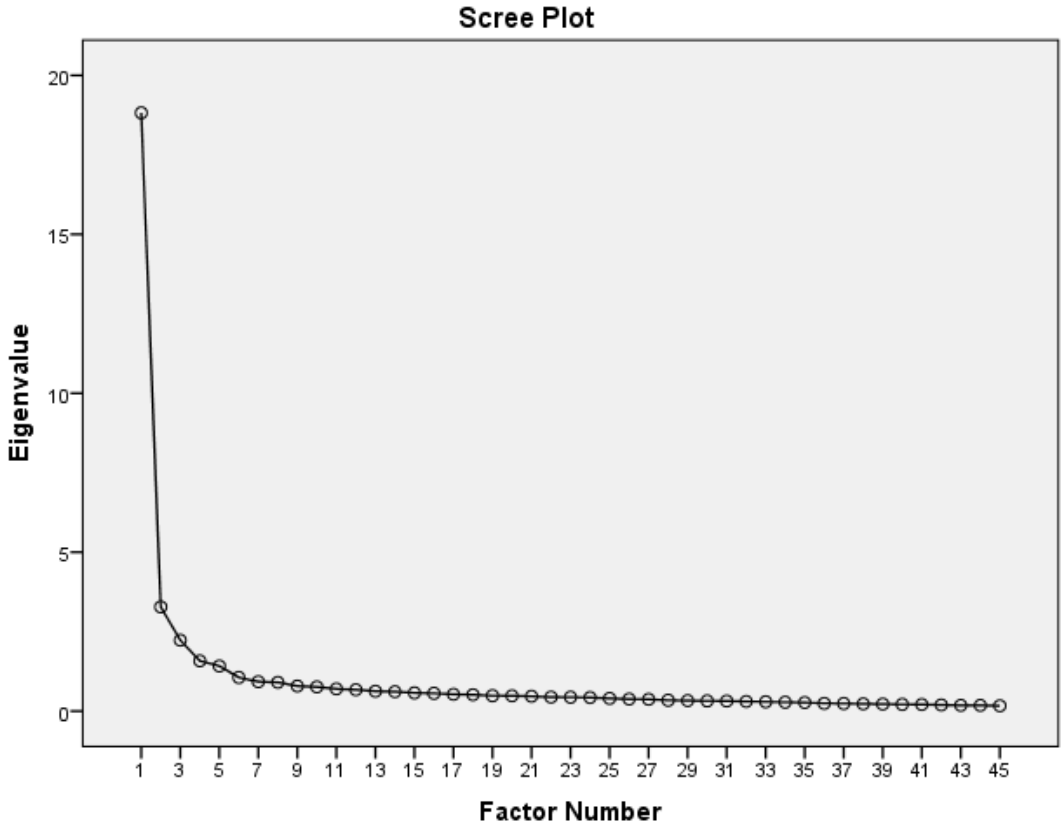

Figure 1. Scree plot for AICS-R items.

