© *Journal of Higher Education Outreach and Engagement*, Volume 17, Number 2, p. 67, (2013)
 ISSN 1534-6104

Learning Outcomes Assessment: Extrapolating from Study Abroad to International Service-Learning

Donald L. Rubin and Paul H. Matthews

Abstract

For international service-learning to thrive, it must document student learning outcomes that accrue to participants. The approaches to international service-learning assessment must be compelling to a variety of stakeholders. Recent large-scale projects in study abroad learning outcomes assessment—including the Georgia Learning Outcomes of Students Studying Abroad Research Initiative (GLOSSARI)—offer precedent from which international service-learning assessment programs may draw. This article outlines five promising practices to guide international service-learning assessment activities: (1) focus on outcomes about learning; (2) employ multiple sources and methods for data collection; (3) invest in compiling credible comparison groups to build the case for a causal relationship between international service-learning and learning; (4) acquire data from multiple and diverse institutions and programs to better generalize and also to warrant conclusions about best program practices; and (5) acquire data from large samples of program participants to provide insights into under-represented groups and program sites.

Introduction

What kinds of assessment practices promise to foster both proliferation and excellence in international service-learning? The purpose of this essay is to extrapolate selected promising practices for assessing student learning outcomes from the more fully established domain of study abroad to the still emerging domain of international service-learning. The primary source for these study abroad assessment practices is the Georgia Learning Outcomes of Students Studying Abroad Research Initiative *(GLOSSARI; Sutton & Rubin, 2004)*, a multi-year, multi-phased project that utilized diverse approaches to assessing learning outcomes over the course of 8 years among students attending more than 30 public institutions and enrolled in scores of study abroad programs.

Operating any international education program—and especially one focused on service to host nation communities—is not for

the faint of heart. The foremost concern is, without question, student safety in environments that are by design unfamiliar and often lacking expected on-campus infrastructure for risk control *(Burak & Hoffa, 2001)*. Program directors are, of course, also concerned about the quality of the learning experience for their students, as well as the benefits for the host community. Numerous reports of service-learning programs evaluate personal, social, and citizenship outcomes for students, but relatively few document academic or intellectual learning outcomes *(Conway, Amel, & Gerwien, 2009)*. Assessing student learning outcomes in service-learning can be quite labor-intensive *(Rama, Ravenscroft, Wolcott, & Zlotkowski, 2000)*; as is commonly practiced for on-campus programs, many international educators fall back on the convenience of student evaluations *(Engle & Engle, 2003)*, which are essentially customer satisfaction surveys. Generally, little is learned about program outcomes from student satisfaction surveys alone; even that hallmark of service-learning, reflective writing, when not carefully structured, may be "useful neither in assessing learning, evaluating programs, nor conducting research" *(Whitney & Clayton, 2011, p. 150)*.

Relegating assessment in international service-learning to convenient, but largely uninformative information sources, is a pedagogical and strategic misstep, however *(Steinberg, 2007)*. As Tonkin *(2011)* enjoins,

> [M]ore needs to be known about whether present [international service-learning] practices are achieving their objectives, or indeed achieving any objectives at all. Not only are . . . practitioners and researchers accountable to funders, institutions, and students, they are also accountable to their hosts and the public good. Thus, research is more than an academic exercise: it is an ethical imperative. *(p. 215)*

Thus, international service-learning will thrive to the degree that rigorous assessment processes hold it accountable to its various stakeholders—students, parents, and host communities as well as university administrators and academic disciplines. In particular, the following five "promising practices" may help guide this work, grounded in the experience of evaluating study abroad.

- From the perspective of adding value to a United States education, the most appropriate metrics for measuring the impact of international service-learning are student learning outcomes, including "hard" institutional

outcomes like college completion rates (where institutionally appropriate).

- International service-learning learning outcomes initiatives should deploy diverse approaches, including studies of students' success in their academic careers.
- The strongest policy and curricular case for international service-learning will require aggregation of findings across multiple programs and multiple institutions. This strategy requires alignment among those various programs and institutions regarding the questions posed and the metrics taken as evidence.
- To draw credible conclusions about the value added attributable to international education experiences, it is necessary to compile credible comparison groups of students who forgo those experiences, or those who choose to fulfill education abroad in differing formats (e.g., credit versus noncredit programs).
- Accumulating a large sample size provides insights about participation and outcomes for less represented groups and about a variety of program variables.

First, a caveat about the object of international service-learning assessment is in order. Student outcomes are not the only important outcomes from international service-learning, and perhaps not even the most important ones. International education that incorporates experiential components and service-learning should also research the impacts on the hosting communities *(Bringle & Hatcher, 2011; Sutton, 2011; Tonkin, 2011; Wells, Warchal, Ruiz, & Chapdelaine, 2011, p. 320)*. However, studies focusing on host communities are still generally rare *(Crabtree, 2008; Cruz & Giles, 2000; Wood, Banks, Galiardi, Koehn, & Schroeder, 2011)*. While acknowledging the importance of attending to this "equally important standard of community benefits" *(Bringle & Hatcher, 2011, p. 17)*, this essay deliberately adopts a student learning outcomes perspective, as that position is foundational to the scholarship of teaching and learning *(Banta, 2002; Marsh, 2007)*.

