IEPS International Education Forum Panel

Study Abroad: Measuring Learning Outcomes

Panel members

Darla Deardorff, Duke University
Jonathan Gordon, Georgia Institute of Technology
Celeste Kinginger, Pennsylvania State University
Susan Vernon-Gerstenfeld, Worcester Polytechnic Institute
Brian Whalen, Forum on Education Abroad

Moderator: Michael Vande Berg, CIEE
Rapporteur: Richard Gaulton, Cornell University

Summary of Presentation Topic

This panel was designed to move beyond describing the need for objective evaluations of study abroad, and to begin to discuss evaluation models that focus on learning outcomes. The panelist were to suggest ways in which the Title VI community can build on these models. Methods of reporting student learning outcomes were also to be discussed.

International educators find intrinsic value in studying abroad, as do an increasing number of undergraduate and graduate students. There are numerous programs which promote both sort- and long-term study abroad experiences. Yet we know that some programs are probably better than others. Study abroad programs have recently come under greater public scrutiny, and more international educators are calling for improved mechanisms for accountability. While it is simple to count numbers of students taking advantage of opportunities abroad, and to categorize programs by their length or their educational intensity, perhaps we are missing the true measure of a program's worth. By what metric may we measure the quality of study abroad programs? What have the students learned through their experiences abroad? How do we know they have learned it, and to what degree? This discussion is intended to move beyond describing the need to evaluate study abroad programs to introduce successful evaluation models that focus on learning outcomes.

Summary of Discussions

(The panel took the form of a free-ranging discussion among panelists and audience members, loosely organized around a series of questions proposed by the panel. The discussion began by noting the rapid increase in the number of studies of education abroad in the past decade and asking what forces are responsible for the increase.)

1. Who are the major stakeholders here, and what sorts of learning outcomes interest or motivate each of these stakeholder groups?

Panel and audience participants noted the pressure for accountability and outcomes assessment from many stakeholders, with different motives:

- program providers, eager to demonstrate the quality of their programs
- legislative and executive branches of the government, as a part of the general emphasis on accountability in education
- prospective employers, looking for assurance that study abroad students have desirable job skills
- accrediting agencies, beginning to consider study abroad as part of an institution's overall academic quality
- faculty, wanting to know if study abroad makes their students better students
- universities themselves, which are expending resources on study abroad at a time of scarcity.
- grant providers, seeking demonstrable outcomes from their grants
- students and parents, wanting to know if study abroad is "worth it" in terms of educational and career goals.

Study Abroad has borrowed the language of "outcomes assessment" from these stakeholders, but should be careful not to adopt uncritically their points of view. Institutions participating in study abroad should define their own goals and assess outcomes in terms of them. We might do well to encourage and assist students to develop and state their own goals, articulate their own leaning outcomes, and assess their study abroad experience in light of them.

2. What sorts of research needs should we be focusing on at this point?

Research in the area needs to be improved methodologically and must avoid the appearance of self-serving research. There is a real need to assess actual outcomes rather than simply to record the level of student satisfaction, to take into account the unreliability of much self-reported data, to compensate for sample bias, and to rely on sound methodology, valid and reliable instruments, mixed methodology, control groups, etc. In other words, research on outcomes assessment in study abroad must conform to the standards of high quality social science research. Since assessment itself consumes resources, we must avoid "over-assessment" – assessment for its own sake --or, even worse, "teaching to the test" in study abroad.

One needn't be an expert in assessment to conduct effective research; there are many potential partners in assessment (career services departments, institutional research offices) that may have the necessary expertise and an interest in helping.

The panel identified three main areas for substantive research:

(a) Assessment of academic progress, often in discipline-specific research, is the key area for many, especially the faculty. Research in language acquisition during study abroad is probably the most mature area of discipline-specific research, but research in other areas tends to be haphazard. Research designed to answer the question "How does education abroad contribute to making a better mechanical engineer, historian, or marketing major?" is difficult, since the tools for assessment in these fields often do not take into account the particular features of the study abroad experience.

(b) Assessment of the cultural learning aspect of study abroad, often defined these days as "intercultural competence," has made some progress by identifying key elements of the concept and moving toward a consensual definition. Much of the past research in this area is subject to question on the grounds of sampling errors, over-reliance on self-reporting, and bias toward socially correct answers. With the development of some more commonly accepted instruments for measuring "intercultural competence," the time is ripe for more ambitious and sophisticated research. Differing opinions on the importance of culturally specific learning as opposed to intercultural competence may complicate the research agenda.

