Jones, D. C., Campbell, M. M., & Acheson, K. (2019). Intercultural learning in semester-long study abroad: A comparative analysis of the effectiveness of one-on-one versus group-mentored interventions. *Journal on Excellence in College Teaching, 30*(4), 7-35.

Intercultural Learning in Semester-Long Study Abroad: A Comparative Analysis of the Effectiveness of One-on-One Versus Group-Mentored Interventions

Daniel C. Jones
Michelle M. Campbell
Kris Acheson
Purdue University

Evidence suggests that facilitated interventions in study abroad enhance gains in intercultural development, but many recent studies are not generalizable due to small numbers and lack of experimental design. In this study, subjects received individual or small-group interventions online in a course with trained intercultural mentors. The authors analyzed pretest-posttest changes on an intercultural assessment, the Intercultural Development Inventory, for treatment groups and a control for between-group differences. This study corroborates previous findings on the effectiveness of mentorship in study abroad and offers innovations for scaling up institutional efforts to support intercultural learning to reach larger numbers of students.

Contemporary research has countered the long-held assumption that intercultural competence development occurs naturally, without intentional support, as a result of participation in a study abroad experience. For example, Pederson (2010) found that during a year-long study abroad, the intervention group demonstrated more growth than non-intervention groups. In a review of literature on immersion versus intervention, Lou, Vande Berg, and Paige (2012) found little evidence in the form of assess-

ment outcomes to support an immersion hypothesis for intercultural learning; rather, they argue that intentional support, also called intercultural intervention, is critical during study abroad programs to promulgate student intercultural gains.

Evidence suggests that facilitated intervention during an undergraduate study abroad program provides greater gains in intercultural development when compared with no intervention. Not only did Lou and Bosley (2012) conclude that intercultural intervention in cultural immersion was required for meaningful gains, they also compared facilitated intervention and unfacilitated intervention in their Intentional, Targeted Intervention (ITI) model. They found that facilitated intervention led to significantly greater gains for undergraduate study abroad students as compared to unfacilitated (that is, student self-guided) intervention. Paige, Harvey, and McCleary (2012) also emphasized the need for facilitated intervention to be intensively responsive to students in an effort to "[cultivate] critical thinking, curiosity, and reflection; challenging and supporting students to advance their learning; and helping to ensure the well-being of the students in their host environment" (p. 332).

While the disciplines of continuous learning or adult education sometimes discuss the differences between individual and group teaching models (see Sessa & London, 2015), research comparing individual to group interventions is most often conducted by clinicians in the psychological sciences, for example, studies on smoking cessation (Stead, Carroll, & Lancaster, 2017) and interventions for antisocial behaviors (Lochman et al., 2015). Such research is rarely done in higher education, likely because the norm for formal intervention delivery tends to be group (class) based, with individual mentoring occurring more rarely (for example, research internships) and often informally.

Previous studies on adding intercultural mentoring to a study abroad have yielded encouraging results but typically have small numbers of around 30-40 subjects (Krishnan, Masters, Holgate, Wang, & Calahan, 2017; Williams, 2017). This study, in contrast, examines the efficacy of both individually and group-mentored interventions over the period of a semester-long study abroad using a relatively large number of subjects compared to similar research projects ($N = 376$). The subjects chosen for this study were selected according to their program length (academic semester or full academic year programs) and voluntary election to participate in structured intercultural mentoring courses.

The primary interest in this study was to determine how effective each of the structured intercultural mentoring courses were in comparison to semester abroad experiences with no intervention. Secondarily, the extent

to which each treatment was effective in comparison to one another was examined. Or, to state the research question more plainly: What, if any, meaningful differences in intercultural growth existed between individualized mentoring, group mentoring, and no mentoring? Additionally, we also examined whether correlations could be found when reviewing intercultural growth along lines of gender and the chosen field of study. Each of these questions is addressed using results from the Intercultural Development Inventory, or IDI (IDI, LLC, 2018). We are aware of no other published study that addresses these questions.

Context of the Study

Purdue University is a land-grant, research institution with an undergraduate enrollment of approximately 30,000, 21% of whom are international students (Purdue University Data Digest, 2019). More than 75% of students are enrolled in a STEM program of study (Purdue University Data Digest, 2019). During the academic year 2016-17, of the 2,044 undergraduate students who participated in study abroad, 418 took part in a semester- or year-long program, according to Purdue University's Study Abroad Office, with the remainder joining faculty-led short-term programs. In 2017-18, 2,289 undergraduate students studied abroad, with 396 abroad for a semester or longer.

In 2012, Purdue University's Faculty Senate adopted various American Association of Colleges and Universities' (AAC&U) Valid Assessment of Learning in Undergraduate Education (VALUE) Rubrics (see Finley & Rhodes, 2013). These include the Intercultural Knowledge and Competence Rubric, as embedded student learning outcomes in the university Core Curriculum (Purdue University, 2012).[1] In response, the Dean of International Programs prioritized shifting the study abroad paradigm to one centered on intercultural learning. New programs were created to train the faculty and staff leading short-term study abroad programs, and online mentorship courses were developed as part of a new intercultural learning scholarship program for semester abroad students.[2] Each of the courses available to students in the program offers one hour of academic credit and a scholarship awarded upon successful completion.

In order to scale up course offerings to serve the nearly 400 students studying in semester-long programs each year, Purdue University's Center for Intercultural Learning, Mentorship, Assessment and Research (CILMAR) entered into a partnership in 2016 with AFS-Intercultural, Inc., a renowned non-profit international youth exchange organization. This partnership provided the AFS Global Competence Certificate (GCC) online

curriculum as a digital text to the university at a reduced cost, allowing a significant percentage of semester abroad students to receive mentoring in a group setting rather than individually. Coursework for the scholarship program is, thus, divided into two separate curriculum groups: (1) courses offered through the Colleges of Science and Engineering, in which students are mentored individually; and (2) a course offered through the Brian Lamb School of Communication for students mentored in groups with the customized GCC curriculum.[3]

In the past two academic years the program has involved 50 mentors from 27 units across the campus. Mentors are recruited from a university master list of Intercultural Development Inventory (IDI) Qualified Administrators to ensure theoretically informed feedback and support that targets the participant's Developmental Orientation—hereafter referred to as stage-based feedback and support or stage-based pedagogy. In addition, mentors in the Communication course are also required to be trained as Qualified Facilitators of the Global Competence Certificate. Both groups of mentors receive further training from the course coordinator, who is an intercultural learning specialist, in a pre-semester mentor orientation. During this orientation mentors are trained in the application of the AAC&U Intercultural Knowledge and Competence VALUE Rubric and IDI stage-based pedagogy, as well as in the mentoring best practice of balancing challenge and support (Sanford, 1968). Thus, in this context, intercultural competence as defined and measured by the AAC&U Intercultural Knowledge and Competence VALUE rubric and the IDI are foundational in the training of mentors, the pedagogical framework of the intervention, and the metrics by which student learning and program success are evaluated. The AAC&U Intercultural Knowledge and Competence VALUE rubric and the IDI are products of well-researched developmental theoretical frameworks of intercultural competence.

