

# The Global Engagement Measurement Scale (GEMS): A New Scale for Assessing the Impact of Education Abroad and Campus Internationalization

**Noel L. Shadowen**

*The University of Delaware, The Institute for Global Studies*

**Lisa P. Chieffo**

*The University of Delaware, The Institute for Global Studies*

**Nancy G. Guerra**

*The University of Delaware, The Institute for Global Studies*

The growing popularity and push for engagement in education abroad programs in the U.S. and globally, coupled with the need to demonstrate tangible benefits, has increased focus on how best to assess impact beyond academic outcomes. Although course grades and learning are important outcomes, education abroad is intended to be more than just coursework in other countries. Indeed, the rationale is that the experience of studying and living in another culture provides important opportunities for learning and development above and beyond the acquisition of academic knowledge at the home campus. Students have the opportunity to study language with native speakers, participate in cultural events, and meet or live with peers and families who have different norms, perspectives, and expectations. Students often experience new routines, confront difficult situations requiring problem solving skills, and cope with psychological stressors as they adjust to new and unfamiliar contexts. These experiences may be brief, or they may span a semester or academic year.

Given these diverse experiences, an important step for assessing impact is to examine carefully how the unique features of education abroad can influence students' development. Although scholars in the field of international education have acknowledged the value of this non-academic learning, as of yet there has been no consensus on the most relevant constructs (Deardorff 2004, 2006; Fantini, 2000; Wastson, Siska, & Wolfel, 2013). A number of different outcomes have been considered, although in recent years there has been a growing focus on the construct of "intercultural competence." Still, this term is quite broad, and there is little agreement on the core elements or how to assess them.

Attempts to disentangle these core elements have resulted in an array of definitions. Some are extremely broad; for example, Ashwill and Hoang Oanh (2009, p. 143) define intercultural competence as "a multifaceted state of being that gives one the ability to adapt to different cultural settings." Others are more specific. In a Delphi study, Deardorff (2006) highlighted two definitions: (1) "knowledge of others; knowledge of self; skills to interpret and relate; skills to discover and/or to interact; valuing others' values, beliefs, and behaviors; and revitalizing oneself" (Byram, 1997, p. 34), and (2) "the ability to communicate effectively and appropriately in intercultural situations based upon one's intercultural knowledge, skills, and attitudes" (Deardorff, 2004, p. 194). In some cases, terms such as "intercultural sensitivity," "cross-cultural competence," and "intercultural communication competence" have been used with similar connotations, making it even more difficult to articulate a set of common constructs and reliable indicators for measurement.

One challenge is to articulate carefully the important and likely non-academic outcomes among a wide array of education abroad programs, each with its own individual objectives. Another challenge is to access valid and reliable measures to examine these outcomes. On the one hand, many universities have developed "home grown" measures for their own use, but these generally are not theoretically based or psychometrically validated. On the other hand, cottage industries sell related measures (with adequate psychometric data) for a fee, but these typically are quite long, require training and/or certification, and are not specific to education abroad outcomes, nor to the specific goals and needs of particular institutions. For example, the Beliefs, Events, and Values Inventory (BEVI; Shealy, 2005) is a 300-item survey developed for use in cross-cultural settings. It includes 18 separate scales that measure constructs such as sociocultural openness, self-awareness, and gender traditionalism. In order to administer the BEVI, universities must complete a one-day certification process and pay an annual usage fee. A number of similar measures are also available, such as the Intercultural Development Inventory (IDI; Hammer, 2010) and the Global Perspectives Inventory (GPI; Braskamp, Braskamp, & Engberg, 2013).

The purpose of the present study was to develop and validate a brief education abroad assessment that focuses on key constructs, is easy to administer, and is publically available. We first identified a small set of constructs that were specific, measurable, and likely to change as a result of short- or long-term education abroad programs. Given that intercultural competence has gained considerable traction as an area of inquiry in the field, we wanted to assess related constructs, but consider other options as well. Towards this end, we conducted informal focus groups with the University of Delaware's Study Abroad office and other international educators on campus. We also reviewed literature from related disciplines. Although our focus was on the evaluation of education abroad, we aimed to develop a scale that could be used broadly to assess the impact of a range of campus internationalization efforts. Ultimately, we focused on four specific constructs: cultural engagement, knowledge of the host site, ambiguity tolerance, and diversity openness.