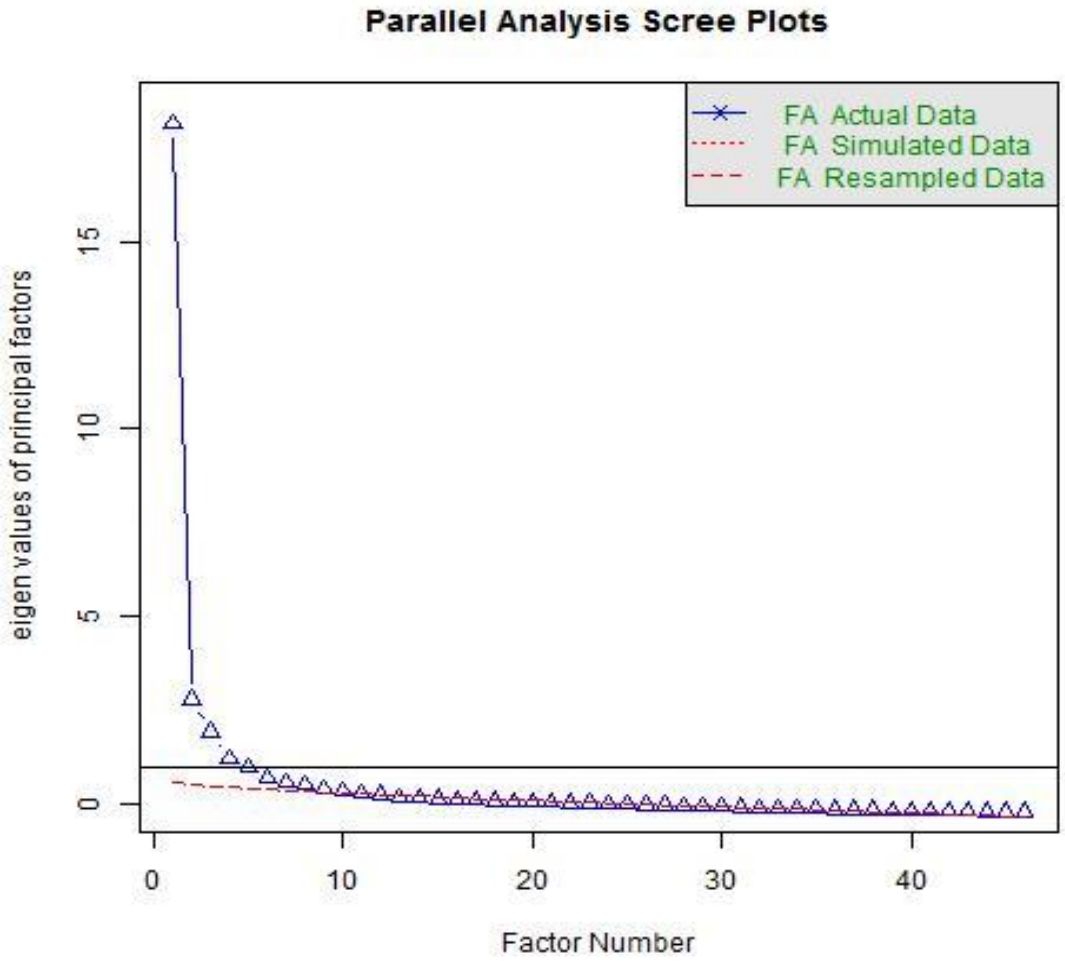

Figure 2. Parallel analysis plot for AICS-R items.

References

American Psychological Association (2003, May). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. American Psychologist, 58(5), 377-402.

American Psychological Association. (2013). APA guidelines for the undergraduate psychology major: Version 2.0. Washington, DC.

Antonio, A. L., Chang, M. J., Hakuta, K., Kenny, D. A., Levin, S., & Milem, J. F. (2004). Effects of racial diversity on complex thinking in college students. American Psychological Society, 15, 507-510.

Arnett, J. J. (2002). The psychology of globalization. American Psychologist, 57 (10), 774.

Arredondo, P., Toporek, R., Brown, S. P., Jones, J., Locke, D. C., Sanchez, J., & Stadler, H. (1996). Operationalization of the multicultural counseling competencies. Journal of Multicultural Counseling and Development, 24, 42-78.

Association of American Colleges and Universities. (2005, December). Wheaton College builds "inclusive excellence" through curriculum innovation. AAC&U News: Insights and Campus Innovations in Liberal Education. Retrieved from www.aacu.org/aacu_news/AACUNEWS05/December05/feature.cfm

Atewologun, D., & Sealy, R. (2014). Experiencing privilege at ethnic, gender and senior Intersections. Journal of Managerial Psychology, 29, 423-439.

Banks, J. A. (1995). Handbook of research on multicultural education. New York: Macmillan.