(c) Research on the long-term outcomes of study abroad. Many returned students testify that study abroad has been a "life-changing experience" -- research measuring those life changes over time therefore becomes necessary. Long-term longitudinal studies, however, are among the most costly and difficult research to conduct, hampered by the difficulties of contacting study abroad participants long after the fact, low response rates, and sample bias. Participant perspectives on the study abroad experience tend to change over time after the experience ends, so research based on cohorts may be a strong approach. Gordon reported on the results of a study of Georgia Tech alumni 3-5 years after graduation, finding that study abroad students had a higher level of career satisfaction, felt better prepared for employment, and earned more than non-study abroad students.

3. What standards or metrics now exist that allow us to assess student learning abroad? What standards still need to be identified?

Effective standards have been developed in some areas, but more needs to be done to develop metrics that can be widely accepted.

In language acquisition, assessment of language-learning progress on the ACTFL scale, based on pre- and post- testing using Oral Proficiency Interviews, provides a standard that is commonly, although certainly not universally, accepted. According to Kinginger,

research in the field tends to focus on documenting development typical of academic learning, downplaying the sociolinguistic and pragmatic dimensions of language learning and the links between language and culture learning likely to emerge in study abroad. Nevertheless, the ACTFL and OPI system is close to being an accepted standard for the field. The multi-year studies of Russian language acquisition in study abroad (Brecht, Davidson, and Ginsberg) are the strongest body of research in this area, and although primarily concerned with language acquisition they have raised many issues concerning the relationship between language acquisition and the structure of study abroad programs and their settings.

Deardorff reported that surveys of experts reveal that there is substantial consensus about the component parts and definition of the concept of Intercultural Competence. With this emerging consensus, the numerous instruments independently devised to measure versions of Intercultural Competence can be consolidated, laying the foundation for comparative research. Bennett's Intercultural Developmental Inventory (IDI) and Shealy's Beliefs, Events, and Values Inventory (BEVI) are two of the more promising instruments currently in use. Many independently-developed instruments are, as they are tested, being consolidated into a relatively few instruments that have the potential for wide acceptance.

Engle and Engle, in their article "Toward a Classification of Study Abroad Programs," have succeeded in identifying key program variables affecting outcomes, such as duration, type of housing, language level of the students, etc. According to Vande Berg, the Engle classification has allowed researchers to focus on the concrete experiential aspects of the study abroad experience, and have already yielded interesting (and counter-intuitive) results. These variables have proven powerful in such studies as the Georgetown Consortium Studies and the State of Georgia study.

Metrics for assessing progress in academic disciplines are understandably not standardized, and designed primarily to measure outcomes in very specific programs. Vernon-Gerstenfeld reported that Worcester Polytechnic Institute has been developing an assessment model based on student journals and blogs after finding that their engineering students engaged in education abroad didn't show progress on standard measures of intercultural competence. It's interesting that WPI and Georgia Tech, two engineering institutions, appear to be leading in developing measures for their discipline.

4. Why are we doing this research?

Outcomes assessment research has been directed at three main goals:

- the scholarly goal of contributing to knowledge
- the political goal of justifying what we do and advocating effectively
- the instrumental goal of improving practice.

There was a lively discussion about the proper relationship among these goals as more institutions devote more attention to assessment of study abroad. There were strongly

expressed views that the political goal has moved to the fore under the increased pressure for "accountability" in many aspects of university life. Like it or not, the political atmosphere in which study abroad occurs today tends to dominate the research agenda.

Members of the audience pointed out that one area that has been little studied is the impact of study abroad on the local communities in which it occurs. The "local stakeholders" are developing their own interest in "outcomes assessment," and we should expect more demands for research in this area as the number of students and programs increases.

Gordon emphasized strongly that assessment research needs to be informed by an institution's particular goals in study abroad, as apart of that institution's "culture of assessment." It is by building outcomes assessment into a system in which we develop and articulate our own goals for academic progress and cultural or cross-cultural learning, based on the statements of mission adopted by our institutions, that we will generate fruitful results. Allowing outcomes assessment to be driven by externally generated goals will lead to half-hearted research and puny results. In a field as complex as study abroad, "teaching to the test" seems very much out of place.

(Additional questions initially posed by the panel)

5. What efforts are being made to implement the results of research studies—to have theory inform practice in the field?

6. Are there existing evaluation or assessment studies, or theoretical constructs, that can serve as research models? Can we identify the characteristics of these studies or constructs that make them particularly successful?

These questions were discussed as the previous questions were addressed.

Best Practices and Future Directions

The best practices in outcomes assessment involve the incorporation of study abroad outcomes assessment into an institution's overall assessment strategy, the development of methodologically sound and replicable research designs. What must be avoided is the production and publication of weak or self-serving research that will undermine the credibility. As Vande Berg noted during his introduction to the discussion, the past decade has seen a spectacular growth in the amount of research conducted and published on the topic of study abroad. The continued development of selective publishing outlets

for research of high quality in the field is vitally important to the development of our research efforts.

Richard Gaulton
Rapporteur
1 April 2008