International Service-Learning

International service-learning and study abroad are two types of international education that often overlap, but do not coincide *(Bringle & Hatcher, 2011)*. Study abroad is associated with

formal credit-bearing classes that presumably impart some body of disciplinary knowledge using the international context as an instructional resource. This sort of study abroad has a long history *(see papers collected in DePaul & Hoffa, 2010)*. Alongside these traditional study abroad programs, service-learning and community-based experiential education elements are increasingly incorporated into international programs "as an effective way to complement and expand on existing study abroad course objectives" *(Kiely, 2011, p. 243)*. Properly performed, such "experiential activities . . . are not add-ons to meet student demand, but core activities that are at the heart of the study abroad experience" *(Steinberg, 2002, p. 223)*. Although there is clearly a wide range of structures and types of international service-learning *(e.g., Jones & Steinberg, 2011; Steinberg, 2002; Tonkin, 2011)*, including credit-bearing and non-credit-bearing experiences, this article follows Bringle and Hatcher's *(2011)* definition of international service-learning as

> A structured academic experience in another country in which students (a) participate in an organized service activity that addresses identified community needs; (b) learn from direct interaction and cross-cultural dialogue with others; and (c) reflect on the experience in such a way as to gain further understanding of course content, a deeper appreciation of the host country and the discipline, and an enhanced sense of their own responsibilities as citizens, locally and globally. *(p. 19)*

Many international service-learning experiences may qualify also as study abroad courses with defined disciplinary learning objectives and credit toward graduation. Sometimes the distinction reflects mainly a matter of degree of emphasis. Indeed, international service-learning and study abroad generally share a great many objectives, particularly those that speak to transformational learning among students. Learning that enhances self-knowledge and intercultural development is central to both international service-learning and study abroad *(Hoff, 2008; Pusch & Merrill, 2008)*, as are certain learning instructional practices such as experiential activity and reflection *(Montrose, 2002; Pagano & Roselle, 2009)*. Indeed, the ascendant rubric "global learning" implies a moral imperative for community engagement *(Hovland, 2006)*. Study abroad and service-learning are both identified as "high impact practices" that enhance student engagement and attendant outcomes such as grades, time to graduation, and advanced study *(Gonyea, 2008; Kuh, 2008)*.

On the other hand, many traditional study abroad programs appropriately adopt learning objectives more specific to a canon of disciplinary knowledge *(Brewer & Cunningham, 2009)*. Presumably a geology course taught in the Peruvian Andes seeks to impart a corpus of orographic information and field methods that differ (at least in degree) from the objectives of an international service-learning course on indigenous natural resource management taught in the same location. Conversely, such a service-learning course on natural resource management might encompass objectives pertaining to conducting community needs assessments that the traditional geology course might not.

Assessing Learning Outcomes in International Service-Learning

In the introductory chapter of their edited volume on international service-learning, Bringle and Hatcher *(2011)* note a "state of confounded rationales, program goals, and program types [that] complicates assessing study abroad outcomes to the point that there is limited high-quality evidence on its outcomes . . . gathered across programs" *(p. 9)*. Indeed, many aspects of study abroad and international service-learning have not been rigorously assessed to date. In his comprehensive review of the state of the latter's research agenda, Tonkin *(2011)* frames a wide range of pressing issues for research and assessment in international service-learning. These include research into "fundamental issues" such as program design, ethics, and the contexts of international service-learning; student recruitment, motivations, and readiness in these programs; faculty practices, attitudes and beliefs; the practice of international service-learning, such as curriculum development, technology, and preparation issues; questions relating to service abroad, such as impact on the hosting community and agencies; and the characteristics and outcomes of student participation in study abroad/international service-learning. While not the only important avenue of investigation, these student learning outcomes are the focus of this essay.

An examination of extant research on student learning outcomes assessment in study abroad and international service-learning highlights the challenges of effective evaluation and shows a need for additional, high quality research in this area, especially studies that are quantitative and those investigating more than a single program *(Bringle, Hatcher, & Williams, 2011)*. As Tonkin notes of the voluminous research on student assessment, "[v]irtually all

of this literature is ancillary to [international service-learning]" *(2011, p. 197)*.