### **Cultural Engagement**

Cultural engagement (or cross-cultural awareness, openness, or sensitivity, as it is variously described in the literature) features prominently in many theories of intercultural competence (Deardorff, 2006). As typically defined, it captures students' worldview or attitudes toward cultural differences, diversity, and exchanges. It includes concepts such as global-mindedness, defined as a worldview that is future-oriented and extends beyond national borders (Hett, 1993). The notion is that increased exposure to different cultures promotes increased understanding not only of new cultures but also of the interconnectedness between various peoples and cultures.

### **Tolerance for Ambiguity**

The construct of tolerance for ambiguity measures students' ability to feel comfortable with and competent in new or novel situations, or when handling novel stimuli or information. It has been defined as "an orientation, ranging from aversion to attraction, toward stimuli that are complex, unfamiliar, and insoluble" (McLain, 2009, p. 977). Education abroad students often are in situations in which they must navigate unfamiliar or confusing circumstances. Indeed, preliminary evidence shows that ambiguity tolerance can increase during the course of international education experiences, perhaps due to the necessity of "stretching" beyond one's comfort level (Dewaele & Wei, 2013).

## Knowledge of the Host Site

Knowledge of the host site is a relatively straightforward assessment of students' self-reported knowledge of their host community. This can include the host culture's norms, governance structure, history, and current issues. This type of functional knowledge of a host community is an important outcome of education abroad, demonstrating that students took the time and effort to acquire in-depth knowledge of their program's context (Chieffo & Griffiths, 2004; Williams, 2009).

## Diversity Openness

Pascarella, Edison, Nora, Hagedorn, and Terenzini (1996, p. 175) defined openness to diversity as an "orientation toward enjoyment from being intellectually challenged by different ideas, values, and perspectives as well as an appreciation of racial, cultural, ethnic diversity." Individuals who tolerate and embrace diversity are able to take new perspectives and work effectively across diverse collaborative groups and settings (Bakalis & Joiner, 2004). The importance of integrating diversity measures in education abroad outcomes has been well acknowledged, although there is little consistency in measurement strategies and findings (Black & Duhon, 2006; Clarke, Flaherty, Wright, & McMillen, 2009; Zhai, 2000). One particular challenge of diversity tolerance measurement is potential social desirability response effects.

## Study Aims and Hypotheses

The aims of the present study were to: 1) examine the psychometric properties of each construct of interest, including the underlying factor structure; 2) perform preliminary validity checks on the new factors and scales; and 3) examine pre-post change in each of these constructs after a five-week education abroad program.

We hypothesized that the underlying factor structure of the Ambiguity Tolerance and Knowledge of the Host Site scales would best fit a one-factor model, the most parsimonious solution. We hypothesized a three-factor higher order model for the Cultural Engagement scale. We also expected to see a two-factor correlated model best fit the Diversity Openness scale, as it was designed to measure attitudes in two different situations. We also hypothesized that all scales would demonstrate preliminary convergent and divergent validity (Kazdin, 1995). Specifically, we expected to see positive correlations between theoretically correlated constructs such as diversity openness and cultural engagement. We did not expect there to be significant correlations between theoretically unrelated constructs such as ambiguity tolerance and knowledge of the host site. Finally, we hypothesized that students' scores on all of the constructs of interest would increase from pre- to post- education abroad experience.

## Method

### Participants and Procedures

Participants were recruited from all students who participated in the University of Delaware's 47 short-term, faculty-led education abroad programs during January 2014 ( $N=1,017$ ). Each student received a survey request via Qualtrics from the Study Abroad Office requesting voluntary participation. A total of 454 students participated (45% of eligible students). The majority of the students were college juniors (45.6%), but sophomores, seniors, freshman, and graduate students also participated. Females constituted 75% of the sample. Most of the students (96.9%) were U.S. citizens and the majority self-identified as non-Hispanic White (93.2%). Students who completed the pre-

study abroad survey were invited to complete the survey again after their education abroad experience in February 2014. Of the 454 participants who completed the pre-survey, 263 (56%) participated in the follow-up survey. Demographic data was similar to the pre-test sample. All research procedures were approved by the University of Delaware's Institutional Review Board.

