

7 “It Changed My Life”: Strategies for Assessing Student Learning

By Donald Rubin and Michael Tarrant

For an international educator, there’s nothing quite like the excitement you get when one of your students on a study abroad program suddenly experiences an epiphany about his or her worldview or role as a citizen in today’s global society. As one of our students remarked, while participating in a sustainability program in Australia’s Northern Territory, “So, somehow the past is not really past for [the Aborigines]. The past is really going on right now, right? That legend from 2,000 years ago is happening at this moment?!”

Most study abroad program directors and administrators probably entered the field specifically for those thrilling moments when students experience such transformations on their programs. Not many, however, become involved in education abroad primarily because they love assessment and program evaluation. For many study abroad professionals, assessment is a chore, an afterthought, a painful task they hope will be dispatched without too much investment. Further, for those program directors who work to compress lots of learning into a short-term course, adding on assessment may seem like an especially low priority.

However, such a lackadaisical attitude toward assessment is rather like getting enjoyment from sailing a yacht without caring much about navigation. You can’t be sure where your students’ knowledge is invested, nor where it is heading, without proper assessment. Assessment is how educators plan their routes, make midcourse adjustments as necessary, and help students find their way into port at the end of the journey. Therefore, it is only by deploying careful and innovative assessment that program directors and administrators can reliably navigate their students toward that epiphany-like learning that we (and they) so highly prize. In particular, we believe that best practices in study abroad assessment call for:

- measuring learning outcomes rather than just student satisfaction;
- tailoring outcome measures to match course objectives and course content;
- deploying a pretest-posttest comparison group design to identify the value added by studying abroad; and

- compiling comparable outcome measures across years of program implementation as part of a process of continuous quality improvement.

Study abroad assessment is also integral to the scholarship of teaching and learning. Assessment is one way of making one's teaching public and subject to peer review and, in that sense, it demands a degree of courage from practitioners. For those who pursue it, program assessment challenges us to be inquisitive and experimental about our teaching. In short, assessment need not be viewed as a burden; instead, it ought to be viewed as enriching our careers as scholars.

Student Satisfaction Surveys

Very often, the approach to assessment treats study abroad primarily as a consumer product. Consistent with this “student-as-customer” approach, program success is frequently judged by just two blunt evaluator metrics: enrollment numbers and customer satisfaction. These metrics are not unimportant, of course. (Indeed, they arguably reflect the realities of higher education administration today.) Institutions and economics impose enrollment thresholds, below which programs are no longer viable. Year-to-year trends in declining (or increasing) enrollment need to be understood. Declining enrollment that is not attributable to economic, competitive, or perhaps, security factors may be a warning sign that students and/or their parents no longer perceive value in the program. Interviews with recent program alumni, or with current students who initially expressed interest but eventually did not enroll, can yield valuable insights about student and family decisionmaking processes.

It can be difficult to interpret scores on satisfaction surveys, however. In the broader world of consumer service ratings, providers these days often enjoin us, “Anything lower than the top score means that my service was unsatisfactory.” On a satisfaction scale where scores can range from 1–5, lots of 1s and 2s ought to be treated as red flags. However, it is often unclear, for example, what an average score of 3.5 means for an item such as, “In-country advising was helpful,” or what it means if “Accommodations met students’ basic expectations.” To address these issues, some best practices for satisfaction surveys include the following recommendations:

- Construct questionnaires in ways that encourage students to provide focused free-response (qualitative) explanations along with numeric ratings.
- Provide students with sufficient in-program time to complete surveys thoughtfully. Or, if students must complete surveys after their return, link survey response to some incentive, such as early receipt of course grades, to ensure a useful response rate.
- Pay attention to year-to-year fluctuations in ratings, especially fluctuations that correspond to changes in the program curriculum or itinerary.

When study abroad programs face mandates to generate assessment data, the default strategy is generally to administer satisfaction surveys to participants. Certainly it is worth

knowing if students perceived their courses to be stimulating, or whether they would have preferred more comfortable lodging or more unstructured time on their own. Some examples of student satisfaction surveys can be found on the website for [Butler University's Institute for Study Abroad](http://www.ifsabutler.org/for-faculty-advisors/student-satisfaction-surveys.html) (<http://www.ifsabutler.org/for-faculty-advisors/student-satisfaction-surveys.html>).