Sutton, Miller, and Rubin *(2007; see also Sutton & Rubin, 2004)* draw sharp distinctions between *learning* outcomes assessment and other kinds of outcomes assessment in international education. Most study abroad outcomes research has examined changes in students' attitudes and personal development, or impact on life choices, as opposed to increased knowledge or skill (i.e., *learning* outcomes per se). The same is true of most international service-learning outcomes assessment research, which has mainly (though by no means exclusively) focused on attitudinal and dispositional outcomes such as development of identity as a global citizen and changes in intercultural sensitivity, global competence, and similar dispositional variables *(Kiely, 2011; Plater, 2011; Tonkin, 2011; Tonkin & Quiroga, 2004)*. To be sure, some international service-learning programs have pursued outcomes in the domains of academic learning (e.g., health professions— *Bentley & Ellison, 2007; Martinez-Mier, Soto-Rojas, Stelzner, Lorant, Riner, & Yoder, 2011*; entomology— *Robinette & Noblet, 2009*; teacher education—*Knutson Miller, & Gonzalez, 2011*), especially language proficiency and knowledge of the host country *(see Kiely, 2004; Steinberg, 2002)*. In surveying research on international service-learning outcomes, Tonkin *(2011)* particularly notes a dearth of attention to learning outcomes "that extend beyond the course level of analysis" *(p. 207)*, such as knowledge, degree attainment, and pursuit of postgraduate education.

Assessing international education learning outcomes beyond the course level invites a variety of research methodologies. In addition to largely qualitative case studies and examinations of student learning artifacts, one might consider administering measures and surveys *(Paige & Stallman, 2007)*, or collecting institutional data such as graduation rates *(O'Rear, Sutton, & Rubin, 2011)*. Standardized performance assessments for measuring liberal arts outcomes like critical thinking and analytical reasoning have also become available over the past decade (e.g., the Collegiate Learning Assessment; see *Arum, Roksa, & Cho, 2011; Bers & Swing, 2010*). Yet the preponderance of research on international service-learning has been undertaken through qualitative research traditions only *(Kiely & Hartman, 2011)*, "with most analyses being descriptive case studies of particular courses and programs" *(Bringle, Hatcher, & Williams, 2011, p. 276)*; the latter also "posit that a quantitative approach to research on [international service-learning] will yield fruitful results that can guide program design, improve practice, test theory, contribute to a knowledge base, and provide a basis for funding and support for program expansion" *(pp. 275–276)*.

Study Abroad Resources for Developing Promising Practices

The world of study abroad is by no means monolithic nor singularly advanced with respect to its assessment practices. To the contrary, many study abroad programs rely on enrollment "body counts" and student evaluations as their primary vehicles for evaluation *(Engle & Engle, 2003)*. Nonetheless, in recent years the field has launched several initiatives aimed at elevating the role of assessment in general, and moving toward more learning outcomes assessment in particular *(Sideli, 2001)*. The impetus for this trend derives from several sources, including the increasing scrutiny placed on study abroad as part of institutional reaccreditation processes. As regional accrediting bodies began routinely to accept study abroad participation as an indicator of institutional excellence, they simultaneously began encouraging institutions to document in greater detail the value added to general education (and other) objectives. In short, the learning outcomes assessment movement in study abroad was driven from the start by the challenge to provide convincing evidence to a variety of external as well as internal stakeholders. Demonstrating the legitimacy and value added of service-learning programs to internal university stakeholders and to community stakeholders has certainly been one important motivation for emphasizing the centrality of evaluation to the broader service-learning enterprise *(Nisbett, Tannenbaum,& Smither, 2009)*, and no doubt this motivation will eventually pervade international service-learning.

The University System of Georgia's GLOSSARI project.

In the year 2000, the Office of International Education at the University System of Georgia began investing in the Georgia Learning Outcomes of Students Studying Abroad Research Initative (GLOSSARI) project to assay learning outcomes accruing from study abroad at its approximately 35 constituent institutions. This system-wide initiative built on a number of strengths, not the least of which was the prior institution of a database for compiling information about every student participating in study abroad over a decade's time. Data regarding over 30,000 study abroad trips—which eventually were matched with nearly 20,000 complete academic records—provided unprecedented credibility for GLOSSARI's conclusions about such matters as the impact of studying abroad on graduation rates. One of the co-authors of this article was the director of research for the GLOSSARI project.

For details of GLOSSARI beyond those presented in Sutton and Rubin *(2004)*, see http://www.glossari.uga.edu.

Other data sources.

Shortly after the University System of Georgia initiated GLOSSARI, the U.S. Department of Education's Office of International Research and Studies funded a cluster of large-scale studies of learning outcomes of education abroad. Along with GLOSSARI, those projects included the Georgetown Consortium Project *(Van deBerg, Balcum, Scheid, & Whalen, 2006)* and the Study Abroad for Global Engagement project *(Paige, Fry, Stallman, Josić, & Jon, 2009)*.

Data Analysis

In reflecting on the applicability of this essay's conclusions to their own instructional settings, readers will of course need to recognize that even though the GLOSSARI institutions ranged across a broad swath of higher education, they are specific to a single state and do not include private colleges.

