

OFFICE OF INSTITUTIONAL ASSESSMENT

Purdue University

Case Study:

Aligning the Assessment Instrument
with the Intervention and the Audience

June 2015



Katherine N. Yngve

Intercultural Learning Specialist

Office of Institutional Assessment

It is a central tenet of transformative education that students learn not through merely experiencing difference, but by critical reflection on their experiences of contact with difference¹. For the assessment professional or education abroad office, then, reflective writing holds out promise of being a rich source of direct evidence of student learning. In practice, however, even with well-designed topic prompts, few education abroad blogs or student journals achieve these standards of excellence. More often they devolve into a serial recitation of exciting daily events rather than the sort of anthropological 'thick description' which could qualify as solid evidence of deep learning. In addition, short-term study abroad, with its often intense pace of experience-seeking, can prove problematic in terms of creating conditions for meaningful reflection. This case study will review one institution's efforts to wrestle with these assessment hurdles, in pursuit of evidence of meaningful learning within the daily journal entries of 13 US participants as they spend two weeks learning and working in China. Although it deals with intercultural learning within a professional development context, we believe that laying bare both our initial assessment “muddle” and our ultimate success will be helpful to all whose education abroad programs incorporate reflective writing assignments.

Context

Since the summer of 2012, Purdue University, a land-grant, research-one institution located in central Indiana, has offered an annual summer short-term professional development seminar to China for selected student support staff members, such as academic advisors, residence life personnel and orientation leaders. This program has come about in direct response to an exponential growth over the last eight years in undergraduate non-domestic students. Originally designed as a sort of “pre-orientation” for incoming Chinese first-year students, the **Global Partners Program** [GPP] is also intended to produce, in staff members with no prior experience in Asia, a better understanding of the values and upbringing of Purdue's thousands of undergraduate students from China, Korea, Taiwan and

¹ Mezirow, J. (1991), *Transformative Dimensions of Adult Learning*, San Francisco: Jossey-Bass

Singapore. Designed to both challenge and support the professional learner during his or her first visit to Asia, the two-week sojourn includes multiple opportunities for Purdue staff to interact with Chinese nationals in both professional and tourism contexts. Post-program evaluations after the initial sojourn suggested, in fact, that it was perhaps a little too intense, causing culture shock challenges for those participants who had little prior travel experience. This led to a change in structure such that six weeks of pre-departure cultural and linguistic preparation has become a permanent part of the program.

Goals

The intended learning outcomes of GPP are for Purdue staff members outside the International Programs Office to: have an increased ability to work effectively with international students from all backgrounds, particularly those from Asia; and to feel better equipped to be strong advocates for undergraduate study abroad, particularly to non-European destinations. Returnees are also required to share their greater knowledge of Asia with staff colleagues, with the hope of creating a ‘ripple effect’ within the student support units across Purdue of knowledge about Asia, comfort with Asian students and familiarity with study abroad options in Asia.

In 2013, Purdue began assessing GPP participants with a pre/post administration of the Intercultural Development Inventory [IDI²] and in 2014 also began requiring daily guided journal reflections while in-country in order to help participants make greater learning gains. In the lead-up to GPP 2014, a matter of concern to the Program’s brain-trust was the fact that the IDI data from the 2013 cohort showed only a relatively modest average gain; with some participants making great forward strides, while others had a significantly lower post-participation score. Further, there was no obvious factor (age, gender, developmental stage prior to departure, etc.) which appeared to adequately explain the different responses to the experience. Yet, as is frequently the case with any study abroad, the GPP returnees all espoused a feeling that their experience was a meaningful and useful learning opportunity.

² Hammer, M, Bennett, M. and Wiseman, R. (2003) Measuring Intercultural Sensitivity, *International Journal of Intercultural Relations*, Vol. 7: 4, pp. 421-443.

Fortunately, the addition of a full-time intercultural learning specialist in Purdue's Office of Assessment in the summer of 2014 allowed for a deeper dive into the outcomes data for this flagship internationalization effort. For the 2014 GPP iteration, it was decided to postpone the second administration of the IDI (until late in the fall semester), but also to analyze the daily journal essays for insights into the stages or critical incidents of participants' intercultural odyssey. In other words, the primary goal for this assessment exercise was to lend clarity to our understanding of the IDI findings. This would, we hoped, give Purdue better insights into the personal and institutional benefits of offering this opportunity to selected staff members. Finally, we hoped to shed light on intercultural development outcomes of short-term education abroad for graduate and professional school students, since many programs for students beyond the BA level incorporate a similar professional component.

Methodology: Reflection + Rubric-based Assessment – Round One

GPP participants in the 2014 cohort were asked to journal nightly during their two weeks in China, beginning on the flight across the Pacific. The prompts for these journals were as follows:

- Describe any significant activities that occurred today.
- How did the events of the day affect you emotionally?
- What did you learn about yourself or others today?
- What was your highlight for the day?
- What was your biggest challenge for the day?
- Anything you would do differently if you could?
- What goals will you set for tomorrow?