Assessing Learning Outcomes: Why Do It?

Neither program popularity nor student satisfaction ratings tell you whether students have actually learned anything from studying abroad, however. Current trends in higher education make accreditation and internal resource allocation dependent on the development of student learning outcomes. The [National Institute for Learning Outcomes Assessment's website](http://www.learningoutcomesassessment.org/TFComponentSLOS.htm) (see <http://www.learningoutcomesassessment.org/TFComponentSLOS.htm>) offers examples of relevant policies from several institutions regarding the importance of assessing student learning outcomes to the overall academic mission of the university. Given that study abroad is becoming more integrated into the home curriculum, the field of education abroad has also increasingly turned its attention to assessing student learning outcomes (Deardorff 2015). Both the [Forum on Education Abroad](#) and [NAFSA: Association for International Educators](#) provide useful resource links (including a bibliography) on study abroad assessment and evaluation. Collectively, these resources comprise a fairly comprehensive overview of the state of knowledge, including methods and findings, regarding student learning assessment. Moreover, these resources offer innovative approaches that can be applied to short-term study abroad programs.

It is costly, in a number of respects, to send students abroad to learn. It is therefore reasonable to expect education abroad programs to document that students are learning something abroad that they are less likely to learn on their home campuses. Program directors and education abroad professionals need to demonstrate that students can learn something more (or different) when, for example, they study Renaissance art in four weeks in Florence, Italy, as compared with studying it for 15 weeks in Florence, South Carolina (Sutton, Miller, and Rubin 2007). It is not sufficient for a program director to fervently believe with all his or her heart that the Italy course bestows much richer appreciation. It is not even sufficient for that program director to collect anecdotal accounts from students attesting to their bedazzlement. Learning outcomes assessment is more systematic, and potentially more useful, than student anecdotes or faculty beliefs.

One way to visualize study abroad learning outcomes assessment is through the example of world languages education, a subject that is increasingly being taken by students on short-term programs (Kinging 2011). Outcomes of students' interest include conventional language skills such as oral proficiency (measured via semistructured oral proficiency interviews), listening comprehension, and grammatical accuracy. In recent years, however, program directors have also examined their students' gains in pragmatics, such as how to make polite requests, and in communication competencies, for example how to be persuasive.

Interestingly, the cumulative body of research on world language learning sometimes flies in the face of many educators' articles of faith. Studying abroad might not always lead

to superior world languages proficiency, relative to studying on campus, particularly if grammatical accuracy is the aim. Similarly, living with a host family may not always improve students' oral proficiency compared with those living in dormitories; surprisingly, the latter often end up speaking in the target language more often (Rivers 1998). In fact, careful assessment of home stay immersion outcomes has been crucial for developing structured interventions with host families to make home stays more satisfying and productive.

For language instructors, unexamined “received wisdom” about where language gains will come from can sometimes be counterproductive. For those working in nonlanguage-based study abroad, the lesson can be similar. “Received wisdom” that asserts that longer duration residential programs always yield superior outcomes to short-term travel programs turns out to be, at best, inconsistently supported by evidence. It is not enough, therefore, to presume that a particular type of program structure, itinerary, or housing configuration will automatically lead to the kind of learning you envision. Intentional connections among learning outcomes, program details, and assessment measures are necessary.

Aligning Measures of Learning with Course Objectives and Learning Experiences

Assessing the Academic Core of the Program

There are several different instruments and methods that you can use to assess learning in an education abroad context. How do you determine the assessment tool that works best for your particular education abroad program? While it may be tempting for a program administrator or faculty director to select a particular assessment instrument or technique simply because it is well known, or because it is logistically manageable, the first source to consult in order to answer that question ought to be the course syllabus. Perhaps you have noticed in the assessment literature that a number of programs have adopted an instrument that measures, say, intercultural competence. Intercultural competence is certainly a laudable outcome for a study abroad program. However, if your course is focused on the marine biology of mangrove forests in Australia, and includes no activities or objectives promoting intercultural competence, then that instrument would not be appropriate as an assessment tool. Instead, look to the learning objectives in the syllabus to guide you in choosing or designing an assessment instrument.