In some senses, the state of the field of international service-learning is similar to that of study abroad prior to the implementation of these large-scale research studies, with many unanswered questions floating on a sea of small-scale, qualitative, and program-specific descriptive research studies. To provide compelling evidence to move beyond this current status quo, then, international service-learning might productively learn from, and consider adopting, approaches to evaluation similar to those successfully implemented in this recent study abroad research. Part of the validation for the promising practices described in the following section derives from the experience of transporting and replicating these methods to other institutions. For example, the GLOSSARI methodology was adopted by the California Community College Student Outcomes Abroad Research initiative (see http://globaled.us/cccsoar/index.asp#top) as well as by a similar project started at San Diego State University.

All three large-scale projects—Study Abroad for Global Engagement, the Georgetown Consortium, and GLOSSARI—addressed two complementary concerns in study abroad. First was providing program directors with evidence-based reasons for adopting particular practices. For example, it has been largely a matter of faith that host national instructors provide for more profound cultural immersion than do home campus instructors *(Engle & Engle, 2003)*, but does instructor nationality make a documentable

difference in learning outcomes? Data for these sorts of questions were lacking. The second concern was providing evidence to skeptical stakeholders of the value that study abroad adds to learning in higher education. This question invites a broad conception of who those stakeholders for international education may be. This group can include legislators and federal education officials who are urged to increase financial aid for study abroad. It includes college administrators and even fellow faculty members who might need to adopt more appropriate calculations for adjusting faculty teaching loads or to adjust course requirements to make it easier for STEM (science, technology, engineering, and mathematics) majors to go abroad, for instance. Stakeholders also include parents (as well as students) who need to be convinced that studying abroad is as rigorous and rewarding as on-campus study.

Based on extensive work developing, implementing, and assessing GLOSSARI, but also grounded in the outcomes and findings of the other two federally funded large-scale education abroad learning outcomes projects *(Paige et al., 2009; Van deBerg et al., 2006)*, this essay proposes that the same sorts of decisions made in these study abroad research projects can be productively applied to international service-learning, even when the exact questions (e.g., features of program design or practice, institutional outcome variables of interest, etc.) may not be identical to those for study abroad. The following five practices seem likely to be fruitful for enhancing the practice and effectiveness of learning outcomes assessment in international service-learning. Although some of these recommendations have previously been suggested *(cf. Tonkin, 2011)*, the study abroad evaluation studies help demonstrate ways in which they can be concretely implemented, modeling possibilities for international service-learning.

Findings

Recommendation: Emphasize Outcomes Pertaining to Student Learning

Attitudinal, dispositional, and developmental outcomes like world-mindedness or cultural relativism are key values for education abroad. For many program directors, witnessing students' empathic responses to another culture is the big payoff. However, the meaning of these constructs is often abstruse, and in practice interpretations are tied closely to the particular instruments used to measure them *(Eyler, 2011)*.

As just one case in point, the Intercultural Development Inventory *(Hammer, Bennett, & Wiseman, 2003)* is quite commonly taken as a general measure of intercultural competence in evaluating international education *(e.g., Georgia Institute of Technology, 2010)*. Yet it was developed to index a particular model of development in ethnorelativism. Adopting high levels of ethnorelativism may or may not be desirable for college students; one can more confidently assert that understanding the nature of ethnorelativism is an appropriate college learning outcome *(Sutton et al., 2007)*. Thus, the International Learning Outcomes instrument developed by GLOSSARI asks students to self-report on such statements as, "When interacting in a foreign country, I know when it is to my advantage to take risks." It does not, however, ask students if they do take appropriate risks when interacting.

A special case can be made for including at least some "hard" indicators among the mix of learning outcome measures. Certain disciplines, for instance, may offer relatively standardized ways to demonstrate subject matter mastery, such as the Russian language proficiency battery promulgated by American Association of Teachers of Russian *(Davidson, 2007)*. However, since other outcome measures can also do double duty for accreditation and supporting institutional (as opposed to solely program- or course-specific) goals, the most convincing indicators of learning may derive from institutional-level data *(Volkwein, 2011)*. Measures such as students' graduation rates, overall grade point averages, pass rates on professional certification and accreditation processes, and career attainment for alumni, as well as impacts on general-education outcomes such as critical thinking or moral reasoning, make a strong and readily understandable case for the variety of international service-learning stakeholders. Phase 4 of GLOSSARI attracted substantial interest because of its findings of a positive effect of study abroad on college completion rates, for example *(O'Rear et al., 2011)*.

Recommendation: Employ Multiple Sources and Methods for Data Collection

By utilizing multiple methods, researchers can triangulate their conclusions and examine a variety of learning outcomes. In international service-learning, both quantitative *(Bringle et al., 2011)* and qualitative *(Kiely & Hartman, 2011)* research traditions, methodologies, and instruments have been recommended, but putting these into place for larger-scale evaluation can be challenging.