Theoretical Frameworks

Intercultural competence is "the ability to communicate effectively and appropriately in intercultural situations based on one's intercultural knowledge, skills, and attitudes" (Deardorff, 2004, p. 194). Developing intercultural competency leads to improved understandings of other worldviews and the ability to manage thoughts, emotions, change, and ambiguity in order to effectively build bridges across cultural diversity. Many contemporary scholars view intercultural competence within a developmental framework—that is, as a set of related competencies that can improve over time as a result of experience, not merely a collection

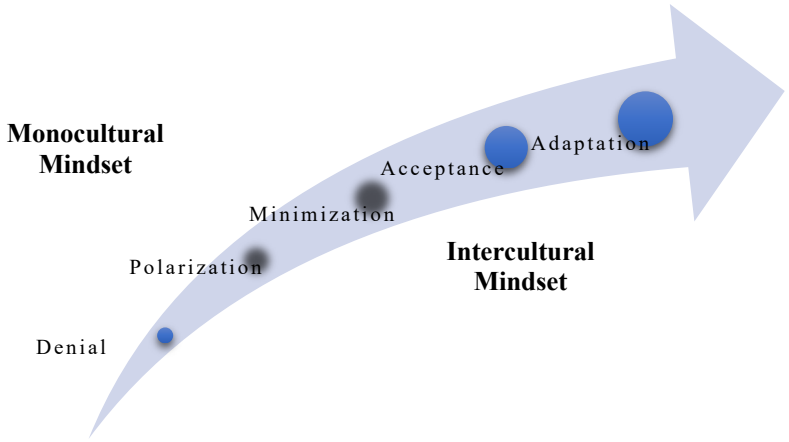
of static personality traits (Hammer, 2015). In other words, intercultural competence is imminently learnable. Myriad definitions of intercultural competence exist in the scholarly literature, with some emphasizing the role of language and localized cultural knowledge, and others focusing on globally transferable components such as positive attitudes toward cultural difference and cultural discovery skills (Acheson & Schneider-Bean, 2019). Most scholars, however, agree that intercultural competence comprises some combination of attitudes/affect, knowledge/cognition, and skills/behavior (Spitzberg & Changnon, 2009) and that the construct is measurable both holistically and broken down into its components (Fantini, 2009).

The DMIS and IDC

One of the foundational models in stage-based intercultural development is the Developmental Model of Intercultural Sensitivity, or DMIS (Bennett, 1986). The DMIS models a developmental spectrum comprising six stages that range from the more ethnocentric orientations of denial, defense, and minimization to the more ethnorelative orientations of acceptance, adaptation, and integration (Bennett, 1993). Several decades ago, Bennett and Hammer built a psychometric instrument based on the DMIS, the Intercultural Development Inventory (IDI), which positions respondents on the spectrum based on their degree of agreement with 50 Likert-scale items (Hammer, Bennett, & Wiseman, 2003). Since then, the IDI's validity and reliability have been well-established across a range of projects, including many completed by external researchers such as Paige, Jacobs-Cassuto, Yershova, and DeJaeghere (2003). A prime example of this type of research is the Georgetown Consortium Project, a large-scale multi-institutional analysis performed by Vande Berg, Connor-Linton, and Paige (2009). The IDI is often used for the purpose of program evaluation in study abroad contexts (for example, Jackson, 2011; Van Der Poel, 2016).

More recently, Hammer (2012) revised the DMIS in light of the large set of existing IDI data, arguing that the Intercultural Development Continuum (IDC), with its five rather than six stages, better aligns statistically with the IDI instrument. Hammer also relabeled the ends of the spectrum with the terms *monocultural mindset* and *intercultural mindset* instead of *ethnocentristic* and *ethnorelative*, as depicted in Figure 1. The IDC was chosen as a theoretical framework over the DMIS for this study because of its alignment with the IDI—that is, because results from the instrument were returned to participants and explained in terms of the IDC.

Figure 1
The Intercultural Development Continuum
(www.idiinventory.com)



Formative and Summative Assessment

Both formative and summative assessment practices are critical components of study abroad program evaluation (Saunders, Hogan, & Olson, 2015). While summative assessment of learning outcomes provides information about the growth and development of students as a result of their participation, formative assessments are strategically important for a number of reasons, including understanding who learners are when they enter a program (Wandschneider et al., 2015), determining variances in student preparedness and background knowledge (Dixson & Worrell, 2016), and motivating learners to set their own learning goals in line with Self-Determination Theory (Chirkov, Vansteenkiste, Tao, & Lynch, 2007). In this study, the program design capitalizes on both the formative and summative potential of the intercultural assessment used, providing feedback to both mentors and students at the beginning of the program and measuring their growth through pretest-posttest comparisons.

Methods

Subjects

This study examines data gathered from 376 undergraduate students who took part in a semester or year-long study abroad programs during fall and spring academic semesters of 2017 and 2018. The treatment group comprised students who opted into a scholarship program to promote intercultural learning in study abroad. Additionally, students included in the control group were recruited by e-mail from among all undergraduate students who took part in a semester abroad program during the same time frame but did not receive any structured intercultural mentored intervention while abroad. All students were required to complete the Intercultural Development Inventory at the beginning (T1) and end (T2) of their semester abroad.

Interventions

Students in the first treatment group received individualized mentoring and support during their semester abroad that was tailored to their individual IDI stage over a 14-week period. Prior to departure, students were individually debriefed on their IDI T1 results by their assigned mentor. Each of the weekly intercultural assignments required students to engage with the host culture to complete activities, followed by guided conversations within an online discussion board thread. A handful of these assignments included discussion among a peer group cohort. The remaining assignments constituted private discussions between the students and their assigned mentor. Following the initial assignment submission, students were provided brief feedback from their mentor (or their peers in whole cohort activities) and were challenged to respond to questions addressing the assignment's competency area(s) and the subject's DO. This follow-up discussion required a minimum of two additional posts from the student before it was considered complete. Detailed descriptions of these assignments along with the course syllabus are available as online supplemental materials in the Intercultural Learning Hub (www.hubicl.org), a free resource site for intercultural teachers, trainers, and mentors.

Students in the second treatment group received group-mentored support during their semester abroad that included attention to a group IDI stage over a 16-week period. Students were grouped into cohorts of 5-15 based upon official program start dates and time zone. Prior to departure, students were given the opportunity to meet their mentors and cohorts in a pre-departure orientation meeting. Students in this treatment group

viewed a group debriefing video that focused on the aggregate mean and distribution of group results. The curriculum of the group-mentored course was the customized GCC, which was designed to help learners develop intercultural competence and deal with cultural adjustment by focusing on four key learning goals: self-awareness, awareness of others, emotional intelligence, and bridging cultural differences (https://afs.org/certificate/). The online curriculum requires students to engage with peers in forums to discuss module content presented primarily through short videos. Additionally, in each of the four units, students engage with the host culture to complete interactive activities followed by a reflection essay. Following each unit, students took part in a live webinar Facilitated Dialogue Session, during which their mentor provided learning support and facilitated connections between the intercultural content and the students' lived experience. More information on the Global Competence Curriculum is available from AFS (https://afs.org/certificate/), and the additional assignments added to the GCC curriculum are also available on the Intercultural Learning Hub (www.hubicl.org).

Instrument

The IDI assessment is the primary source of data collected in this study. The decision to use the IDI as the sole instrument for this study is based upon the tool's shared developmental origins with the Intercultural Knowledge and Competence VALUE Rubric[4] combined with Purdue University's adoption of the Intercultural Knowledge and Competence VALUE Rubric as a framework for curricular alignment. Additionally, Purdue University has found through program evaluations that "intentionally using the Intercultural Knowledge and Competence VALUE Rubric together with the IDI predicts significant increases in intercultural competency as measured by the IDI" (Calahan, 2018). Ultimately, the IDI most closely aligns with the learning outcomes assessed in this study.