Banks, J. A., Banks, C. A. M., Cortés, C. E., Hahn, C. L., Merryfield, M. M., Moodley,

K. A., Osler, A., Murphy-Shigematsu, S. & Parker, W. C. (2005). Democracy and diversity: Principles and concepts for educating citizens in a global age. Seattle, WA: Center for Multicultural Education.

Barger, S. D. (2002). The Marlowe-Crowne affair: Short forms, psychometric structure, and social desirability. Journal of Personality Assessment, 79, 286-305.

Barter, C., & Reynold, E. (2000). "I wanna tell you a story": Exploring the application of vignettes in qualitative research with children and young people. International Journal of Social Research Methodology, 3, 307-323.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. Educational Measurement: Issues and Practice, 17, 10-22.

Bonilla-Silva, E. (2006). Racism without racists: Color-blind racism and the persistence of racial inequality in the United States. Lanham, MD: Rowman & Littlefield.

Braskamp, L.A., Braskamp, D.C., Merrill, K. C. & Engberg, M.E. (2013). Global Perspective Inventory (GPI):Its purpose, construction, potential uses, and psychometric properties [Measurement Instrument]. Retrieved from http://gpi.central.edu/supportDocs/manual.pdf

Carr, J. F. (2007). Diversity and disciplinary practices. In J. Branche, J. Mullennix, & E. Cohn (Eds.), Diversity across the curriculum: A guide for faculty in higher education (pp. 30-37). Bolton, MA: Anker.

Cattell, R. B., & Tsujioka, B. (1964). The importance of factor-trueness and validity, versus homogeneity and orthogonality in test scales. Educational and Psychological Measurement, 24, 3–30.

Centre for Information for Language Teachers (CILT). (2008). National occupational

standards for intercultural working: Standards for working with people from different countries or diverse cultures. London, UK: National Centre for Languages.

Chenoweth, K. (1998). Poll confirms that Americans want diversity on campuses. Black Issues in Higher Education, 15, 12

Colby, S. L., & Ortman, J. M. (2015). Projections of the Size and Composition of the US Population: 2014 to 2060. Retrieved from: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf

Constantine, M. G., & Ladany, N. (2000). Self-report multicultural counseling competence scales: Their relation to social desirability attitudes and multicultural case conceptualization ability. Journal of Counseling Psychology, 47, 155.

Corbitt, J. N. (1998). Global awareness profile (1). Nicholas Brealey Publishing.

Cornwell, G .H. & Stoddard, E. W. (1999). Globalizing knowledge Connecting international and intercultural studies. Washington DC: Association of American Colleges and Universities.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. U. Chi. Legal F., 139.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Cross-Cultural World-Mindedness Scale (n.d.) Retrieved from http://www.statisticssolutions.com/cross-cultural-world-mindedness-scale/

Davidson, C. N. (2011). Now you see it: How the brain science of attention will transform the way we live, work, and learn. New York, NY: Viking.

Davis, S. L., & Finney, S. J. (2006). A factor analytic study of the cross-cultural adaptability inventory. Educational and Psychological Measurement, 66, 318-330.

Deardorff, D. K. (2006). The identification and assessment of intercultural competence as a student outcome of internationalization. Journal of Studies in International Education, 10, 241-266.

Deardorff, D. K. (2011). Assessing intercultural competence. New Directions for Institutional Research, 2011, 65-79.

Der-Karabetian, A. (1992). World-mindedness and the nuclear threat: A multinational study. Journal of Social Behavior & Personality.

Der-Karabetian, A., & Metzer, J. (1993). The cross-cultural world-mindedness scale and political party affiliation. Psychological reports, 72, 1069-1070.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. Journal of Personality and Social Psychology, 56, 5.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. Journal of Personality and Social Psychology, 82, 62.

Dunn, D. S., Brewer, C. L., Cautin, R. L., Gurung, R. A. R., Keith, K. D., McGregor, L. N., Nida, S. A., Puccio, P., & Voigt, M. J. (2008). The undergraduate psychology curriculum: Call for a core. In D. F. Halpern (Ed.). Undergraduate Education in Pschology: A Blueprint for the Future of the Discipline (pp. 47-61). Washington, DC: American Psychological Association.