GLOSSARI developed and administered one self-report survey and one direct test of learning to several thousand current study abroad participants. It also administered standardized measures of intercultural sensitivity and development and of critical thinking. These included the Cross Cultural Adaptability Inventory *(Kelley & Meyers, 1999)*, the Intercultural Development Inventory *(Hammer et al., 2003)*, and the California Critical Thinking Test *(Faccione, 2000)*. A handful of classes participated in a mixed methods (qualitative and quantitative) study of student learning artifacts such as examinations, essays, and group multimedia projects. Finally, GLOSSARI undertook a large-scale retrospective examination of institutional data (time to graduation, grade point average, change of major).

In a similar vein, the Georgetown Consortium, because of its special interest in learning additional languages, administered hundreds of structured oral proficiency interviews and also administered written tests and measures of intercultural development *(Van deBerg et al., 2006)*. The Study Abroad for Global Engagement project focused on long-term impact on study abroad alumni *(Paige et al., 2009)*. In addition to sending out several thousand mail surveys, it also conducted scores of intensive interviews. While the specifics of the methodology will be constrained and guided by the institutional or program variables of interest, consciously incorporating multiple methods into international service-learning research will help ensure robust, triangulated conclusions for the kinds of questions identified as salient for the field by researchers such as Tonkin *(2011)*.

Recommendation: Accumulate Data from Diverse Institutions and Programs

In a similar vein, assessing learning outcomes across a broad range of institutions and programs holds promise for producing more robust, generalizable, and replicable findings. As a system-wide initiative, for instance, GLOSSARI collected data from more than 30 colleges and universities. These included large Research I institutions, 4-year liberal-arts colleges, community colleges, and three historically Black institutions. This substantial number of institutions fielded scores of study abroad programs each year representing all regions, disciplines, and durations. For more labor-intensive data collecting procedures (pre- and post-study abroad surveys, for example), a smaller representative set of institutions (and the programs housed therein) participated.

In addition to the improved generalizability of such findings, this diversity also allows fine-grained comparisons among, say, students studying in Asia versus those studying in the South Pacific. Within the small subset of courses that international service-learning represents, a growing number of opportunities remain for developing research projects investigating programs from multiple universities in the same host country (e.g., programs in Costa Rica or South Africa), or even service-learning programs that engage students in similar service experiences in more than one international setting.

Recommendation: Compile Credible Comparison Groups

Much research in international education simply describes student dispositions upon program completion and then attributes those dispositions to studying abroad (or to service-learning). More convincing research designs should at least compare students' performance after studying abroad with their performance just prior to studying abroad. But how does one know that increments in, say, Russian pronunciation accrued by students spending a month in St. Petersburg, Russia, are significantly higher than those achieved by students taking an intensive Russian class in St. Petersburg, Florida? As all students mature and learn across time (hopefully), it is necessary to compare students studying abroad with their peers who did not.

The trick, however, is determining just who counts as a "peer" to usefully compare with a study abroad (or international service-learning) participant. Students who study abroad are a select group on dimensions such as choice of major, socioeconomic status, grade point average, prior cosmopolitanism, and progress toward degree; similar distinctions could presumably be made for those who choose to take part in service activities through their international experiences. Skeptics of study abroad often point to the distinctiveness of its enrollees to refute claims about the value of international education. To ascertain the value-added dimension of studying abroad or of international service-learning, research designs must minimize these "confounding" factors as possible explanations for learning outcomes. Thus, for example, one ought to compare students who studied abroad as juniors with students who were juniors at the same time, but took the path of on-campus classes. Studies that simply compare graduation rates between first-year cohorts who studied abroad and those who did not are unconvincing. How meaningful is it to compare students who have

survived 2 years of college with those just starting out? (Across all U.S. institutions, 33% of college students stop at the end of their first year.)

In conducting GLOSSARI, immense effort was expended in compiling credible comparison groups. These carefully constructed comparison groups sometimes resulted in smaller effects for studying abroad than are shown by other studies, but the causal arguments these comparisons warrant are more compelling. In the international service-learning context, possible comparisons could be made not only between students who participate in a service-learning experience in a domestic or international version of the same course, but also between students in a given course taught abroad with and without service components *(cf. Eyler, 2011)*.

Recommendation: Accumulate Large Samples

Although many valuable insights are available only through small-*n*, intensive qualitative analyses, there is power (literally) in large sample sizes. With an initial sample size of more than 30,000 for some analyses, GLOSSARI was able to drill down and draw meaningful conclusions about participation rates (e.g., nearly 10% of the total were graduate students) and types (e.g., apparent heritage motivations among Asian Americans, who were disproportionately represented at Asian sites). Outcomes could be disaggregated for other subgroups (e.g., the improvement in graduation rate was especially pronounced for African Americans). With such a large sample size, a substantial number of GLOSSARI participants were financial aid recipients, and it was even possible to determine the effect of unmet financial need on program participation (e.g., even after statistically eliminating unmet financial need, African Americans were substantially less likely to study abroad than their White counterparts).