Specifically, data for this study were limited to academic and demographic information paired with IDI T1 and T2 scores for treatment and control group subjects. The IDI is a well-researched (Engelking, 2018; Vande Berg et al., 2009; Watson & Wolfel, 2015), valid, and credible tool (Hammer, 2011, 2012; Hammer et al., 2003; Wiley, 2016) that is widely used to measure intercultural development in both business and education. The online psychometric assessment consists of 50 multiple-choice items and demographic questions.

Analysis of the data examines the Perceived Orientation, the Developmental Orientation, and the Orientation Gap retrieved from the T1 and

T2. The Perceived Orientation (PO) score represents how a participant subjectively views their own behavior when encountering cultural differences. In contrast, the Developmental Orientation (DO) score represents actual, objective behavior when encountering cultural differences. The Orientation Gap (OG) score is the difference between the PO and DO, providing insight into the subject's self-awareness and the accuracy of their self-evaluation.

The IDI is used as both a formative and summative assessment tool in the structured mentored interventions. The IDI is utilized formatively to provide students with feedback on their own, or their group's, pre-intervention intercultural mindset. These individual or group debriefings facilitate setting learning goals and support the motivation to intentionally pursue intercultural development while abroad. Then, the theoretical framework of the IDI, the Intercultural Development Continuum (IDC), is incorporated into the mentored intervention and course curricula. Summatively, the comparison of T1 and T2 reveals the efficacy of intervention on the intercultural development of students.

Data Collection

During the study abroad application process, students self-selected the scholarship program and an appropriate course option according to the requirements for their particular major/minor program of study (individually mentored course, group-mentored course, or a course offered by a program provider). Once students were accepted into the scholarship program, they were provided instructions via confidential e-mail for completing the pre-IDI assessment survey before the start of their study abroad program. A similar e-mail was sent to each student shortly before the end of their program. Control group subjects were administered the T1 and T2 using the same method and timeline. T1 and T2 data for each subject were paired with demographic data provided by Purdue University's study abroad office.

Strict criteria were created for selecting valid data for this study. Data for subjects were excluded if they did not receive structured intercultural mentoring, dropped out of the course or semester abroad program, or failed to complete the T1 and T2, in whole or in part ($n = 23$). Additionally, data from students who received the scholarship by completing a pre-approved course administered by external program providers were also excluded ($n = 15$). One subject was excluded as an outlier after receiving a T2 DO score that was far below the minimum 55-point threshold for IDI scores.

Statistical Analysis

SPSS Statistics 25 was used to analyze data, and Microsoft Excel was used to generate tables and charts. For a comparison of each group in this study to one another, a one-way analysis of variance (ANOVA) in conjunction with a Post Hoc Tukey HSD Test with $\alpha = 0.1$ was used to find means that are significantly different from group to group. This alpha provides statistical significance with a 90% degree of certainty.

With quasi-experimental research such as this study, which uses convenience samples and relies on existing treatment groups rather than random assignment, false negatives could easily occur with strict observance of a very low *P*-value. Also pertinent is the strong theoretical basis for assuming that mentorship will increase gains in intercultural learning; confirming this trend that has been demonstrated elsewhere (albeit with small *N*s) should not require a stringent *P* threshold.[5] Therefore, in this study, the significance threshold was set at $P < .1$, with the acknowledgment that the *P*-value is one indicator of the value and surety of findings, not a verdict of significance to be accepted without question.

The focus of this comparative analysis was the mean IDI change for each group. Additionally, data were analyzed by cross tabulating the relationships between multiple covariates, such as treatment groups, semester, gender, citizenship and academic majors according to specific data subsets (for example, pre- and post-perceived orientations, perceived orientation change, pre- and post-developmental orientations, developmental orientation change, orientation gap change, and the like). These cross tabulations allowed for closer examination of data subsets.

Results

Results for mean scores among control group subjects who received no intercultural intervention were consistent with the findings from previous research on study abroad (Lou & Bosley, 2012; Paige et al., 2012; Vande Berg, Quinn, & Menyhart, 2012), where average change in scores range between +0.7 and +6.65 points on the IDI for non-intervention groups. Figure 2 shows each of the groups' average DO gains by semester. Note that variations between spring and fall semesters may be impacted by differences in enrollment numbers.

Table 1 shows the analysis of variance between the average change in DO scores for individually mentored (INDIV) and group-mentored (GROUP) treatment groups and the control group (CONTROL). Table 2 displays the results of the Post Hoc Tukey HSD Test, which provides comparisons for each pair of subject groups.

Figure 2
Average DO Gains by Group and Semester

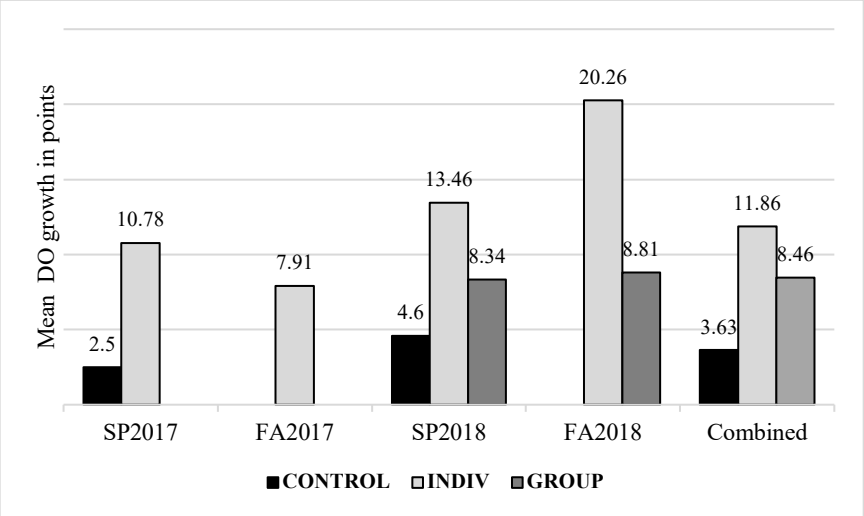


Table 1
ANOVA Results

	Sum of Squares	df	*Mean Square*	F	*Sig.*
Between Groups	2715.242	2	1357.621	6.313	.002
Within Groups	80213.703	373	215.050		
Total	82928.946	375			

The resulting *p*-values show statistical significance in average differences among all groups with a 90% degree of certainty. The Post Hoc Tukey HSD Test reveals that the greatest statistical significance is found between the control and the individually mentored treatment group, with $p = 0.002$. The difference between the control and the group-mentored treatment groups is less statistically significant, with $p = 0.092$, and between the individually mentored and the group-mentored treatment groups, with $p = 0.097$. By comparing the standard errors of each pair, it was found

Table 2
Post Hoc Tukey HSD Test
(Dependent Variable: Change in Mean Developmental Orientation)

					90% Confidence Interval	
(I) Group	*(J) Group*	*Mean Difference (I-J)*	*Std. Error*	*Sig.*	**Lower Bound**	**Upper Bound**
CONTROL	INDIV	-8.23596*	2.37698	.002	-13.1293	-3.3426
	GROUP	-4.83618*	2.30590	.092	-9.5832	-.0892
INDIV	CONTROL	8.23596*	2.37698	.002	3.3426	13.1293
	GROUP	3.39977*	1.64196	.097	.0196	6.7800
GROUP	CONTROL	4.83618*	2.30590	.092	.0892	9.5832
	INDIV	-3.39977*	1.64196	.097	-6.7800	-.0196

Note. *The mean difference is significant at the 0.1 level.

that the sample variability in the mean difference between INDIV and GROUP treatments was the least variable, indicating that conclusions made using the current data would remain valid. This suggests that if this study were to be replicated, the results would be reliably similar for each of these pairs, with the highest reliability predicted in the comparison of INDIV and GROUP.