In fact, the exercise of choosing an assessment protocol for a particular program might very well inspire you to revise the course objectives. You might come to realize, for example, that the course objectives need to be more specific in order to be more measurable. Or you might realize that the objectives fail to adequately exploit the unique attributes of your international site. For the marine biology program exploring mangrove forests in Australia, the program director may decide to revise the course objectives and experiences if he or she determines that intercultural competency should play a more important role in the course. For example, the director could include a goal relating to intercultural communication and then include meetings with indigenous landowners as a way for students to develop or practice such skills. Once those objectives and experiences are integrated into the program,

then it makes sense to assess students' growth in intercultural competence within this marine biology course.

The authors of this chapter teach a study abroad course on the theme of “progress,” but have found it difficult to assess whether students had learned about progress in the ways intended (Tarrant, Rubin, and Stoner 2014). One method that we devised taught students techniques for digital storytelling by having them create three-minute digital stories to reflect their understanding of progress and how it may have changed as a result of their study abroad experience. Several of these digital stories about the concept of “progress” are available on the University of Georgia's [Discover Abroad](http://www.discoverabroad.uga.edu/index.php/media/digital-stories/) website (<http://www.discoverabroad.uga.edu/index.php/media/digital-stories/>). As the course leaders, we then developed a formal rubric to rate the digital stories along nine specific dimensions grouped under (1) critical thinking, (2) developing evidence, and (3) cultural understanding. We discovered that students scored lowest on the dimension related to cultural curiosity—comparing how their notions of progress might be viewed through other cultural lenses. After analyzing the assessment results, we identified that that particular aspect of our objective needed more deliberate instructional intervention.

Learning Outcomes in an Institutional Context

One recent thread of study abroad assessment aligns with the notion that study abroad is a “high-impact” experience that enhances student engagement and, thus, improves indicators of general academic success such as GPA and timely college completion (Rubin, Sutton, O’Rear, Rhodes, and Raby 2014). Demonstrating the impact of a study abroad program obviously demands a long-term commitment to data collection. It also requires collaboration with your institutional research office in order to connect the assessment of individual program goals and student learning in the context of broader university initiatives. The effort is worth the investment, however. If you can demonstrate that participation in your study abroad program improves students’ graduation rates, for example, your upper-level administration will be eager to hear the news. Efforts such as the federally funded Georgia Learning Outcomes of Students Studying Abroad Research Initiative (GLOSSARI) and California Community College Student Outcomes Abroad Research (CCC-SOAR) projects (see www.glossary.uga.edu and <http://gloaled.us/ccsoar/>), in which constituent institutions combine their information into a systemwide database, can be particularly powerful tools in influencing broader educational policymaking (Stearns 2009). Similarly, linking assessment efforts to institutional (as well as the international education office’s) missions should be a key objective. For example, demonstrating the effect of study abroad on global citizenry helps support the more than 40 percent of U.S. universities that reference globalization in their strategic plans (Stearns 2009).

Finally, assessment and program design should go hand in hand. Without direct translation of assessment findings into practice—for example, advising prospective study abroad students about program selection on the basis of anticipated learning outcomes, rather than simply country of destination—the powerful benefits of studying abroad will not be realized.

Tools for Assessing Outcomes in Education Abroad

Program directors and administrators can choose from a wide palette of established instruments, as well as unique observational strategies, to assess student learning outcomes (Paige and Stallman 2007). The following section is a sampling of some of the assessment approaches that we find to be the most helpful.

Course Content Knowledge Tests

Most courses seek to impart some content domain of information or skill that can be subjected to meaningful examination. The methods that are used to judge student achievement in courses on campus can be adapted to short-term study abroad. For example, in our University of Georgia programs that are focused on sustainability, we have administered a nine-item short-answer test that assesses students' knowledge of sustainability. (Question: What is included when corporations report their triple bottom line? Answer: Environmental, social, and financial performance.) Content knowledge can be measured reliably using open-ended or essay tests, as well as short-answer or multiple choice tests. (See Clay 2001 for a succinct guide to creating classroom tests.) Some faculty may believe that formal content examinations are not appropriate in short-term, experiential, field-based classes. We argue, to the contrary, that if course objectives include mastery over some knowledge domain, then formal testing—using open-ended essays questions or multiple choice questions, recall questions as well as synthesis questions—is quite appropriate in short-term study abroad.