Although gathering a large set of data for international service-learning in particular is a challenging task, many of the fundamental research questions posed by Tonkin *(2011)* and others for the field could more readily and credibly be investigated through large-sample analysis. International service-learning courses are often quite small, but many programs are repeated annually; thus, collecting multiple years of data as well as collaborating across programs and institutions, as mentioned above, can help boost the explanatory power of such analyses.

Conclusion

In setting out a research agenda for international service-learning, Tonkin *(2011)* reiterates the importance of investigating outcomes beyond student satisfaction with programs, especially focusing on longer-term variables of institutional interest:

> Outcomes assessment is crucially important if study abroad and international experiences are to find a firm foothold in the curriculum and if curricular designers are to make wise decisions that earn the support of the executive leadership of the campus. Research needs to determine how [international service-learning] contributes to a student's readiness and preparedness to learn after returning to the home campus. *(p. 208)*

Bringle et al. *(2011, pp. 285-287)* argue that future "good research" on international service-learning should be guided by theory, involve clearly defined constructs, account for differences among groups, use psychometrically defensible measures with multiple indicators, use multiple methods with converging results across different methods, apply designs that result in confidence in the conclusions reached, and have "implications for teaching and learning in general." Achieving these goals for assessing international service-learning will be challenging. Attention to student academic outcomes, to using clearly defined and psychometrically defensible outcome measures, and to building a compelling evidence-based case for international service-learning stakeholders requires systemic approaches to evaluation. To attain that systematicity in building large databases across programs, institutions, and research methods, international service-learning educators will need an organizational hub and a commitment to collaboration among program administrators at various sites. Obtaining that commitment is difficult not only because good learning outcomes research requires an infusion of resources, but also because it requires courage to voluntarily submit one's program to an evaluation regime.

One sign that the time is ripe for this kind of concerted effort to anchor international service-learning in hard evidence of academic learning outcomes lies in the spate of recent influential critiques questioning the value added by higher education *(Bok, 2006; Keeling & Hersch, 2011)*. Many of these critiques center on a perceived loss of focus on the core mission of academic learning. Arum and Roksa *(2011)*, for example, contend that a contemporary college education

in the United States typically fails to impart critical thinking skills such as the capacity to make a compelling argument. In response to such critiques, consortia such as the New Leadership Alliance for Student Learning and Accountability *(2012)* are promoting systematic learning outcomes assessment much as this essay proposes for international service-learning. The recent precedent of systemic, large-scale outcomes assessment in the domain of conventional study abroad may provide models that the field of international service-learning can draw upon, and then improve.

References

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Arum, R., Roksa, J., & Cho, E. (2011). *Improving undergraduate learning: Findings and policy recommendations from the SSRC-CLA Longitudinal Project*. New York, NY: Social Science Research Council. Retrieved July 1, 2012, from http://eric.ed.gov/PDFS/ED514983.pdf

Banta, T. W. (2002). Characteristics of effective outcomes assessment: Foundations and examples. In T. W. Banta & Associates, *Building a scholarship of assessment* (pp. 261–283). San Francisco, CA: Jossey-Bass.

Bentley, R., & Ellison, K. J. (2007). Increasing cultural competence in nursing through international service-learning experiences. *Nurse Educator, 32*(5), 207–211, doi:10.1097/01.NNE.0000289385.14007.b4

Bers, T., & Swing, R. L. (2010). Championing the assessment of learning: The role of top leaders. In G. L. Karmer & R. L. Swing (Eds.), *Higher education assessment: Leadership matters* (pp. 3-26). Lanham, MD: Rowan & Littlefield.

Bok, D. (2006). *Our underachieving colleges*. Princeton, NJ: Princeton University Press.

Brewer, E., & Cunningham, K. (2009). Capacity building for study abroad integration: The institution and the faculty. In E. Brewer & K. Cunningham (Eds.), *Integrating study abroad into the undergraduate curriculum: Theory and practice across the disciplines* (pp. 209–221). Sterling, VA: Stylus Publishing.

Bringle, R. G., & Hatcher, J. A. (2011). International service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 3–28; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Bringle, R. G., Hatcher, J. A., & Williams, M. J. (2011). Quantitative approaches to research on international service learning: Design, measurement, and theory. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 275–290; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Burak, P. A., & Hoffa, W. W. (Eds.). (2001). *Crisis management in a cross-cultural setting*. Washington, DC: NAFSA: Association of International Educators.

Conway, J. M., Amel, E. L., & Gerwien, D. P. (2009). Teaching and learning in the social context: A meta-analysis of service learning's effects on academic, personal, social, and citizenship outcomes. *Teaching of Psychology, 36*, 233–245.

Crabtree, R. D. (2008). Theoretical foundations for international service-learning. *Michigan Journal of Community Service Learning, 15*(1), 18–36.

Cruz, N., & Giles, D. E. (2000). Where's the community in service-learning research? *Michigan Journal of Community Service Learning*, Fall (special issue), 28–34.