This analysis is further supported by an examination of the differences in average DO progression, stasis, and regression between groups as seen in Figure 3. The DO constitutes the primary data point for determining the efficacy of each intervention. The change in DO score from T1 to T2 reveals orientation stasis, the progression toward more intercultural mindsets or a regression toward an orientation of more monocultural mindsets.

Figure 4 illustrates that both treatment groups obtained greater gains in DO change in comparison to the control group. Notably, stasis is cut by nearly half from CONTROL to treatment groups. While there is a greater increase in the percentage of students progressing one whole orientation stage in GROUP than in INDIV, INDIV shows even greater gains in the percentage of students progressing two whole orientation stages or more. In total, 57% of INDIV students progressed one orientation or more compared to 52.7% among GROUP students. Also of note is the difference in combined regression for each group. INDIV had a combined percentage of 5.6% who regressed as compared to 13.7% in GROUP.

Figure 3
Group Pre- and Post-Developmental Orientations

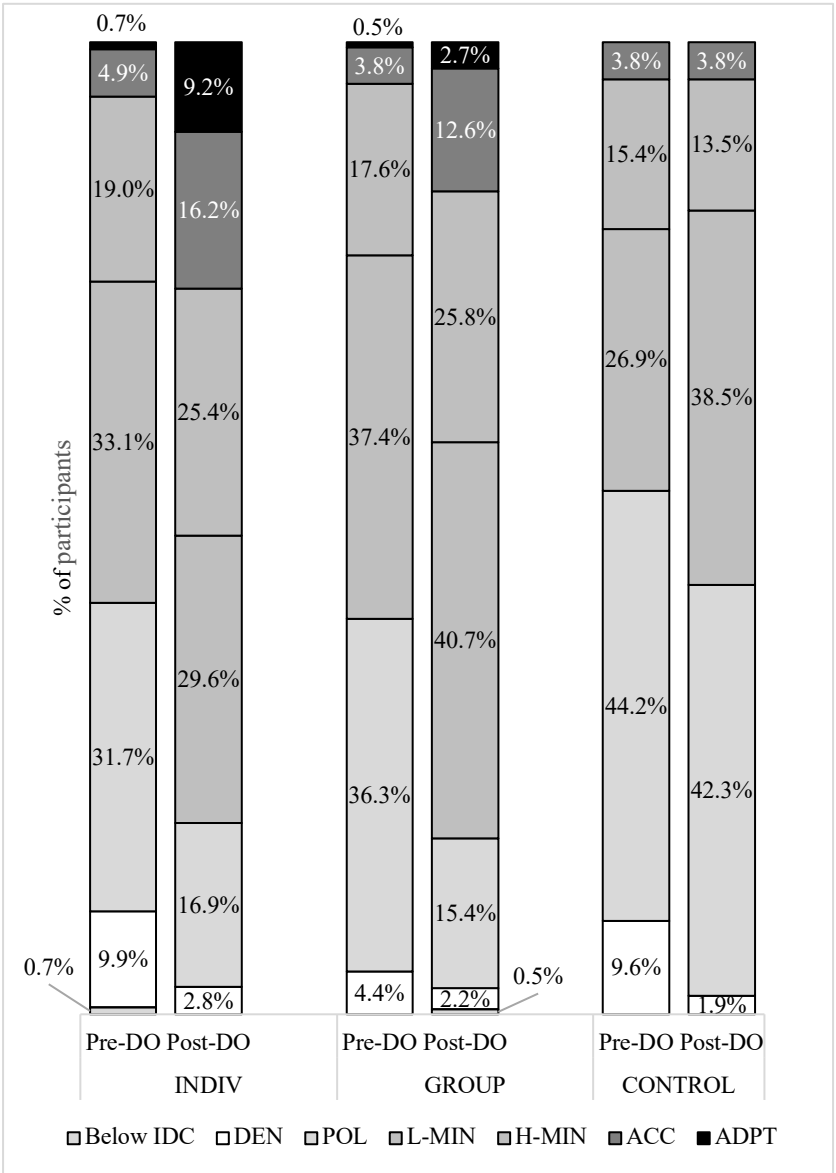
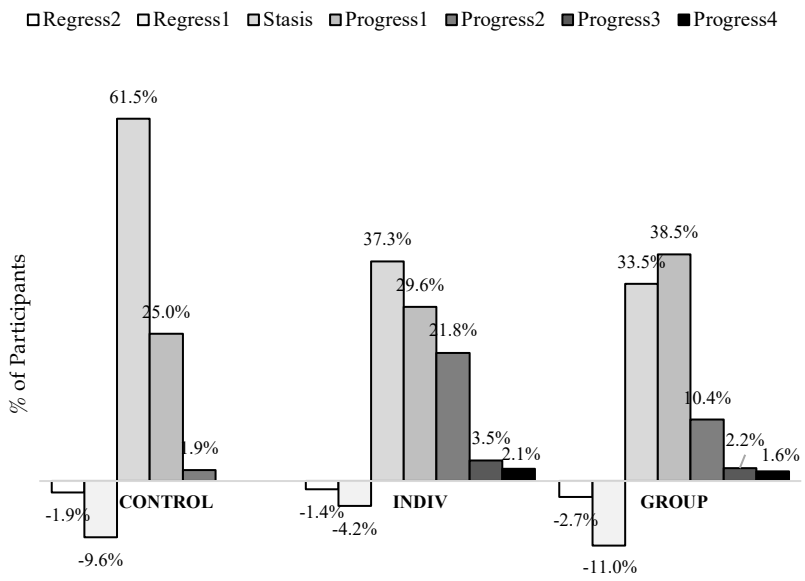


Figure 4
Average DO Change by Group



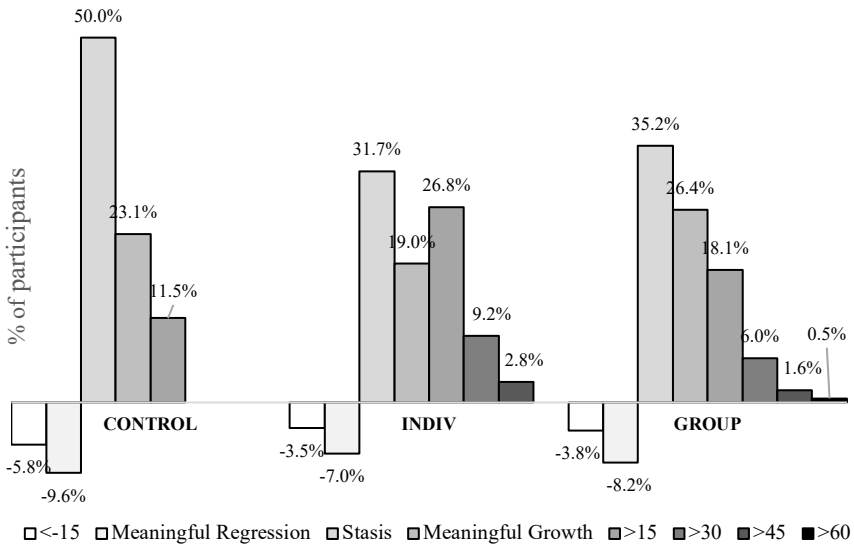
Note. Progress# indicates the number of whole orientations increased from pre- to post-DO. Regress# indicates the number of whole orientations decreased from pre- to post-DO.