“Can-Do” Self-Reported Achievement

In language-based study abroad, oral proficiency interviews (OPIs) administered and rated by personnel trained to exacting standards are generally considered the most convincing evidence of achievement. Similarly, in (nonlanguage) study abroad, generally, oral interviews and personal oratories (such as digital stories) are compelling ways to assess students' understanding of material. However, OPIs and personal interviews are costly in terms of time and money. Therefore, carefully developed and validated measures in which students self-assess their understanding can be practical alternatives. The downside? Sometimes students' self-assessments of their proficiency or knowledge can actually decline from inflated predeparture views once they are challenged by actual immersion in the local environment.

Global Citizenship

Ultimately, the mission of study abroad is to foster a global citizenry. And toward that end, we enjoin program directors and administrators to include global citizenship as an objective to be taught and assessed for every study abroad program. Broadly, we identify at least three key components of global citizenship: (1) awareness, (2) responsibility, and (3) engagement in the interdependency of all living things on the planet. Measures of global citizenship attitudes, beliefs, and behavioral dispositions are as varied as the definitions of

the construct. Such assessment tools include the Global Perspectives Inventory (Braskamp, Braskamp, and Merrill 2009) and the Global Citizenship Scale (Morais and Ogden 2011).

Cultural Sensitivity, Tolerance, Ethnorelativism

Many programs take as a measure of success the degree to which students’ progress away from ethnocentric value judgments of other cultures and move toward an acceptance of cultural pluralism. Several popular measures of such outcomes are used with frequency by sending institutions (see table on page XX). One such measure is the [Intercultural Development Inventory](#) (IDI), an assessment tool that measures intercultural sensitivity and is widely used across many schools and organizations. As with all assessment tools, we urge directors and administrators to examine carefully the content domain of each instrument to determine whether it matches course and overall program or institutional objectives.

Intercultural Competence

Intercultural competence depends on cultural sensitivity, but it is not the same construct. Competence requires knowledge and the capacity to actually function effectively in another culture. Intercultural competence can be discerned via self-report scales such as the [GLOSSARI International Learning Outcomes](#) (ILO) scale. The ILO contains Likert-type scales such as, “When interacting in a foreign country, I know how to talk my way out of difficult situations,” and “I know how to buy toothpaste or a can opener in a foreign country.”

A Sampling of Instruments Measuring Intercultural Competence	
Intercultural Development Inventory (IDI) idiinventory.com	A 50-item questionnaire to assess intercultural sensitivity and competence (with scores placed along a six-stage developmental scale from denial to integration). The IDI has been used with more than 10,000 individuals, in 30 countries and more than 170 universities.
Global Perspectives Inventory (GPI) www.gpi.hs.iastate.edu	GPI measures a global view of student learning and development across three broad dimensions: cognitive (how do I know), intrapersonal (who am I), and interpersonal (how do I relate to others). The GPI is endorsed by the Association of American Colleges & Universities and widely adopted by institutions nationally (used by more than 42,000 students).
Beliefs, Events, Values Inventory (BEVI) www.thebevi.com	An extensive scale of almost 500 questions designed to assess students’ belief orientations and worldviews and their receptivity to change following experiential learning and events. The scale was piloted with approximately 2,000 undergraduates internationally and has undergone continued refinement over the past decade.
National Survey on Student Engagement (NSSE) http://nsse.indiana.edu	A measure of active involvement in learning focusing on (1) engagement indicators (the time and effort students put into their studies) and (2) high-impact practices (strategies used by institutions to increase involvement, such as service-learning and research projects). The NSSE forms part of the <i>College Student Report</i> that has been administered to 5.5 million students since 2000.

(continued)

A Sampling of Instruments Measuring Intercultural Competence (continued)

Global Citizenship	A self-report scale of 47 items to assess three dimensions of global citizenry: social responsibility, global civic engagement, and global competence. A theoretically grounded scale, comparable to the GPI.
International Learning Outcomes (ILO) glossari.uga.edu	The Georgia Learning Outcomes of Students Studying Abroad Research Initiative (GLOSSARI) ILO survey links study abroad to academic accountability and indicators of student learning.
Global Engagement Measurement Scale (GEMS)	This instrument developed in-house at the University of Delaware measures change in various areas of intercultural competence (cultural engagement, tolerance for ambiguity, knowledge of host site, diversity acceptance), as well as resilience.
Other Scales	An overview of other student learning outcomes assessment tools can be found in Roy, Wandschneider, and Steglitz 2014.