Since, as part of a newly-adopted core curriculum, Purdue had defined 'intercultural effectiveness' as a six-component construct based on the AAC&U VALUE rubric for Intercultural Knowledge and Effectiveness³, the Purdue version of said rubric was adopted for evaluating these journal entries. This rubric comprises cognitive, affective, and behavioral elements, with a 4-point scale of defined achievement for each of the following components: cultural self-awareness, knowledge of non-US

³ The VALUE rubrics, co-developed by teams of faculty and academic professionals at over 100 US universities as part of a two-year project (2007-2009), are available for use and adaptation from the AAC&U website. See <https://www.aacu.org/value/rubrics>

cultural frameworks, empathy, verbal & non-verbal communication abilities, curiosity and openness.

Journal entries were de-identified, and an attempt was made to code sentences, comments and vignettes in line with the above six components of intercultural effectiveness as well as developmental level.

Findings of the Round One Assessment

Unfortunately, apart from giving assessors another way (in addition to the IDI) to quantify each participant's degree of openness and the growing complexity of their understanding of the Chinese way of life, the use of this rubric when applied to the reflection journals did not yield deep insights into the participants' process of acquiring or augmenting intercultural effectiveness during their time in China. For the most part, responses to the journal prompts did not regularly go into sufficient depth or analysis to allow alignment with the developmental levels on any of the six components of the VALUE rubric. Use of the rubric did, however, allow us to notice that the GPP participants went into their China experience with relatively well-developed skills of *empathy*, *curiosity*, and *communicative flexibility*; the journals also told us that most had an appreciation for Chinese history, art and cultural achievements. While it was useful to know which component skills of intercultural effectiveness our group had possessed prior to departure, it was also not surprising to detect evidence that experienced student support professionals have strong people skills or that staff at a university with a well-articulated global mission would find an overseas experience to be beneficial.

Sadly then, the findings derived from using this particular rubric did not allow the team to make bold statements about the benefits of GPP to Purdue. On the plus side, since the IDI does not isolate component skills along the route to an intercultural mindset, using the AAC&U rubric did permit us to know slightly more than we did before about each participant's metaphorical 'tool kit' for effectiveness, but not about how learning occurred. Still, despite not yielding what we had hoped for when analyzed with the VALUE rubric, the comments in the journals definitely suggested that the participants thought they were experiencing meaningful intercultural encounters. Was there another analytical tool which could help us make better sense of the qualitative data, even if our average

participant was not much given to verbose flights of critical self-analysis in the journal entries?

Methodology: Reflection + Rubric-based Assessment – Round Two

Returning to the literature, we found a plausible research approach in an article entitled “*Personal travel blogs as texts for studying intercultural interactions: a pilot test case study of an American sojourner’s blogs from Zimbabwe*”⁴. In this article, the authors propose a simple four-frame categorization system for textual data from travel blogs, based on an analysis of common topical themes which recur in travelers’ stories: Culture Shock, Cross-Cultural Comparison, Communication Challenge and Cultural Adaptation. This rubric draws heavily on the works of Gudykunst, Ting-Toomey & Kim, and operationalizes each of the four elements of the rubric via a set of guideposts which assist in categorizing qualitative data [expressing feelings of anxiety, deprivation, stress or homesickness, for example, are defined under the heading “culture shock”]. Pilot-testing of coding the Purdue GPP journal data using this categorization system was promising, but also suggested the need for two tweaks to the Malleus-Slaterry model for better inter-rater replicability and reliability. First, we agreed that, since this was a short-term program it would be more sensible to look for incidents of positive adaptive behavior (for example, trying to use one’s limited Chinese at a curbside food stand) rather than for fully-fledged narratives of Cultural Adaptation. Second, we found it necessary to add a fifth category, Cultural Appreciation (e.g., “I love the way Chinese respect older people.”) in order to codify input that did not fit well into the other four categories.

Findings of the Round Two Assessment

Coding the journal entries into the above 5 categories gave a completely different perspective of the participants’ learning process from attempting to analyze the same data using the AAC&U rubric. Now we were able to spot which participants started out on the lower side of the intercultural development range, as well as when they began to break out of their shell (and why), and to confirm

⁴ Malleus, R. and Slaterry, E. (2014), *Personal travel blogs as texts for studying intercultural interactions: a pilot test case study of an American sojourner’s blogs from Zimbabwe*, *SpringerPlus* 3:211 <http://www.springerplus.com/content/3/1/211>

that by the end of the two-week period most participants were able, at least occasionally, to see events from both their home-culture perspective and that of a Chinese national. More importantly, we were also able to see something about the group learning process which most survey-based assessment instruments do not probe: that often the impetus for breaking out of one's shell was the support of a more adventurous trip participant. In other words, as the program progressed, participants who had appeared by their initial journal writings to be more introverted or unsure of their capacity for sustained positive interaction with Chinese nationals were either the recipients of targeted assistance by the more extroverted or confident participants, or they independently realized that they could rely on their peers to help them overcome the perceived difficulties of interaction. In other words, participating in GPP helped student support staff form an effective community of practice, dedicated to stretching beyond one's professional comfort level in pursuit of higher standards of empathy and comfort with difference. This finding revealed an institutional success story which we could report to higher administration with pride, by providing evidence of positive GPP outcomes beyond the level of the individual learner.