A closer look at T1 and T2 DOs in Figure 4 allows us to see the nature of the changes taking place in each group. In both treatment groups, the monocultural orientations of Denial and Polarization are reduced by at least half, if not more, while more intercultural orientations like High Minimization, Acceptance and Adaptation are increased. CONTROL is most telling because it reveals that in the absence of facilitated intervention, by and large students tend to have their mindsets reinforced by their experiences abroad. The exception to this trend is found with CONTROL subjects who began in the Denial stage. The percentage of subjects who began in this particular orientation decreased from T1 to T2, likely due to intercultural exposure.

Because the analysis above focuses on the change from one orientation to another regardless of the number of points changed, it is limited in accurately representing the degree of change. A change from one orientation

to another could result from movement of only a few points on the scale. A more accurate view of this degree of change is gained by examining the difference in raw scores. In the analysis represented by Figure 5, categories of "Meaningful Regression" or "Meaningful Growth" indicate a change of 7.00 to 14.99 points, with 15.00 points being the point range for each orientation. As stated before, "stasis" is any change that falls between ±6.99 points. Beyond this initial measurement, growth and regression is divided into increments of 14.99 points.

Figure 5
Developmental Orientation Scores: Degree of Meaningful Change



This analysis reveals that INDIV showed meaningful growth among 57.8% of its students and GROUP showed meaningful growth among 52.6% of its students. The most notable difference between the two treatment groups were greater gains in growth of 15, 30, and 45 points for INDIV.

A closer examination of subject numbers at each orientation in T1 and T2 data provides further illustration of how the intervention types impact students' DO, as seen in Table 3. By examining the patterns of change from T1 to T2, a more precise view of growth and regression can be seen.

In Table 3 INDIV students saw the greatest increases when they began in Polarization. Also those who began in Low-Minimization had the greatest number experiencing orientation stasis. Looking at T1 and T2 totals, a clear picture forms of which orientations experience the greatest change: those on the extreme ends of the continuum. Analysis of data for GROUP students shows similar results.

Table 3
Individually Mentored Intervention
Pre- to Post-Developmental Orientation Change

Pre-Developmental Orientation	*Post-Developmental Orientation*						
	DEN	**POL**	**L-MIN**	**H-MIN**	**ACC**	**ADPT**	**Total**
Below IDC	1						1
DEN	1	6	7				14
POL	1	14	14	10	3	3	45
L-MIN	1	3	20	12	9	2	47
H-MIN		1	1	13	7	5	27
ACC				1	4	2	7
ADPT						1	1
Total	4	24	42	36	23	13	142

By examining the change in Orientation Gap, which includes a pairing of individual DO and PO scores, it can be determined whether students are becoming more or less accurate or realistic in their subjective self-evaluations. Among INDIV, 22.1% more students decreased the size of their gap compared to CONTROL, and 2.6% fewer INDIV students than CONTROL subjects showed an increase in their gap. GROUP students showed similar, albeit slightly less, improvement in closing their OG.

In addition to analyzing intergroup data, we made analyses of subject covariates focusing on gender, academic field, and citizenship. An analysis of gender comparison using degree of meaningful change in DO revealed no meaningful difference, with each degree of change falling within less

than 10 percentage points for male and female students. Likewise, a separate analysis of treatment groups revealed no meaningful difference between genders.[6]

In order to reliably examine the impact of intercultural intervention according to academic fields, analysis had to be carried out on both treatment groups combined because of insufficient numbers in each field when separated according to intervention type. Thus, an analysis of the degree of meaningful change in DO according to the academic field covariate revealed that among students in STEM fields, Agriculture showed the greatest percentage of combined growth (72.2%), followed by those in the College of Science (58.4%) and Engineering (53.5%). Notably, Purdue Polytechnic Institute students showed the least percentage of combined growth (33.3%) and, among all academic units, the greatest amount of stasis (55.6%).

The same analysis revealed that for students in the non-STEM fields, the College of Liberal Arts showed the greatest percentage of combined growth (59%), followed by the School of Management (50%). The College of Health and Human Sciences had the least amount of combined growth (44.7%) and the greatest amount of stasis (48.3%).[7] Purdue Polytechnic Institute data also presents results that require further investigation. It is unclear as to why technology students in particular exhibit more stasis than other subgroups. This is especially intriguing when Purdue Polytechnic Institute results are compared to Engineering results.

When the same analysis of the academic field covariate was carried out for each treatment group separately, the only academic fields that contained sufficient numbers for comparison were in the College of Science and the College of Engineering. In the individually mentored intervention, more College of Science students experienced meaningful growth (73.0%) than College of Engineering students (57.0%), but more College of Engineering students experienced meaningful regression (29.9%). Both experienced stasis in about a quarter of their populations. In the group-mentored intervention, more College of Engineering students experienced meaningful growth (48.7.0%) than College of Science students (33.3%), but more College of Science students experienced meaningful regression (22.2%) than did College of Engineering students (15.8). Both experienced stasis in over 40% of their populations.[8]

An analysis of the citizenship status covariate using degree of meaningful change in DO revealed that each of the groups attained meaningful growth in roughly half of all students. Students who are US citizens and permanent residents showed the least percentage of total meaningful regression (10.4% and 11.1% respectively), while international students showed twice the percentage of total regression (24.4%). Stasis among

U.S. citizens and permanent residents was also fairly similar, albeit with a larger margin (37.1% and 33.3%, respectively). International students displayed stasis among only 26.8% of subjects.

Results for permanent residents are likely skewed due to the small group size (n = 9). For this reason, the permanent resident group was paired with the international group. The rationale behind pairing these two groups was based upon the similarity in intercultural experiences as opposed to legal definitions of citizenship. The result of this analysis reveals a much more equal comparison of the two major groups. The one notable difference in this analysis was found in the percentage of total meaningful regression. International and permanent resident subjects combined had a similar total percentage of meaningful regression (12%) compared to domestic students (10.4%), which is a greatly reduced gap compared to the above analysis.

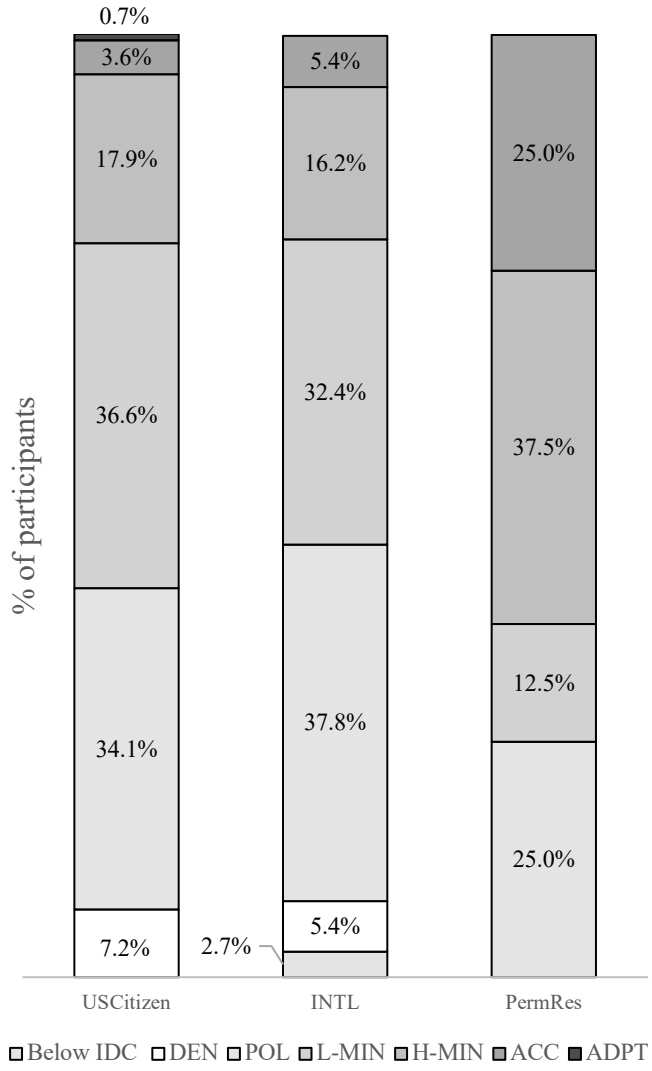
Interestingly, upon examination of the pre-DO scores represented in Figure 6, permanent residents were found to have a much narrower range of beginning orientations that trended more toward intercultural mindsets than the other two groups, beginning with a smaller percentage of students at polarization and low minimization and a higher percentage of students at high minimization and acceptance compared to the other two groups. Again, this is likely due to the much smaller group size for permanent residents. When returning to the analysis of combined scores for international students with permanent residents, the orientations appear in similar percentages for both groups.