## Measures

An initial Global Engagement Measurement Scale (GEMS) was developed to be pilot tested with the participants detailed above. Reliability estimates are reported below in the results section, as part of the psychometric analyses. The final measure is presented in Appendix A.

### *Cultural engagement.*

We reviewed existing instruments that assess constructs similar to cultural engagement, particularly the Hett Scale of Global Mindedness (Hett, 1993). Hett (1993) characterized global-mindedness as a worldview that is future-oriented and extends beyond national borders. The piloted cultural engagement scale contained 16 items focusing on students' feelings of engagement with other cultures, pluralistic points of view, and ethno-relative worldviews. Students rated items on a four-point Likert scale from 0 (*strongly disagree*) to 3 (*strongly agree*). Some items reflected ethno-centric points of view and were reverse coded. Higher scores on this measure reflect higher values of the construct. A sample cultural engagement question is "I like trying to understand people's behaviors in the context of their own culture."

### *Ambiguity tolerance.*

We reviewed several measures related to ambiguity tolerance. We adapted several items from McLain's (2009) Multiple Stimulus Types Ambiguity Tolerance Scale-II (MSTAT-II), an instrument based on psychological theory. We also developed two new items that related ambiguity tolerance to situations relevant to education abroad. These were "I enjoy exploring new places where I am unfamiliar with the geography or with the people" and "I generally prefer the beginning of a project when there are many potential paths and outcomes rather than the end of a project when the outcome is clearer." Participants rated items on a four-point Likert scale from 0 (*Never*) to 3 (*Always*). There were a total of 14 items on the initial scale.

### *Knowledge of the host site.*

Our seven-item knowledge of the host site scale focused on the extent to which students hold cultural, social, or political knowledge of their host site. Participants rated items on a four-point scale from 0 (*Strongly disagree*) to 3 (*Strongly agree*). A sample item is "I have sufficient knowledge of my host site to explain a current issue there to a friend or family member who has never been there."

### *Openness to diversity.*

The openness to diversity scale measured the extent to which students felt comfortable interacting with different people. We created the questionnaire items based on a literature review of this construct (Pascarella, Edison, Nora, Hagedorn, & Terenzini, 1996). The openness to diversity measure consists of two sets of questions. Both sets were rated on a four-point scale from 0 (*No*) to 3 (*Definitely*). The first set of eight items asked students to think about diversity in terms of if they could *become good friends* with a person distinctly different from themselves (for example, "if the other person held different religious beliefs"). The second set of eight items asked respondents if certain characteristics would negatively impact their *judgment of a person who was going to marry their best friend or sibling* (for

example, “if the other person was from a different cultural background”). By asking respondents to imagine actual real-world scenarios, we hoped to realistically tap into students’ potential biases and avoid the demand characteristics of typical diversity tolerance scales (Ganster, Hennessey, & Luthans, 1983).

## Results

### Aim 1: Underlying Factor Structure

Analyses were organized around study aims. First, we used Confirmatory Factor Analysis (CFAs) to examine the underlying factor structure of the GEMS instrument subscales in Mplus version 7 (Muthén & Muthén, 2011). The maximum likelihood estimator with robust standard errors (MLR) estimator was used, as it is robust to potential non-normality. We used CFAs instead of Exploratory Factor Analyses (EFAs) because we had a priori hypothesized factor solutions.

We examined standard parameters to assess CFA model fit and to compare models (Raykov & Marcoulides, 2008). These fit indices and approximate cut-off values included: (1) small chi square values, (2) a Root Mean Square Error of Approximation (RMSEA) close to .06 or below, (3) a Standardized Root Mean Square Residual (SRMR) close to .08 or below, and (4) Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values approaching .95 or greater (Browne & Cudeck, 1993; Hu & Bentler, 1999; Kenny & McCoach, 2003).

Regarding individual items, we examined factor loadings to assess the fit of items. Individual items were removed if they did not load above .40 (Tabachnick & Fidell, 2007). Before beginning factor analyses, we also examined data from the piloted scales to determine if items should be deleted due to poor reliability, poor loadings, or poor feedback from students. One-factor models were tested to determine factor loadings.