Dweck, C.S. (2007). Mindset: The new psychology of success. New York: Ballantine Books.

Dweck, C. S. (2010). Even geniuses work hard. Educational Leadership, 68, 16-20.

Dweck, C. S. (2012). Mindsets and human nature: Promoting change in the Middle East, the schoolyard, the racial divide, and willpower. American Psychologist, 67, 614.

Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. Journal of Applied Psychology, 37, 90-93.

Eliason, M. J., & Raheim, S. (2000). Experiences and comfort with culturally diverse groups in undergraduate pre-nursing students. Journal of Nursing Education, 39, 161-165.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J., (1999). Evaluating the use of exploratory factor analysis in psychological researcher. Psychological Methods, 4, 272-299.

Finch, J. (1987). The vignette technique in survey research. Sociology, 21, 105-114.

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation models. In G.R. Hancock & R.O. Mueller (Eds.). A second course in structural equation modeling. Greenwich, CT: Information Age Publishing, Inc.

Giles, M. B., & Sherman, T. M. (1982). Measurement of multicultural attitudes of teacher trainees. The Journal of Educational Research, 75, 204-209.

Fantini, A. E. (2009). Assessing intercultural competence. The SAGE handbook of intercultural competence, 456-476.

Field, A. (2005). Discovering statistics with SPSS (2nd ed.). London: Sage.

Field, A. (2009). Discovering statistics using SPSS. (3rd ed.). London: Sage.

Global Awareness Profile, About (2015, August) Retrieved from https://globalawarenessprofile.wordpress.com/about/

Green, D., Callands, T. A., Radcliffe, A. M., Luebbe, A. M., & Klonoff, E. A. (2009). Clinical psychology students' perceptions of diversity training: A study of exposure and satisfaction. Journal of Clinical Psychology, 65, 1056-1070.

Gruenfeld, D.H., & Hollingshead, A.B. (1993). Sociocognition in work groups: The evolution of group integrative complexity and its relation to task performance. Small Group Research, 24, 383–405.

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. Harvard Educational Review, 72 330-336.

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). The Intercultural Development Inventory: A measure of intercultural sensitivity. International Journal of Intercultural Relations, 27, 421-443.

Hays, D. (2008). Assessing multicultural competence in counselor trainees: A review of instrumentation and future directions. Journal of Counseling & Development, 86, 95-101.

Hays, P. A. (1996). Addressing the complexities of culture and gender in counseling. Journal of Counseling & Development, 74, 332–338.

Hays, P. A. (2008). Addressing cultural complexities in practice: Assessment, diagnosis, and therapy (2nd edition). Washington, D.C.: American Psychological Association.

Hayward, F. M. & Siaya, L. M. (2001). Public experience, attitudes, and knowledge: A report on two national surveys about international education. Washington, DC: American Council on Education.

Henderson, S. J., Newman, R., Wedding, D., Chung, W. V., Day, T., Foster, S. L., Kelly, L., Mullen, P., Nickel, T., Ozyurt, S., Platt, J.J., Pinterits, E.J., Shoemate, A., & Webb, K. S. (2015). Conceptualization and initial testing of the Alliant Intercultural Competency Scale (AICS). International Psychological Bulletin, 30, 19-27.

Henderson, S. J., Horton, R. A., Saito, P. K., & Shorter-Gooden, K. (2014, July). Validation of assessment vignettes and scoring rubric of multicultural and international competency in faculty teaching. Multicultural Learning and Teaching. Published online ahead of print.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179-185.

Humes, K., Jones, N. A., & Ramirez, R. R. (2011). Overview of race and Hispanic origin, 2010. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Hunter, B., White, G. P., & Godbey, G. C. (2006). What does it mean to be globally competent? Journal of Studies in International education, 10, 267-285.