Davidson, D. E. (2007). Study abroad and outcomes measurements: The case of Russian. *The Modern Language Journal, 91*(2), 276–280.

DePaul, S., & Hoffa, W. W. (Eds.). (2010). *A history of study abroad: 1965–present*. Carlisle, PA: Forum on Education Abroad.

Engle, L., & Engle, J. (2003). Study abroad levels: Toward a classification of program types. *Frontiers: The Interdisciplinary Journal of Study Abroad, 9*, 1–20.

Eyler, J. (2011). What international service learning research can learn from research on service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 225–241; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Faccione, P. A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic, 20*(1), 61–84.

Georgia Institute of Technology. (2010). Strengthening the global competence and research experiences of undergraduate students: Impact report of Georgia Tech's Quality Improvement Plan. Retrieved from http://www.accreditation.gatech.edu/wp-content/uploads/2010/11/11_GT__2010_QEP-Annual_11_10_2010_WEB-FINAL__.pdf

Gonyea, R. M. (2008, November). *The impact of study abroad on senior year engagement*. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Jacksonville, FL. Retrieved from http://cpr.iub.edu/uploads/Gonyea_StudyAbroad.pdf

Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The Intercultural Development Inventory. *International Journal of Intercultural Relations, 27*(4), 421–443.

Hoff, J. G. (2008). Growth and transformation outcomes in international education. In V. Savicki (Ed.), *Developing intercultural competence and transformation: Theory, research, and application in international education* (pp. 53–73). Sterling, VA: Stylus Publishing.

Hovland, K. (2006). *Shared futures: Global learning and liberal education*. Washington, DC: Association of American Colleges and Universities.

Jones, S. G., & Steinberg, K. S. (2011). An analysis of international service learning programs. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 89–112; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Keeling, R. P., & Hersh, R. H. (2011). *We're losing our minds: Rethinking American higher education*. New York, NY: Palgrave Macmillan.

Kelley, C., & Meyers, J. (1999). The Cross-Cultural Adaptability Inventory. In S. Fowler & M. Mumford (Eds.), *Intercultural sourcebook: Cross-cultural training methods* (Vol. 2, pp. 53–70). Yarmouth, ME: Intercultural Press.

Kiely, R. (2004). A chameleon with a complex: Searching for transformation in international service-learning. *Michigan Journal of Community Service Learning, 10*(2), 5-20.

Kiely, R. (2011). What international service learning research can learn from research on international learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 243–273; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Kiely, R., & Hartman, E. (2011). Qualitative research methodology and international service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 291–317; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Knutson Miller, K., & Gonzalez, A. M. (2011). Challenges and rewards associated with service-learning in international contexts: Pre-service teacher outcomes. *Journal of Research on Service Learning and Teacher Education, 1*(2), 56–68. Retrieved from http://www.jrslte.com/documents/miller.pdf

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges and Universities.

Marsh, P. A. (2007). What is known about student learning outcomes and how does it relate to the scholarship of teaching and learning? *International Journal for the Scholarship of Teaching and Learning, 1*(2), 1–12. Retrieved from http://academics.georgiasouthern.edu/ijsotl/v1n2/essays/marsh/Essay_Marsh.pdf

Martinez-Mier, E. A., Soto-Rojas, A. E., Stelzner, S. M., Lorant, D. E., Riner, M. E., & Yoder, K. M. (2011). An international, multidisciplinary, service-learning program: An option in the dental school curriculum. *Education for Health, 24*(1), 1–12. Retrieved from http://www.educationforhealth.net/publishedarticles/article_print_259.pdf

Montrose, L. (2002). International study and experiential learning: The academic context. *Frontiers: The Interdisciplinary Journal of Study Abroad, 8*, 1–15.

New Leadership Alliance for Student Learning and Accountability. (2012). *Committing to quality: Guidelines for assessment and accountability in higher education*. Washington, DC: Author. Retrieved from http://www.newleadershipalliance.org/images/uploads/committing%20to%20quality.pdf

Nisbett, N., Tannenbaum, S. C., & Smither, B. (2009). Advancing service learning through program evaluation. In T. Kelshaw, F. Lazarus, & J. Miner (Eds.), *Partnerships for service-learning: Impacts on communities and schools* (pp. 103–128). San Francisco, CA: Jossey-Bass.

O'Rear, I., Sutton, R. C., & Rubin, D. L. (2011). The effect of study abroad on college completion in a public university system. Retrieved from http://www.glossari.uga.edu

Pagano, M., & Roselle, L. (2009). Beyond reflection through an academic lens: Refraction and international experiential education. *Frontiers: The Interdisciplinary Journal of Study Abroad, 18*, 217–229.

Paige, R. M., & Stallman, E. R. (2007). Using instruments in education abroad outcomes research. In M. C. Bolen (Ed.), *A guide to outcomes assessment in education abroad* (pp. 137–162). Carlisle, PA: The Forum on Education Abroad.