#### *Cultural engagement.*

We first tested the model fit of a parsimonious, one-factor model with all items loading on to a single factor. This model did not meet the relevant criteria for goodness-of-fit. Table 1 details all fit indices. Next a two-factor model and a three-factor model were tested, based on prior literature. Although the inclusion of three-factors seemed to improve the fit of the model, neither model’s fit indices met the relevant fit criteria. Finally, we tested our hypothesized model, a three-factor nested model that included “cultural engagement” as a higher-order construct. Based on the standards for fit indices, this hypothesized higher-order model seems to best fit the data (Table 1).

The higher-order three-factor model was validated and makes sense theoretically, as all three factors load on to the larger construct of “cultural engagement”. The three factors were “pluralism” (four items,  $\alpha = .85$ ) “interconnectedness” (three items,  $\alpha = .85$ ), and “global mindedness” (ethnocentrism, reverse coded, five items  $\alpha = .71$ ). “Pluralism” loaded on to the higher-order “cultural engagement” construct highest, at .93. “Interconnectedness” loaded at .61, and the loading for “global mindedness” was .24. Cronbach’s alpha for this total scale was  $\alpha = .77$ , indicating adequate internal reliability.

#### *Ambiguity tolerance.*

For ambiguity tolerance, we first tested the hypothesized one-factor model; fit indices for this model did not meet the standards for a well-fitted model (Table 1). We also tested an alternative two-

factor model which, based on psychometric standards, was a better fit for the data. The two factors were named “flexibility” (three items,  $\alpha = .71$ ) and “preference for new situations” (four items,  $\alpha = .73$ ). The factors were correlated to each other at .84. Internal reliability for the total ambiguity tolerance scale was high at  $\alpha = .84$ .

#### *Knowledge of the host site.*

The knowledge of the host site scale was tested with the hypothesized one-factor model. The fit indices for this model indicated that psychometric standards were met ( $\chi^2 = 11.18$ ,  $df = 5$ ,  $CFI = .99$ ,  $TLI = .99$ ,  $SRMR = .01$ ,  $RMSEA = .05$ ) and the hypothesized model was a good fit for the data. Alpha reliability for the Knowledge of Host Site scale was  $\alpha = .87$ , indicating high reliability.

#### *Diversity Openness.*

A one-factor parsimonious model was first tested with the Diversity Openness scale. The fit indices for the model (Table 1) did not meet standards for a well-fitting model. Additionally, we tested the two factors as two individual uncorrelated one-factor models. Next, we tested the hypothesized two-factor correlated model. The two factors were correlated at .21. The alpha reliability of the two-factor correlated model is high at  $\alpha = .90$ . The model fit indices for the two-factor correlated model indicate that there is adequate model fit, although the RMSEA index is higher than would be expected. Although it was not possible to directly compare the two-factor correlated model with the uncorrelated individual scales models, it appears that the model fit indices are similar. We chose the two-factor correlated model as our final model, as it is the more theoretically sound model. See Table 1 and Table 2 for factor loadings and model comparisons.

### **Aim 2: Preliminary Validity Checks**

Descriptive statistics (Table 3) were calculated on the initial sample of students before they went on their education abroad program ( $n = 454$ ), according to scale construction based on the CFAs presented above.

Next, we computed bivariate correlations for the subscales' individual factors to determine preliminary convergent and divergent validity (Table 4). The constructs were correlated in the expected direction and provide preliminary evidence for scale validity. Sub-factors which would be expected to correlate because of theoretical similarity, such as interconnectedness and preference for new situations, were in fact correlated (at  $r = .40$   $p < .001$ ). Sub-factors that were not theoretically related, such as ethnocentrism and flexibility, were not significantly correlated. The Pearson correlations provide preliminary divergent and convergent validity for the constructs and sub-factors.

### **Aim 3: Pre-to-Post Scores**

Finally, we examined pre- to post-education abroad change on the four scales using paired-samples *t*-tests (Table 5). These analyses were conducted using IBM SPSS 20 software. For each *t*-test, we also calculated an estimate of effect size (ES) using a Cohen's *d* equation adjusted for the correlated nature of the paired sample (Dunlap, Cortina, Vaslow, & Burke, 1996). ES estimates were interpreted according to Cohen's (1988) standards, with 0.20-0.49 being a small effect, 0.50- 0.79 a medium effect, and 0.80 and above a large effect. We expected that scores would increase following participation in education abroad.