Institutions with the resources and expertise may want to consider developing in-house assessment instruments, which have the advantage of being developed for particular institutional purposes and also cost-free administration. On the other hand, developing, piloting, and refining one's own instrument can be time consuming. "Off-the-shelf" questionnaires have typically been tested numerous times and can be more quickly implemented.

Assessing the Added Value of Study Abroad

As a case example, let's say you administer a measure of intercultural competence to your students on the final day of their four-week sojourn studying art history in Florence, Italy. (You have done so because the measure matches one of the learning objectives listed in the course syllabus.) And suppose that the average score is 42.3 out of a possible 50. What will you have learned as a result of this assessment? Unfortunately, not much. Because you did not obtain baseline scores for your students before they went abroad, you don't know if the final-day scores represent a meaningful gain, a plateau, or even a loss. The number 42.3 has no intrinsic meaning; it needs to be compared with something else.

You decide to give the same test to an on-campus art history course. At the end of the semester, those students score an average of 36. Looking good for your study abroad assessment? Not really; common knowledge says that study abroad enrolls "better" and more internationally-minded students. Skeptics will conclude that your Italy students scored higher because they were more competent before they ever boarded the flight to Florence.

Now let's say you did administer this same measure before your students ever left for Italy. It turns out that their average predeparture score was 38.1. With that final-day score of 42.3, it looks like your students gained, on average, 10 percent during the course of their program. How do you know if that 10 percent gain is a lot, or even if it is associated with participating in the program? Inferential statistics will tell you that an increase of 4.2 points could be just a chance event. But even if the predeparture to final-day increment of 4.2 points for the Italy course turns out to be statistically significant, it is still possible that

students who studied art history back on campus might display the same bump (maybe even a bigger one) in intercultural competence. To feel confident that studying in Italy added meaningful value, you need to show that the 4.2-point increment that the Italy students achieved exceeded the increment that their on-campus peers may have also achieved. You must also show that the difference in the size of those increments was not likely due to chance. In short, to obtain convincing evidence of value added, you need to design your assessment as a pretest-posttest comparison group quasiexperiment. The same design principle holds true whether your assessment data are quantitative or qualitative. (The most promising assessment strategies employ mixed quantitative and qualitative methods.)

The pretest-posttest comparison group model for gathering assessment data is a powerful strategy for program evaluation. That power is amplified when assessment is designed not just as a series of annual snapshots, but rather as an ongoing project with continuity across academic seasons and years. That is, administering the same or similar measures to students in consecutive cohorts offers several advantages. Obviously that kind of longitudinal sampling (longitudinal in the sense that the same or similar programs are tracked over time) yields a lot of data. Conclusions based on years of experience rather than on one-shot studies are that much more credible (Dwyer 2004).

Longitudinal data collection allows program administrators to plot the impact of curricular or logistical changes. For example, if a study abroad course that has been taught in consecutive years implements a writing-intensive mode of instruction, the success of that instructional innovation is best judged by looking across the two years preceding and following the change. If a bump in, say, a measure of student engagement can be associated with the change, and that elevated engagement is sustained as that innovation becomes part of the program routine, then the switch to more intensive writing for students would seem well warranted. An ongoing program of assessment should maintain some consistent substrate of measures and methods across time.

The (Limited) Time Factor

Any educator who has ever run a short-term study abroad program comes to view limited time as the prime enemy. For assessment in short-term programs, time is the enemy in two respects. First, some administrators and program directors may anticipate that a program of three, four, or five weeks' duration does not provide a big enough "dose" of instruction or immersion or dissonance to register a discernible change from predeparture to the final day. The administrators and directors may worry that they won't be able to demonstrate student growth. Second, if you have only three, four, or five weeks of contact with your students, you are loath to waste any of that limited time administering assessments.

The impact of program duration on study abroad learning outcomes remains an issue begging for further investigation. Acknowledging mixed results in the research, it does appear that more time spent abroad is often better than less. And yet, plenty of well-conceived short-term study abroad programs do consistently move the needle on

well-constructed outcomes assessments (Tarrant and Lyons 2011). For example, even relatively short-term world languages programs can evince measurable learning.