Discussion

Analyses of the data demonstrate that this study is consistent with findings from previous studies regarding the impact of facilitated intervention on intercultural development as measured by the IDI. Reviewing the results displayed in Table 1, a familiar pattern emerges confirming the findings of these studies: The more facilitated is the intercultural intervention, the greater are the gains in students' intercultural development. The greater average gains resulting from the individually mentored intervention (+11.86) compared to the group-mentored intervention (+8.46) suggest that the one-on-one nature of the individually mentored intervention provides more impactful mentoring as assessed by the IDI. Nevertheless, the analysis of the data consistently shows that group-mentored intervention provides the next best option, especially when addressing issues of scaling up.

The fact that students in the individually mentored intervention are debriefed on their individual IDI results and repeatedly discuss their

Figure 6
Pre-Developmental Orientation Scores by Citizenship Status



behaviors in relation to their DO likely has a greater impact on their ability to progress along the continuum. It is also worth noting the sharp

increase in 2018 in the individually mentored course which correlates to the addition of two additional turns (posts) minimum in the assignment follow-up discussion.

Consideration was also given to the potential outcomes of decreasing the Orientation Gap for each intervention type. The assumption is that a narrowing of the gap between actual and perceived behavior indicates greater self-awareness and, therefore, greater intercultural development gains. This was another area in which the individually mentored intervention seemed to have greater impact, although by only a slight margin over the group-mentored intervention. Again, the slightly greater gains in the individually mentored intervention compared to the group-mentored intervention seem to be due to differences in IDI debriefing results and stage-based mentoring.

When interpreting the analysis of gains according to academic enrollment of all students who received intercultural intervention, the dominant pattern suggests that STEM students made greater gains than non-STEM students. This outcome may seem counterintuitive for some, but it makes sense when considering starting orientations of each group. Non-STEM students, on average, had greater numbers beginning in High-Minimization through Adaptation compared to their STEM counterparts. This left them with less room to develop along the continuum by comparison. It is also worth noting that because of stereotypes about social abilities among STEM students, STEM units tend to provide more institutional support for programs that assist this population in developing interpersonal skills. By way of example, the Colleges of Science and Engineering have historically provided more explicitly embedded intercultural learning opportunities within their curricula.

When the academic field covariate was separated by intervention type, the data revealed further evidence in support of the individually mentored intervention having greater impact on intercultural development compared to group-mentored intervention. The individually mentored course resulted in greater growth, less regression, and significantly less stasis for both College of Science and Engineering students.

Challenges

There have been numerous revisions over the life of this program. Students enrolled voluntarily, submitted work on BlackBoard Learn and via e-mail, and received a certificate of completion. Feedback was provided, but the extent to which this feedback was engaged by students remained unclear. As the intervention was improved, course credit was added

along with a scholarship. At the same time credit and the scholarship were added, minor revisions were made to individual assignments and the two-turn follow-up conversation requirement was added. To enhance the discursive nature of this multi-turn conversation, submissions were changed from assignment uploads in the virtual learning environment to discussion board posts.

Aside from the difference in mentoring approaches (individual vs. group), curricular variables contributed to the impact of the interventions on development. While students in both courses are assessed using the IDI as a pre- and post-test, the students enrolled in the individually mentored course are individually debriefed by their mentor and are, therefore, aware of their specific Developmental Orientation. For logistical reasons, the students enrolled in the group-mentored course receive a group IDI debriefing that provides them only with a group mean and distribution percentages. The result is that group-mentored students can make only an educated guess as to their DO. This is likely a factor that limits their ability to progress in their intercultural development.

Both interventions suffer from the limitations of the virtual learning platforms employed. While the familiarity students have with BlackBoard Learn as Purdue University's course management software is an advantage, it can be difficult for instructors to navigate. The group-mentored course required students to use a new virtual learning platform and an unfamiliar video-chat software. Additionally, limitations of the Bridge App platform required AFS-Intercultural, Inc., to code new data retrieval software in order to provide mentors with necessary tools for monitoring student work.

The capacity for an institution to provide quality instruction to a large number of students is limited by many factors, most notably staffing. One early challenge that had to be overcome was providing training to each of the staff members slated to mentor in this program. This presented challenges of scheduling and funding for each training. Were it not for the administrative and financial support made available, the number of staff members mentoring students in this program would not have been sufficient.

The variation among mentor expertise presents an equally confounding challenge. Due to increased enrollment in the individually mentored course, volunteer mentors were recruited from partner units across campus. While mentors were required to be Qualified Administrators of the IDI, intercultural instruction is not the primary role of their positions at the university. As such, their expertise in and commitment to the course varied. Efforts have been made to minimize the number of volunteer

mentors while also incentivizing and rewarding those volunteers who provide consistent, high quality mentoring. Again, this solution is not available to all institutions, and the researchers recognize the advantageous position CILMAR is in by having access to such financial support. The degree to which this variation in curricular implementation affects subject development is unclear. Similarly, changes made to curricula and mentor training may affect subject development. Whether that is for good or ill is, as yet, undetermined and requires further study.

Limitations of the Study

The limitations of this study fall into three categories: subject variables, mentor variables, and intervention type. In addition to continuing this study to collect more data and further investigate the effectiveness of these specific intervention types, future research should explore each of these areas of limitations. While this study includes a relatively large number of students over the span of two academic years, the diversity among students is limited by the type of students attracted to semester abroad programs. Demands of scaling up have resulted in variation among mentors. All mentors receive foundational intercultural training, but mastery of this training and application of mentoring skills have yet to be examined fully. As the CIEE: Seminar on Living and Learning (Vande Berg et al., 2012) study suggests, "cultural mentors," such as those utilized here, need support and in-depth mentoring themselves to best support successful study abroad interventions. The types of intervention, as with most curricula, have undergone repeated revision with the intent of improving efficacy. The extent to which specific curricular changes have resulted in improved growth has not be causally established. Additionally, the instructional delivery method provides multiple opportunities for comparative research. What differences, if any, can be found among online only, hybrid, and in-person/onsite mentoring approaches?