Cultural engagement scores increased significantly from pre- to post- experience, with a large ES

(0.52; Cohen, 1988). Similarly, the scores for ambiguity tolerance increased significantly from the pre- to post-test, with a medium ES (0.40). For knowledge of the host site, scores increased significantly from pre- to post-test, and had a large ES (2.10). Finally, for diversity openness there was a significant, but quite small, increase from pre- to post-test, ES = 0.14.

## Discussion

The first aim of this study was to examine the psychometric properties of the GEMS. The results indicate that in line with our hypothesis, the Cultural Engagement scale best fits a higher order three-factor model, with “cultural engagement” the higher order factor. The lower-order factors of global-mindedness, interconnectedness, and pluralism fit with the conceptualization of cultural engagement. These seem to represent a set of attitudes and worldviews that taken together reflect a person’s level of engagement. The Ambiguity Tolerance scale best fits a two-factor correlated model, with two highly correlated factors of flexibility and preference for new situations. This interesting finding suggests that ambiguity tolerance is comprised of two distinct sub-skills. In line with our hypothesis, we found that the Knowledge of the Host Site scale best fits a one-factor model, with all items loading highly. The Diversity Openness scale was determined to best fit a two-factor correlated model. The two-factors and a low correlation (.21) between the factors demonstrate that instead of an overarching concept of “openness”, people may have different attitudes or “openness” that depends on the particular situation (for example classmates versus friends). All of the scales demonstrated adequate to strong internal reliability and factor loadings.

Our second aim was to examine preliminary evidence for validity of the new scales and factors. The results support our hypothesis that constructs that are similar to one another and should be positively correlated do in fact exhibit significant positive correlations, such as global mindedness and diversity openness. Additionally, the constructs not expected to significantly correlate do exhibit non-significant correlations.

The third aim was to conduct a preliminary analysis of pre- to post-education abroad scores on the constructs of interest. As hypothesized, the results supported a post-education abroad increase in scores on all scales. The largest ES was for the knowledge of the host site construct. The diversity openness construct had the smallest sensitivity to change. These differences in ES indicate that objective knowledge may be the quickest or easiest construct to affect, while deeper attitudes or worldviews, such as diversity openness, may require stronger intervention.

Thus, the GEMS instrument has a robust factor structure, reliable scales, and appears to be sensitive to changes in targeted non-academic outcomes following an education abroad program. We were able to delete 17 items, reducing the total item burden to 37 total questions. This GEMS instrument could be used for a range of education abroad programs, including short-term and long-term experiences. This instrument could also be used to measure the impact of campus internationalization efforts for international and domestic students. The instrument measures theoretically and practically important constructs for the development of culturally competent and skilled higher education students, has been preliminarily validated, is easy to administer, and is available free of charge.

A limitation of the GEMS is that it presents a culture-specific notion of the essential skills of

intercultural competence. Further testing of this measure with students from different cultures would be useful. This instrument also represents only some of the many constructs related to intercultural competence. Other constructs such as intercultural communication competence, language proficiency, and personal identity were not included. Additionally, this study did not include a control group comparison, and the GEMS could be more or less sensitive to change in a controlled trial. Future work will be required to conduct measurement invariance tests for factor structure according to time and demographic variables.

According to Deardorff's (2006) study, intercultural competence is best assessed through a mix of quantitative and qualitative measures. Questionnaire data represent one mechanism to quantify outcomes. Qualitative assessments such as student journals, narrative diaries, or portfolios are another method. These assessments would provide detailed and important information about how students learn and incorporate intercultural competence, and can enrich quantitative findings. However, questionnaires administered pre- and post- experience also provide valuable information on how students change and what types of education abroad experience contribute most to this change. For institutions with large numbers of participants on a broad array of programs with diverse objectives, ongoing qualitative assessment may simply not be practical due to the volume of data and the time required for analysis.

As universities expand their education abroad and general campus internationalization efforts, rigorous assessment continues to gain importance and attract the attention of campus policymakers. It is therefore essential to document the impact that these experiences can have on student learning and development. Ideally, assessment should be ongoing, sustainable, and cyclical in order to lead to programmatic changes and improvement over time. For those who find that the constructs measured by the GEMS align with their own campus objectives, there is the potential for other institutions to use the GEMS, compare results, and broaden the growing base of knowledge in the field of education abroad research.