Experience also indicates that a robust assessment regimen need not drain away too much instructional time, even in short-term study abroad programs. Predeparture assessments can be administered during on-campus orientation sessions, and final reflective essays or digital stories can be e-mailed in the days following the end of the program. Student response rates are usually high if these assignments are required in order to receive a course grade. It is our experience, however, that response rates for survey questionnaires tend to be low, unless those questionnaires are administered during course meeting times. Finally, much assessment data can derive from classroom activities that are simply part of the curriculum. For example, students who participate in the University of Georgia's Discover Abroad programs (described previously) have been required to conduct small, independent research projects while in residence at the Great Barrier Reef. These reports—several of which have provided input for a variety of citizen science projects on coral health and monitoring—have yielded indicators for learning objectives about critical thinking and knowledge of interspecies interactions.

Program Assessment for Sustained Impact

Rubrics and guides to assessment, including study abroad assessment, often start off by reminding educators that they must design assessment protocols that reflect both an overarching purpose and the needs of various stakeholders. As such, it is important in study abroad assessment to clearly enumerate evaluation purposes (such as accreditation, marketing, faculty evaluation, and resource justification) and identify and understand the needs of key stakeholders, including administrators, funders, skeptical colleagues, and parents. Such information will help determine which measures, strategies, and sampling plans are likely to be most effective.

For many university-based study abroad offices that serve as umbrella departments for individual faculty-run short-term programs, key stakeholders for assessment include upper-level administrators, and a crucial purpose for assessment is advocacy for greater resources and optimal policies. Here, we are moving beyond assessment of individual programs in order to improve instructional quality to system-level assessment of an entire study abroad portfolio. The principles for systemwide assessment have many similarities to the methods prescribed in this chapter for individual programs.

- Systemwide metrics for assessment should match institutional objectives for international education. If an institutional mission statement refers to producing global citizens, then the study abroad office should administer a measure of global citizenship across many, if not all, study abroad programs.
- Pretest-posttest comparison group designs yield the most compelling evidence.
- While quantitative data are convincing, augmenting that data with pointed narratives derived from qualitative analyses makes for memorable results.

- Longitudinal data can be used to warrant the kinds of support you seek. For example, if you can demonstrate that student learning outcomes rose during the years when study abroad programs were allowed to retain more tuition dollars for student services, and then those outcomes fell when those tuition dollars were withheld, you have a powerful tool for advocating tuition return policies.
- Assessment findings should be disseminated broadly. Offer to write an article about study abroad assessment for the alumni magazine. Send your findings to the office of institutional advancement. Make sure your materials make their way into accreditation reports. And of course, display assessment results prominently on the international education website as evidence that you are keeping your eye on the return on investment for your students and institution.

In developing this initiative to use study abroad assessment for institutional advocacy, you will need allies. You can find them at the institutional research office, on the college committees on general education outcomes, and among directors for other high-impact projects such as residential learning communities and undergraduate research.

Conclusion

We began this chapter likening assessment to our means of navigation for directing students to profound, transformational learning. Do you have a nagging concern that students are having a fine time on your study abroad program, but something may still be missing? What if enrollment is good and satisfaction ratings are high but perhaps the students are not learning in an optimal fashion? A robust, ongoing assessment enterprise is really the only way to identify areas of instruction and/or program design that are in need of improvement or ripe for further enhancement. In this chapter, we have argued that an ongoing assessment across student cohorts is a preferred approach to confirm that efforts at quality improvement are succeeding in getting students where we want them to go. Such ongoing (as opposed to one-off) assessment instills confidence that programs are having a sustained impact, and a pretest-posttest comparison (versus one-time) group enables greater rigor in assessment design. This chapter has shown that a longitudinal, control comparison group design assessment is quite feasible for study abroad practitioners who are committed to assessment as a form of scholarship and to teaching as an expedition of discovery for educators as well as for students.