Some suggested areas for further research into subject variables include internal and external factors. This study is limited in the variety of demographic and experiential backgrounds of its subjects. Most students in the study come from affluent backgrounds and have greater access to financial support and, therefore, greater opportunity to travel. Even among this demographic the amount and type of travel differs and plays an as-yet indeterminate role in the students' intercultural development. Also worth consideration is the role, if any, that pre- and post-study abroad processes and programs may have on intercultural development and retention of this development. Additionally, the degree of similarity

between the students' own culture and that of the host culture may play an important role. Within the population of students, citizenship draws into question the degree to which a subject's prior experiences of living and studying abroad affects their openness to intercultural development. Of particular interest and need is an examination of the impact, if any, that beginning Developmental Orientation may have on developmental growth and regression.

Extrinsically, the degree to which the size and type of an institution may impact the success of interventions also needs to be explored. Specifically, the question may be whether small liberal arts colleges may find it easier to implement top-down edicts for involvement and support among its staff. As members of an institutional culture that endorses learning for learning's sake, employees may be more willing and supportive of such edicts that align with their own educational values. At larger research institutions, the focus tends to be on professionalization and may require more attention to the impact intercultural competence has on employability. An additional challenge at large universities is the tendency for units to become siloed and resistant to large-scale change. While buy-in may be easier to achieve at smaller institutions, the impact may be greater at larger institutions.

Additionally, variables in subject behaviors present further questions. Incentives for taking the intercultural learning course have been shown to directly impact enrollment, but the degree to which it impacts how seriously the students take the course remains an unexplored avenue. What impact would making the course a requirement of all study abroad and growing the curriculum from a one to a 3-credit hour course have on IC development, mentoring, and inter-unit partnerships? While there is a significant number of students for whom this program is a requirement, self-selection of study abroad may play an equal part in affecting intercultural development. Students who are more intrinsically motivated to interact with local cultures are expected to develop more. However, the degree to which this interaction is taking place is unknown and difficult to measure accurately. What impact do voluntary interaction, positive experience, and satisfaction have on student engagement and development? Within the population of this study, there is space to explore more deeply the differences among STEM fields compared to the liberal arts and the social sciences. What role, if any, does students' academic culture play in how they develop? Each of these variables presents challenges for closer examination and falls beyond the scope of this study.

Conclusions

Many scholars agree that facilitated intercultural intervention yields better results than no, or unguided, intervention. This study has addressed the question of which intervention type is best. As with many other institutions, Purdue University's Office of International Programs began with providing boutique programs to students that focused at first on integration and, with the founding of CILMAR in 2016, moved to intercultural competence. While smaller class sizes and individualized instruction translate to better learning, group-mentored intervention provides the opportunity to serve larger numbers of the student population. Group-mentored interventions likely fall into a sweet spot between boutique programs and individually mentored interventions in regard to maximizing student gains and enrollment size. This study reiterates, however, that one-on-one facilitated intervention remains the superior approach for maximizing intercultural gains during semester-long study abroad.

Institutions considering implementing such programs to impact the intercultural development of their students should investigate their own institutional goals and available resources and address the challenges of providing both individualized and group mentoring approaches. Assessing the needs of the institution, programs, and, most important, students points to a need for a broad and varied approach to facilitate the greatest impact with the resources available. While one-on-one facilitated intervention is superior, a group-mentored approach is also effective, especially when compared with no intervention in study abroad. Additionally, study abroad programs are not the only possible site of intercultural learning. One new frontier of intercultural learning is globally connected classrooms, also known as virtual exchange and Collaborative Online International Learning (COIL; see coil.suny.edu for more information). Institutions with limited resources for study abroad programs may also look inward to the intercultural nature of their institution and strategically assess the opportunities for intercultural learning "at home." Such strategic assessment may reveal that an institutional and holistic approach to intercultural competence training through the university's curricula may afford the greatest potential for impacting the intercultural development of all students.

Footnotes

[1]This underpins the use of the term "intercultural competence" and reflects an institutional understanding of the rubric as comprising the

knowledge(s), skills, and attitudes that can best be categorized as competencies.

[2]For the purposes of this study, the term "semester abroad program" refers to enrollment in a program approved by Purdue University's Study Abroad office that fulfills the enrollment requirement for a fall or spring term registration. Practically speaking, this describes programs that require residency in a foreign country for 3-6 months and include academic enrollment in a host institution and/or an approved internship. Additionally, the use of the term "mentoring" refers to the instructional and support role of faculty and staff for individuals or groups.

[3]Scholarship applicants also have the option of taking a pre-approved intercultural course offered by one of seven study abroad program providers for transfer credit(s). Because these courses are not administered by this research group, they have not been included in this study.

[4]"Milton Bennett's Developmental Model of Intercultural Sensitivity (1986, 1993) and Darla K. Deardorff's intercultural framework (2009) influenced the development of the AAC&U's Intercultural Knowledge and Competence VALUE rubric, and, as a result, it aligns very well with the Intercultural Development Inventory (IDI), which is used to assess intercultural competency or stage development (Hammer, Bennett, & Wiseman 2003)" (Calahan, 2018, para. 16).

[5]The setting of *P*-value thresholds for statistical significance is a fiercely contested and critiqued practice, with some scholars advocating for lowering the threshold to $p < .1$ in certain cases, for instance, in non-medical research (Hinkle, Wiersma, & Jurs, 2002); others demanding that the bar be raised as high as $p < .005$, especially for new discoveries (Benjamin et al., 2018); and still others suggesting doing away completely with such means of assuming the generalizability and reproducibility of study results so as not to encourage "*P* hacking, selective reporting and publication bias" (Amrhein & Greenland, 2018, p. 4). The approach we endorse is to strategically set and then justify a threshold in each study, dependent on purpose and context.

[6]It is our opinion that such analyses along the lines of gender are more harmful than useful, because they rely almost exclusively on narrow, binary definitions of gender.

[7]Because of a small sample size, data for the College of Education ($n = 1$) were omitted. Members of the control group were omitted in order to focus on the impact of the interventions on the development of participants as opposed to passive development.

[8]Results may be skewed in this analysis because of the small sample size. In the individually mentored course, the College of Science was represented by only 15 students compared to 107 in the College of Engineering. Likewise, in the individually mentored course, the College of

Science was represented by only 9 students compared to 76 in the College of Engineering.

References

Acheson, K., & Schneider-Bean, S. (2019). Representing the intercultural development continuum as a pendulum: Addressing the lived experiences of intercultural competence development and maintenance. *European Journal of Cross-Cultural Competence and Management*, *5*(1), 42-61.

AFS-Intercultural. (n.d.). Retrieved from https://www.afsnext.org/gcc/" https://afs.org/certificate/

Amrhein, V., & Greenland, S. (2018). Remove, rather than redefine, statistical significance. *Nature Human Behaviour*, *2*, 4.

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E. J., Berk, R., . . . Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, *2*, 6-10.

Bennett, M.J. (1986). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Cross-cultural orientation: New conceptualizations and applications* (pp. 27-70). Lanham, MD: University Press of America.

Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed.; pp. 21-71). Yarmouth, ME: Intercultural Press.

Calahan, C. (2018). A six-year journey of global learning faculty and student development. Retrieved from https://www.aacu.org/peerreview/2018/Winter/Calahan

Chirkov, V., Vansteenkiste, M., Tao, R., & Lynch, M. (2007). The role of self-determined motivation and goals for study abroad in the adaptation of international students. *International Journal of Intercultural Relations*, *31*(2), 199-222.