## References

- Ashwill, M. A., & Hoang Oanh, D. T. (2009). Developing Globally Competent Citizens: The Contrasting Cases of the United States and Vietnam. In D.K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 141-157). Thousand Oaks, CA: Sage Publications.
- Bakalis, S., & Joiner, T. A. (2004). Participation in tertiary study abroad programs: the role of personality. *International Journal of Educational Management*, 18 (5), 286-291.
- Black, T. H. & Duhon, D. L. (2006). Assessing the impact of business study abroad programs on cultural awareness and personal development. *Journal of Education for Business*, 81 (3), 140-144. doi: 10.3200/JOEB.81.3.140-144
- Braskamp, L. A., Braskamp, D. C., & Engberg, M. E. (2013). *Global Perspective Inventory (GPI) Manual*. Chicago, IL: Global Perspectives Institute Inc.
- Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In: K.A. Bollen & J.S. Long (Eds.) *Testing structural equation models*. (pp.136-162). Beverly Hills, CA: Sage.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Philadelphia, PA: Multilingual Matters.
- Chieffo, L. & Griffiths, L. (2004) Large-scale assessment of student attitudes after a short-term study abroad program. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10, 165-178.
- Clarke, I., Flaherty, T. B., Wright, N. D., & McMillen, R. M. (2009). Student international proficiency



- from study abroad programs. *Journal of Marketing Education*, 31 (2), 173-181. doi: 10.1177/0273475309335583
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2<sup>nd</sup> ed. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Deardorff, D. K. (2004). *The identification and assessment of as a student outcome of internationalization at institutions of higher education in the United States*. Unpublished doctoral dissertation. North Carolina State University, Raleigh, NC.  
<http://repository.lib.ncsu.edu/ir/bitstream/1840.16/5733/1/etd.pdf>
- Deardorff, D. K. (2006) Identification and Assessment of Intercultural Competence as a Student Outcome of Internationalization. *Journal of Studies in International Education*, 10(3), 241–266.
- Deardorff, D. K. (2009). Implementing Intercultural Competence Assessment. In D. K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 477-491). Thousand Oaks, CA: Sage Publications.
- Dewaele, J. & Wei, L. (2013). Is multilingualism linked to a higher tolerance for ambiguity? *Bilingualism: Language and Cognition*, 16, 231-240. Doi:10.1017/S1366728912000570
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1 (2), 170-177.
- Fantini, A. E. (2000). A central concern : Developing intercultural competence. *SIT Occasional Paper Series*, 1. Retrieved April 1, 2014, from: <http://www.sit.edu/publications/docs.pdf>.
- Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social desirability response effects: Three alternative models. *Academy of Management Journal*, 26 (2), 321-331.
- Hammer, M. R. (2010). *The Intercultural Development Inventory manual*. Berlin, MD: IDI.
- Hett, E. J. (1993). *The development of an instrument to measure globalmindedness*. Unpublished Doctoral Dissertation, University of San Diego.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6 (1), 1-55. Doi: 10.1080/10705519909540118.
- Institute of International Education. (2013). Open Doors Report on International Educational Exchange. Retrieved from <http://www.iie.org/opendoors>.
- Kazdin, A. E. (1995). Preparing and evaluating research reports. *Psychological Assessment*, 7 (3), 228-237.
- Kenny, D. A. & McCoach, D. B. (2003). Effect of the number of variables on measures of fit indices in structural equation modeling. *Structural Equation Modeling*, 10, 333-351.
- Li, Y. (2013). Cultivating student global competence: A pilot experimental study. *Decision Sciences Journal of Innovative Education*, 11, 125-143.
- McLain, D. L. (1993). The MSTAT-I: A new measure of an individual's tolerance for ambiguity. *Educational and Psychological Measurement*, 53 (1), 183-189.
- McLain, D. L. (2009). Evidence of the properties of an ambiguity tolerance measure: The Multiple stimulus types ambiguity tolerance scale-II. *Psychological Reports*, 105(3), 975-988.
- Muthén, L. K. & Muthén, B. O. (1998-2011). *Mplus User's Guide*. 6<sup>th</sup> ed. Los Angeles, CA: Muthén & Muthén.
- Pascarella, E. T., Edison, M., Nora, A., Hagedorn, L. S., & Terenzini, P. T. (1996). Influences on students' openness to diversity and challenges in the first year of college. *Journal of Higher Education*, 67 (2), 179-195.
- Raykov, T. & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. New York, NY: Routledge.
- Shealy, C. N. (2005). Justifying the justification hypothesis: scientific-humanism, Equilintigration (EI) theory, and the Beliefs, Events, and Values Inventory (BEVI). *Journal of Clinical Psychology*, 61 (1), 81-106.

- Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics*. 5<sup>th</sup> ed. Boston, MA: Allyn & Bacon.
- Wastson, J. R., Siska, P., & Wolfel, R. L. (2013). Assessing gains in language proficiency, cross-cultural competence, and regional awareness during study abroad: A preliminary study. *Foreign Language Annals*, 46 (1), 62-79. Doi: 10.1111/flan.12016
- Williams, T. R. (2009) The reflective model of Intercultural Competency: a multidimensional, qualitative approach to study abroad assessment. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 18, 289-306.
- Zhai, L. (2000). *The influence of study abroad programs on college student development in the College of Food, Agriculture, and Environmental Science at the Ohio State University*. Unpublished doctoral dissertation. Ohio State University, Columbus.

## Appendices

**Table 1. Model fit indices for constructs**

Construct	Specified Model	$\chi^2$	<i>df</i>	CFI	TLI	SRM R	RMSEA
Cultural Engagement	1-factor	729.64	54	.61	.53	.13	.18
	2-factor	432.08	53	.78	.73	.08	.13
	3-factor	226.30	51	.90	.87	.07	.09
	3-factor nested	120.92	51	.96	.95	.05	.06
Ambiguity Tolerance	1-factor	69.91	14	.93	.89	.04	.10
	2-factor	45.46	13	.96	.94	.03	.08
Host Site Knowledge	1-factor	11.18	5	.99	.99	0.01	0.05
Diversity Openness	1-factor	2429.79	65	.49	.38	.26	.30
	2-factor corr.	455.05	64	.92	.90	.04	.12
	2-factor uncorr. (Friends)	86.65	14	.97	.95	.03	.11
	2-factor uncorr. (Siblings)	99.96	9	.96	.93	.03	.16

**Table 2. Factor loadings for scales**

Scale	Item Num.	Factor Loading	1 Factor Loading	2 Factor Loading	3 Factor Loading	Factor Correlations
Cultural Engagement	1	.85				
	2	.76				
	3	.79				
	4	.69				
	5		.76			
	6		.80			
	7		.86			
	8				.52	
	9				.67	
	10				.69	
	11				.51	
	12				.49	
						Cultural Engage by: Factor 1: .93 Factor 2: .61 Factor 3: .24
Tolerance for Ambiguity	1	.69				
	2	.59				
	3	.76				
	4		.67			
	5		.45			
	6		.69			
	7		.75			
						F1 WITH F2: .84
Knowledge of Host Site	1	.81				
	2	.69				
	3	.81				
	4	.90				
	5	.65				

Scale	Item Num.	Factor Loading	1	Factor Loading	2	Factor Loading	3	Factor Correlations
Diversity Openness	1	.63						
	2	.81						
	3	.78						
	4	.75						
	5	.91						
	6	.92						
	7	.87						
	8				.80			
	9				.88			
	10				.79			
	11				.71			
	12				.94			
	13				.92			
								F1 with F2: .21

**Table 3. Descriptive statistics for constructs**

Construct	Mean	SD	Range	$\alpha$	Factors
Cultural Engagement	37.18	4.42	19-48	.77	4
Knowledge of Host Site	10.85	4.11	5-20	.87	1
Ambiguity Tolerance	19.54	3.45	8-28	.84	2
Diversity Openness	46.69	6.33	23-52	.90	2