Jeffreys, M. R. (2000). Development and psychometric evaluation of the transcultural self-efficacy tool: A synthesis of findings. Journal of Transcultural Nursing, 11, 127-136.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis.

Educational and Psychological Measurement, 20, 141-151.

Kelley, C. & Meyers, J. (2007). Cross-Cultural Adaptability Inventory. Encyclopedia of Special Education.

Kline, P. (1999). The handbook of psychological testing (2nd ed.). London: Routledge.

Linstone, H. A., & Turoff, M. (Eds.). (1975). The Delphi method: Techniques and Applications. Reading, MA: Addison-Wesley.

Narvaez, D., & Hill, P.L. (2010). The relation of multicultural experiences to moral judgment and mindsets. Journal of Diversity in Higher Education, 3, 43-55.

Nelson Laird, T. F. (Fall 2014.) Reconsidering the inclusion of diversity in the curriculum. Diversity & Democracy, 17(4). Retrieved at http://www.aacu.org/diversitydemocracy/2014/fall/nelson-laird

Neville, H. A., Lilly, R. L., Duran, G., Lee, R. M., & Browne, L. (2000). Construction and initial validation of the Color-Blind Racial Attitudes Scale (CoBRAS). Journal of Counseling Psychology, 47, 59.

Maccoby, E. E., & Maccoby, N. (1954). The interview: A tool of social science. Handbook of Social Psychology, 1, 449-487.

Marchesani , L. S., & Adams, M. (1992). Dynamics of diversity in the teaching learning process: A faculty development model for analysis and action. In M. Adams (Ed.), Promoting diversity in college classrooms. New Directions for Teaching and Learning, 52, 9-20.

McIntosh, P. (1988). White privilege: Unpacking the invisible knapsack. In V. Cyrus (Ed.), Experiencing race, class, and gender in the United States (209–213). Mountain View, CA: Mayfield Publishing.

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.

Munroe, A., & Pearson, C. (2006). The Munroe Multicultural Attitude Scale Questionnaire: A New Instrument for Multicultural Studies. Educational and Psychological Measurement, 66, 819-834.

Olson, C. L., Evans, R., & Shoenberg, R. E. (2007). At Home in the World: Bridging the Gap. Washington D.C.: American Council on Education. Retrieved at https://bookstore.acenet.edu/products/home-world-bridging-gap-between-internationalization-and-multiculturaleducation

Paulhus, D. L. (1991). Measurement and control of response bias. In Robinson, J. P. & Shaver, P. R. (Eds.), Measure of personality and social psychological attitudes (17-59). San Diego, CA: Academic Press.

RAND. (2004). RAND study says demographic shifts, technological change and economic globalization will affect future U.S. workforce and workplace [Press Release].

Reynolds, W. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. Journal of Clinical Psychology, 38, 119-125.

Stevens, J. P. (2002). Applied multivariate statistics for the social sciences (4th ed.). Hillsdale, NJ: Erlbaum.

Suedfeld, P., Tetlock, P.E., & Streufert, S. (1992). Conceptual/integrative complexity. In C.P., Smith (Ed.), Motivation and personality: Handbook of thematic content analyses (pp. 393–400). New York: Cambridge University Press.

Wang, Y. W., Davidson, M. M., Yakushko, O. F., Savoy, H. B., Tan, J. A., & Bleier, J. K. (2003). The scale of ethnocultural empathy: Development, validation, and reliability. Journal of Counseling Psychology, 50, 221.

Wilks, T. (2004). The use of vignettes in qualitative research into social work values. Qualitative Social Work, 3, 78-87.

World Health Organization. (2015). Globalization. Retrieved November 16, 2015, from http://www.who.int/topics/globalization/en

Zalaquett, C. P., Foley, P. F., Tillotson, K., Dinsmore, J. A., & Hof, D. (Summer 2008). Multicultural and social justice training for counselor education programs and colleges of education: Rewards and challenges. Journal of Counseling & Development, 86, 323-329.