Deardorff, D. K. (2006). The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States. *Journal of Studies in International Education, 10*, 241-266

Dixson, D., & Worrell, F. (2016). Formative and summative assessment in the classroom. *Theory Into Practice, 55*(2), 153-159.

Engelking, T. L. (2018). Joe's laundry: Using critical incidents to develop intercultural and foreign language competence in study abroad and beyond. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *30*(2), 47-62.

Fantini, A. E. (2009). Assessing intercultural competence: Issues and tools. In D. K. Deardorff (Ed.), *The SAGE handbook of intercultural competence* (pp. 456-476). Los Angeles, CA: Sage.

Finley, A., & Rhodes, T. (2013). *Using the VALUE rubrics for improvement of learning and authentic assessment.* Washington, DC: American Association of Colleges and Universities.

Hammer, M. R. (2011). Additional cross-cultural validity testing of the Intercultural Development Inventory. *International Journal of Intercultural Relations, 35*, 474-487.

Hammer, M. R. (2012). The Intercultural Development Inventory: A new frontier in assessment and development of intercultural competence. In M. Vande Berg, M. Paige, & K. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 115-136). Sterling, VA: Stylus.

Hammer, M. R. (2015). The developmental paradigm for intercultural competence research. *International Journal of Intercultural Relations*, *48*, 12-13.

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

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2002). *Applied statistics for the behavioral sciences*. Boston, MA: Houghton Mifflin.

IDI, LLC. (2018). IDI general information: The Intercultural Development Inventory® (IDI®). Retrieved from https://idiinventory.com/publications/the-intercultural-development-inventory-idi/

Jackson, J. (2011). Host language proficiency, intercultural sensitivity, and study abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *21*, 167-188.

Krishnan, L. A., Masters, C., Holgate, H., Wang, C., & Calahan, C. A. (2017). Structured study abroad enhances intercultural competence. *Teaching and Learning in Communication Sciences and Disorders, 1*(1), Article 5.

Lochman, J. E., Dishion, T. J., Powell, N. P., Boxmeyer, C. L., Qu, L., & Sallee, M. (2015). Evidence-based preventive intervention for preadolescent aggressive children: One-year outcomes following randomization to group versus individual delivery. *Journal of Consulting and Clinical Psychology, 83*(4), 728-735.

Lou, K. H., & Bosley, G. W. (2012). Facilitating intercultural learning abroad: The intentional, targeted intervention model. In M. Vande Berg, M. Paige, & K. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 335-359). Sterling, VA: Stylus.

Lou, K. H., Vande Berg, M., & Paige, R. M. (2012). Intervening in student learning abroad: Closing insights. In M. Vande Berg, M. Paige, & K. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 411-419). Sterling, VA: Stylus.

Paige, R. M., Harvey, T. A., & McCleary, K. S. (2012). The maximizing study abroad project: Toward a pedagogy for culture and language learning.

In M. Vande Berg, M. Paige, & K. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 308-334). Sterling, VA: Stylus.

Paige, R. M., Jacobs-Cassuto, M., Yershova, Y. A., & DeJaeghere, J. (2003). Assessing intercultural sensitivity: An empirical analysis of the Hammer and Bennett Intercultural Development Inventory. *International Journal of Intercultural Relations*, *27*, 467-486.

Pedersen, P. J. (2010). Assessing intercultural effectiveness outcomes in a year-long study abroad program. *International Journal of Intercultural Relations*, *34*(1), 70-80.

Purdue University. (2012). Meeting of the University Senate; Senate Document 11-7 Undergraduate Core Curriculum Amended and Approved (pp. 32-41). Retrieved from Retrieved from https://www.purdue.edu/provost/students/s-initiatives/curriculum/documents/Senate_Document_11-7CoreCurriculum_Amended_and_Approved_1.pdf

Purdue University Data Digest. (2019). Undergraduate admissions: Student enrollment data, Fall 2018. Purdue University Office of Admissions. Retrieved from https://admissions.purdue.edu/academics/enrollment.php

Sanford, N. (1968). *Where colleges fail: A study of student as person*. San Francisco, CA: Jossey-Bass.

Saunders, K. P., Hogan, J., & Olson, C. L. (2015). Contextualizing the assessment journey and tools for the trek. In V. Savicki & E. Brewer (Eds.), *Assessing study abroad: Theory, tools, and practice (*pp. 83-102). Sterling, VA: Stylus.

Sessa, V. I., & London, M. (2015). *Continuous learning in organizations: Individual, group, and organizational perspectives*. New York, NY: Psychology Press.

Spitzberg, B. H., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. Deardorf (Ed.), *The Sage handbook of intercultural competence* (pp. 1-52). Thousand Oaks, CA: Sage.

Stead, L. F., Carroll, A. J., & Lancaster, T. (2017). *Group behavior therapy programmes for smoking cessation*. London, UK: John Wiley & Sons.

Vande Berg, M., Connor-Linton, J., & Paige, R. M. (2009). The Georgetown Consortium Project: Interventions for student learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad, 18,* 1-75.

Vande Berg, M., Quinn, M., & Menyhart, C. (2012). An experiment in developmental teaching and learning: The Council on International Educational Exchange's Seminar on Living and Learning Abroad. In M. Vande Berg, M. Paige, & K. Lou (Eds.), *Student learning abroad: What our students are learning, what they're not, and what we can do about it* (pp. 383-407). Sterling, VA: Stylus.

Van Der Poel, M. H. (2016). When does study abroad foster intercultural competence: A study in search of conditions. *European Journal of Cross-Cultural Competence and Management*, *4*(2), 168-185.

Wandschneider, E., Pysarchik, D. T., Sternberger, L. G., Ma, W., Acheson, K., Baltensperger, B., . . . Hart, V. (2015). The Forum BEVI Project: Applications and implications for international, multicultural, and transformative learning. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *25*, 150-228.

Watson, J. R., & Wolfel, R. (2015). The intersection of language and culture in study abroad: Assessment and analysis of study abroad outcomes. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *25*, 57-72.

Wiley, A. (2016). Intercultural Development Inventory (IDI): Independent review. Retrieved from https://idiinventory.com/wp-content/uploads/2017/12/2016-IDI-Validation-ACS-Independent-Review-Phase-3.pdf

Williams, T. R. (2017). Using a PRISM for reflecting: Providing tools for study abroad students to increase their intercultural competence. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *29*(2), 18-34.

Daniel C. Jones *has a master's degree in German linguistics and a Ph.D. in German literature from Purdue University. An Intercultural Learning Specialist in Purdue's Center for Intercultural Learning, Mentorship, Assessment, and Research, Daniel currently coordinates programs on intercultural competence in study abroad programs for students as well as faculty and staff.* **Michelle M. Campbell** *earned a Ph.D. in English in 2019 from Purdue University. She is a Communications Consultant for Graduate Communications and Intercultural Programs in the Pratt School of Engineering at Duke University, where she teaches graduate-level writing and oral communication. Before her position at Duke University, she was a graduate assistant for intercultural learning at Purdue's Center for Intercultural Learning, Mentorship, Assessment, and Research. She has broad research interests in multilingual and intercultural communication, graduate education, and cultural studies.* **Kris Acheson** *earned a Ph.D. in Intercultural Communication from Arizona State University in 2008. A former faculty member at Georgia State University and Fulbright Scholar to Honduras, she currently directs the Center for Intercultural Learning, Mentorship, Assessment, and Research at Purdue University, with a courtesy faculty appointment in Purdue's Brian Lamb School of Communication. She is widely published in a number of disciplines, with her most recent research interests centering on the development and assessment of intercultural competence.*