Lessons Learned

Our first 'lesson learned' was that rarely can one get meaningful information about the weather by using a yardstick, or to put it more plainly, it's important to match one's assessment tool to both the learning objectives *and* to the learner's developmental stage. In this case, the VALUE rubric for Intercultural Knowledge and Effectiveness, although it operationalizes the component skills of Intercultural Competence quite clearly (and, at Purdue, constitutes the "baseline" definition for Intercultural Effectiveness as part of the Undergraduate Core Curriculum), was not a particularly useful tool for this assessment purpose. We believe that there are three main reasons for this misalignment: target group for the rubric, existing professional skillset of participants, and the way GPP participants engaged with the reflection questions. To be more precise:

- 1) The VALUE rubrics, as their unpacked acronym (Valid Assessment of Learning in Undergraduate Education) reveals, were developed for use primarily with college students between the

approximate ages of 18 and 22. While it is not the case that the acquisition of intercultural competence is simply a maturational process, it is demonstrably true that some of the skills most commonly associated with intercultural success (for example, critical thinking and self-awareness) are often not fully developed at the typical age of college entry. Purdue GPP participants, on the other hand, have all been practicing professionals, few of them under the age of 30.

- 2) Although there are often at least one or two GPP participants each year who have not travelled outside the USA before, there are many who have travelled or studied abroad at least once despite having no direct experience in Asia. Further, even those with little prior travel experience usually have a well-developed sense of global awareness; presumably as a result of having attained at least one college degree before choosing to work at Purdue, a very internationalized university.
- 3) For a variety of reasons, the reflection questions did not often generate data which aligned well with the VALUE rubric for Intercultural Knowledge and Effectiveness. We believe this is because the rubric assumes that the learner who is being assessed is not yet in the developmental state of mastery or high proficiency when it comes to openness, self-awareness, empathy, communicative competence, etc. By contrast, most GPP participants had a number of these skills well in hand prior to departure, or they would not have been successful in their current jobs! In addition, it was plain that, on any given evening during the two-week experience, anywhere from one to several of the participants would be too cognitively drained by the day's events to engage in answering the reflection prompts in anything but the briefest possible manner. Reflection is, after all, a cognitively demanding process, and so is converting reflective ruminations into effective written communication. In short, between the sometimes telegraphic nature of participants' responses and the fact that most participants had high interpersonal skill levels before leaving, it was often difficult to discern evidence of intercultural development across the period of the experience.

The above is by no means meant to suggest that the VALUE rubric is a poor assessment tool; indeed, Purdue faculty have found it quite helpful when used with short-term undergraduate education abroad.

Our second ‘lesson learned’ is that daily reflective writing is not ideal for everyone. In an intense short-term learning context for professionals, it may be more effective to alternate an end-of-the-day verbal group debrief over the beverage of choice with regular opportunities for more measured written reflection or an option for regular guided blogging (or audio-blogging). We note also that in some circumstances it is perfectly feasible to record the group debrief sessions in order to perform qualitative outcomes assessment at a later date. We believe that the Malleus-Slattery rubric offers a good alternative tool for analysis of written or transcribed data from journals, blogs and group debriefs; particularly in cases where the assessor is well-acquainted with intercultural learning theory.

Finally, our third ‘lesson learned’ was that we needed to think about further ways of assessing group or institutional outcomes of the Program (see below). The IDI and many quantitative instruments of intercultural assessment are intended to measure only individual learner development.

Summary

Using qualitative methods for this group showed us that there was an evolving community of practice which needed further challenge and support—as well as a different measure of analysis! This was an important finding that we might not have determined without attempting to use rubric-based assessment. As a result, we devised and administered a focus group protocol during a fall semester 2014 gathering of several annual cohorts of GPP returnees. We also resolved to administer a formative instrument (the Intercultural Effectiveness Survey⁵) to GPP 2015 participants (in addition to the IDI), with a formal pre-departure group debrief in order to help them better understand their existing personal tool-kit for individual and group intercultural effectiveness. But describing these further assessment pursuits would be a different case study!

⁵ Mendenhall, M, Stevens, M., Bird, A. and Oddou, G. (2008), *Specifying the Content Domains of the Intercultural Effectiveness Scale*, Chesterfield, Mo: Kozai Group.