References and Resources

- Beliefs, Events, Values Inventory: <http://www.thebevi.com/>
- Braskamp, Larry A., David C. Braskamp, and Kelly C. Merrill. 2009. "Assessing Progress in Global Learning and Development of Students With Education Abroad Experiences." *Frontiers: The Interdisciplinary Journal of Study Abroad* 18:101-118.
- Butler University Institute for Study Abroad. "Student Satisfaction Surveys." Indianapolis, IN: Butler University Institute. <http://www.ifsa-butler.org/for-faculty-advisors/student-satisfaction-surveys.html>.
- The Center for Global Education. "California Community College Student Outcomes Abroad Research Project." Carson, CA: Center for Global Education. <http://globaled.us/cccoar/>.

- Clay, Ben. 2001. "Is This a Trick Question: A Short Guide to Writing Effective Test Questions." Kansas Curriculum Center. <http://docplayer.net/76059-Is-this-a-trick-question-a-short-guide-to-writing-effective-test-questions.html>.
- Deardorff, Darla K. 2015. *Demystifying Outcomes Assessment for International Educators: A Practical Approach*. Sterling, VA: Stylus Publishing.
- Dwyer, Mary M. 2004. "More Is Better: The Impact of Study Abroad Duration." *Frontiers: The Interdisciplinary Journal of Study Abroad* 10:151–163.
- The Forum on Education Abroad: <https://forumea.org/>
- Georgia Learning Outcomes of Students Studying Abroad Research Initiative: <http://glossari.uga.edu>
- Global Perspectives Inventory: <http://www.gpi.hs.iastate.edu>
- Intercultural Development Inventory: <https://idiinventory.com>.
- Kinginginger, Celeste. 2011. "Enhancing Language Learning in Study Abroad." *Annual Review of Applied Linguistics* 31, s1:58–73.
- Morais, Duarte, and Anthony Ogden. 2011. "Initial Development and Validation of the Global Citizenship Scale." *Journal of Studies in International Education* 15, 5:445–466.
- NAFSA: Association of International Educators: <http://www.nafsa.org>.
- National Institute for Learning Outcomes Assessment. "Transparency Framework." Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. <http://www.learningoutcomesassessment.org/TFComponentSLOS.htm>.
- National Survey on Student Engagement: <http://nsse.indiana.edu>.
- Paige, R. Michael, and Elizabeth M. Stallman. 2007. "Using Instruments in Education Abroad Outcomes Assessment." In *A Guide to Outcomes Assessment in Education Abroad*, ed. Mell C. Bolen. Carlisle, PA: The Forum on Education Abroad.
- Rivers, Wilga. 1998. "Is Being There Enough? The Effects of Homestay Placements on Language Gain During Study Abroad." *Foreign Language Annals* 31, 4:492–500.
- Roy, Pamela, Elizabeth Wandschneider, and Inge Steglitz. 2014. "Assessing Education Abroad Outcomes: A Review of the BEVI, IDI, and GPI." White Paper. East Lansing, MI: Michigan State University Office of Study Abroad.
- Rubin, Donald L., Richard C. Sutton, Isaiah O'Rear, Gary Rhodes, and Rosalind Latiner Raby. 2014. "Opening the Doors of Education Abroad to Enhance Academic Success for Lower Achieving Students." *IIE Networker*, Fall 2014:38–41. <http://www.nxtbook.com/naylor/IIEB/IIEB0214/index.php#/38>.
- Stearns, Peter N. 2009. *Educating Global Citizens in Colleges and Universities: Challenges and Opportunities*. New York: Routledge.
- Sutton, Richard C., Ann N. Miller, Donald L. Rubin. 2007. "Research Design in Assessing Learning Outcomes of Education Abroad Programs." In *A Guide to Outcomes Assessment in Education Abroad*, ed. Mell C. Bolen. Carlisle, PA: The Forum on Education Abroad.
- Tarrant, Michael, and Kevin Lyons. 2011. "The Effect of Short-Term Educational Travel Programs on Environmental Citizenship." *Environmental Education Research* 18, 3:403–416.
- Tarrant, Michael A., Donald L. Rubin, and Lee Stoner. 2014. "The Added Value of Study Abroad: Fostering a Global Citizenry." *Journal of Studies in International Education* 18, 2:141–161.
- University of Delaware Institute for Global Studies. "Program Assessment." Newark, DE: University of Delaware. <http://www1.udel.edu/global/studyabroad/information/assessment.html>.
- University of Georgia Discover Abroad. "Digital Stories." Athens, GA: University of Georgia. <http://www.discoverabroad.uga.edu/index.php/media/digital-stories/>.