**Table 4. Bivariate correlations**

Variables	1	2	3	4	5	6	7
1. Cultural Engagement (Pluralism)	--	.51***	.19***	.24***	.26***	.23***	.11*
2. Cultural Engagement (Interconnectedness)		--	.12*	.33***	.40***	.14***	.30***
3. Cultural Engagement (Global Mindedness)			--	.08	.08	.29***	-.01*
4. Ambiguity Tolerance (Flexibility)				--	.62***	.10*	.28***
5. Ambiguity Tolerance (New situation preference)					--	.12*	.29***
6. Diversity Openness						--	-.11*
7. Host Site Knowledge							--

Note. \* $p < .05$ , \*\* $p < .01$  \*\*\* $p < .001$

**Table 5. Paired Samples *t*-tests with effect sizes**

	Mean difference	<i>SD</i>	<i>t</i> -statistic	<i>df</i>	Cohen's <i>D ES</i>
Cultural Engagement	1.88	3.78	7.53***	228	0.52
Knowledge of Host Site	7.28	4.37	25.63***	235	2.10
Ambiguity Tolerance	1.45	3.41	6.42***	228	0.40
Diversity Openness	1.04	7.30	2.13*	224	0.14

*Note.* \* $p < .05$ , \*\* $p < .01$  \*\*\* $p < .001$

## Appendix A. University of Delaware's Global Engagement Measurement Scale

Please mark the response that most closely aligns with your feelings about the following statements:

Strongly Disagree                      Disagree                      Agree                      Strongly Agree

1. It is interesting to spend time talking with people from other cultures.
2. People in my home country are entitled to the standard of living they can afford, even if it has a small negative impact on the environment.
3. In addition to being a citizen of my own country, I think of myself as a global citizen.
4. It is important that universities promote understanding among students of varying backgrounds.
5. I support policies that maintain the present system of distribution of the world's wealth and resources.
6. The needs of my home country and its citizens should be the highest priority when my country's leaders negotiate with other countries.
7. I like trying to understand people's behaviors in the context of their own culture.
8. I feel a strong connection with humanity worldwide.
9. People in my home country can learn from people in other parts of the world.
10. I identify as a member of the worldwide community.
11. My home country's values are most likely the best in the world.
12. I feel irritated when people from other countries do not understand how things are done in my home country.

Please check the response that you think most closely aligns with your personality.

Note: "Ambiguity" refers to something that is open to more than one interpretation, something that is uncertain, or something that is not clearly defined.

Never                      Sometimes                      Frequently                      Always

13. I am comfortable with ambiguous situations.
14. I prefer situations where there is some uncertainty about potential outcomes or solutions.
15. I enjoy solving problems that must be viewed from multiple perspectives.
16. I feel that I can handle ambiguous situations.
17. I prefer to have new experiences rather than familiar or routine experiences.
18. I feel competent navigating ambiguous situations.
19. I enjoy exploring new places where I am unfamiliar with the geography or the people.

Please answer the following questions regarding your knowledge of your program site:

No                      Maybe                      Probably                      Definitely

20. I have sufficient knowledge of my host site to explain a current issue there to a friend or family member who has never been there.
21. I can discuss with confidence at least two historic events that are important to the population of my host site.



22. I can discuss with confidence the system of government and politics in my host site.  
 23. I have sufficient knowledge of my host site to have a discussion about a current social or economic issue of importance to the area.  
 24. I can explain with confidence what the community is like at my host site.

Imagine you just met a new person who is a potential friend. Could you become good friends with this person if you discovered that the person...

No                      Maybe                      Probably                      Definitely

25. ... holds different political beliefs and opinions than you?  
 26. ... is from a different socioeconomic background than you?  
 27. ... has different religious beliefs than you?  
 28. ... has a different sexual orientation than you?  
 29. ... is from a different cultural background than you?  
 30. ... is from a different racial or ethnic background than you?  
 31. ... is from a different country than you?

Imagine that your sibling or best friend is considering marrying someone. You meet this person and discover that he or she has one of the following characteristics. Would it *negatively* impact your judgment of this person if he or she:

No                      Maybe                      Probably                      Definitely

32. ... holds different political beliefs and opinions than you?  
 33. ... is from a different socioeconomic background than you?  
 34. ... has different religious beliefs than you?  
 35. ... is from a different cultural background than you?  
 36. ... is from a different racial or ethnic background than you?  
 37. ... is from a different country than you?