Paige, R. M., Fry, G. W., Stallman, E. M., Josić, J., & Jon, J.-E. (2009). Study abroad for global engagement: The long-term impact of mobility experiences. *Intercultural Education, 20*(Supplement 1), S29–S44.

Plater, W. M. (2011). The context for international service learning: An invisible revolution is underway. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 29–56; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Pusch, M. D., & Merrill, M. (2008). Reflection, reciprocity, responsibility, and committed relativism: Intercultural development through international service learning. In V. Savicki (Ed.), *Developing intercultural competence and transformation: Theory, research, and application in international education* (pp. 297–321). Sterling, VA: Stylus Publishing.

Rama, D. V., Ravenscroft, S. P., Wolcott, S. K., & Zlotkowski, E. (2000). Service-learning outcomes: Guidelines for educators and researchers. *Issues in Accounting Education, 15*(4), 657–692.

Robinette, M. S., & Noblet, R. (2009). Service-learning in entomology: Teaching, research, and outreach domestically and abroad. *Journal of Higher Education Outreach and Engagement, 13*(4), 135–153.

Sideli, K. (2001, Spring). Outcomes assessment and study abroad programs: Commentary on the results of a SECUSSA/IIE electronic sampling. *International Educator, 10*, 30.

Steinberg, M. (2002). "Involve me and I will understand": Academic quality in experiential programs abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad, 8*, 207–229.

Steinberg, M. (2007). The place of outcomes assessment in higher education today and the implications for education abroad. In M. C. Bolen (Ed.), *A guide to outcomes assessment in education abroad* (pp. 7–22). Carlisle, PA: The Forum on Education Abroad.

Sutton, R. C., Miller, A. N., & Rubin, D. L. (2007). Research design in assessing learning outcomes of education abroad programs. In M. C. Bolen (Ed.), *A guide to outcomes assessment in education abroad* (pp. 23–60). Carlisle, PA: The Forum on Education Abroad.

Sutton, R. C., & Rubin, D. L. (2004). The GLOSSARI project: Initial findings from a system-wide research initiative on study abroad learning outcomes. *Frontiers: The Interdisciplinary Journal of Study Abroad, 10,* 65–82.

Sutton, S. B. (2011). Service learning as local learning: The importance of context. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 125–144; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Tonkin, H. (2011). A research agenda for international service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 191–224; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Tonkin, H., & Quiroga, D. (2004). A qualitative approach to the assessment of international service learning. *Frontiers: The Interdisciplinary Journal of Study Abroad, 10*, 163–182.

Van deBerg, M., Balcum, A., Scheid, M., & Whalen, B. J. (2006). The Georgetown University Consortium Project: A report at the half-way mark. *Frontiers: The Interdisciplinary Journal of Study Abroad, 10*, 101–116.

Volkwein, J. F. (2011, September). *Gaining ground: The role of institutional research in assessing student outcomes and demonstrating institutional effectiveness* (NILOA Occasional Paper, No. 11). Urbana, IL: University for Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Retrieved from http://www.learningoutcomeassessment.org/occasionalpapereleven.htm

Wells, C., Warchal, J., Ruiz, A., & Chapdelaine, A. (2011). Ethical issues in research on international service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 319–343; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Whitney, B. C., & Clayton, P. H. (2011). Research on and through reflection in international service learning. In R. G. Bringle, J. A. Hatcher, & S. G. Jones (Eds.), *International service learning: Conceptual frameworks and research* (pp. 145–187; IUPUI Series on Service Learning Research, Vol. 1). Sterling, VA: Stylus.

Wood, C. A., Banks, S., Galiardi, S., Koehn, J., & Schroeder, K. (2011). Community impacts of international service-learning and study abroad: An analysis of focus groups with program leaders. *Partnerships: A Journal of Service-Learning & Civic Engagement, 2*(1), 1–23. Retrieved from http://geo.appstate.edu/sites/default/files/communityimpacts.pdf

About the Authors

Donald L. Rubin is emeritus professor of speech communication, of language and literacy education, and of linguistics at the University of Georgia. He is also senior researcher at the University's Center for Health and Risk Communication. Rubin teaches mainly in areas of health communication, intercultural communication, language assessment, and English for Specific Purposes. His current areas of research activity include health literacy and health message design, student learning outcomes accruing from studying abroad, and attitudinal responses to non-native speakers of English. Rubin earned his M.A. in speech communication from the University of Minnesota and his Ph.D. in education from the University of Minnesota.

Paul H. Matthews is the assistant director of the University of Georgia Office of Service-Learning, helping coordinate faculty development, track institutional reporting of service-learning and community engagement, conduct research and evaluation, coordinate service-learning initiatives, and teach service-learning coursework. His research interests include outcomes and processes of service-learning, education of second language learners, tutoring, and faculty development. Matthews was a Fulbright fellow to the University of Passau (Germany), earned his M.A. in Latin American studies from the University of Texas at Austin, and his Ph.D. in language education from the University of Georgia.