The Georgetown University Consortium Project: A Report at the Halfway Mark[1]

Michael J. Vande Berg
Georgetown University

Al Balkcum
University of Minnesota

Mark Scheid
Rice University

Brian J. Whalen
Dickinson College

For those of us collaborating on the project, the Georgetown University Consortium Project (GCP) represents both an end and a beginning. In its research design this three-year assessment study is an end, the product of a series of conversations among education abroad professionals at Georgetown University, the University of Minnesota, Rice University and Dickinson College.[2] The study is a beginning as well, the start of a process that will allow us to shed light on what types of experiences abroad promote meaningful knowledge and skills, and thereby to improve study abroad programming at our institutions.

Our four institutions differ significantly in student profile, structure and approaches to undergraduate education; in fact, we have deliberately crafted the study with multiple and diverse institutions in mind. What our institutions have in common is more important: a long-standing commitment to the international education of students and a history of rapid increases in study abroad enrollments. We also share a conviction that participation in this study will provide the measure of future success not merely by the crude metric of annual study abroad enrollment increases but through assessments that speak to the quality of our students' educational experiences abroad. This report written at the half-way mark of the study—a year and a half into it—presents evaluative reflections on the progress and process of the GCP to date.

Inter-Institutional Collaboration

Each of our institutions has demonstrated interest in research on student learning abroad. Two different sorts of considerations led us to decide to collaborate formally in designing and carrying out this study. First, we wanted to design a study whose findings would be significant for future study abroad programming at other institutions. Findings about the learning of students from four different types of institutions would be more likely to have a positive impact than a study focusing only on students from any single institution. Our review of education abroad assessment research literature had revealed that earlier research typically focused on the learning of students enrolled in a single institution or on students from multiple institutions enrolling in the same type of study abroad program. We continue to be struck, as we were when we first began to discuss how to design this study, by how little that earlier research has had an impact on study abroad programming. It has been all too easy for one institution to discount another institution's research on the grounds that the study's students or programs are so different that the results are simply not relevant. A study focusing on students from multiple institutions—in the GCP's case, two medium-sized private research universities, one large public university, and one private liberal arts college—who would be enrolling in a wide variety of programs abroad would be far more difficult to dismiss.

The conceptual basis of the GCP was further strengthened by our common work in the Forum on Education Abroad. The Forum mission of "contributing to the growth of knowledge and understanding about education abroad and its impact in learning and skills development"(2001) underpins the basic goals of the GCP. The four institutions in the GCP are active Forum members,[3] and the original planning for the study took place around several of the Forum's early planning meetings. Our discussions about the importance of committing ourselves to assessing students' learning abroad—one of the Forum's five principal goals—provided a fertile intellectual backdrop that enriched our thinking about student learning, and determined the GCP's focus on the two overarching questions that subsequently guided our research design efforts.

Broad Research Questions

Domain and Disciplines in Study Abroad

The first fundamental question that influenced our research design is one implicit in any study that assesses student learning: What is it that our students are learning while abroad?

We concluded early in our discussions that our study ought to respond to this question more comprehensively than earlier outcomes research. Most earlier research had

focused on students from a single institution or on a single type of program abroad, and many earlier studies had focused on only one learning domain—with much of it focusing on second-language acquisition.[4] The GCP would broaden the scope of earlier studies by focusing on student learning in three domains: gains in second-language oral proficiency, gains in intercultural sensitivity, and learning within a disciplinary context. In choosing to focus on three domains instead of one, we were and remain strongly influenced by our perceptions about faculty interest in study abroad.

We recognize that increasing numbers of faculty at US institutions are more deeply concerned about learning abroad than had been the case in the recent past. Rapid and continuing increases in U.S. student enrollments abroad have considerably increased the visibility and impact of study abroad on many campuses. Faculty who had once shown no interest in, or who had perceived scant value in, studying abroad, may have taken little notice when only a small percentage of their institution's students temporarily departed for classrooms overseas. However, more students now return to their home campuses sharing stories about their experiences—with many stories focusing on the extraordinary things the students report experiencing *outside* the academic setting. With study abroad shifting from a marginal activity involving relatively few students to an activity attracting ever increasing numbers of students, enrollments are surpassing a threshold of tolerance for some faculty. The more skeptical among them are wondering aloud why their institutions are encouraging still more students to go abroad.

Initially, the design of the study responded to this concern by considering the two learning domains that have traditionally been associated with studying abroad: second-language acquisition and intercultural learning. It would document learning in both domains, and provide an understanding about what sort of relationship might exist between these two. However, it became evident that documenting learning in these two domains would satisfy only a part of the faculty at many institutions.

The rapid gains in study abroad enrollments during the past decade and more have both stimulated, and been stimulated by, significant increases in the number of non-traditional students opting to participate. Enrollments among business students have, for example, increased from some 11% of total U.S. enrollments to nearly 20% in less than two decades, and enrollments among engineering and science students, though less dramatic, have been significant as well. Many non-traditional students have also proved to be more inclined to participate in programs of shorter duration than had earlier generations of students. Faculty members in business and other non-traditional disciplines who often lead these short-term programs, as well as many of their colleagues, are relatively less interested than faculty in the humanities and social sciences—the disciplinary areas traditionally associated with study abroad—in second-language and intercultural learning.

A faculty member in Civil Engineering, for example, may place little value in student language and intercultural learning, but still encourage students to participate in a three-week program that focuses on the theory and practice of bridge building through trips to three or four sites in Eastern Europe. This faculty member sees value in having students go abroad—but the comparative, intellectual value of this sort of program may or may not be associated with the second language and intercultural goals traditionally identified with studying abroad. In order to respond to the educational values and interests of these non-traditional faculty members, the study design added a third domain: learning in a disciplinary context. Three of the four broad disciplinary areas included—business, engineering and the sciences—are not traditionally disciplines with large enrollments in study abroad. We chose to include a fourth—the humanities—as well since we anticipated that a high percentage of students who would be participating in the study would be majoring within this broad area. Our commitment to assessing outcomes in these four areas would, as we will see, both enhance and complicate our study.

Learning Environments and Study Abroad

The second fundamental question that guided the design of our study, one that has received little attention in earlier research, is: What types of learning environments abroad tend to support or promote student learning?

The great majority of earlier assessment research has treated "study abroad" generically, implying that differences between one type of program and another were irrelevant. However, it seemed intuitively obvious that different sorts of experiences abroad will impact student learning in different ways, and we incorporated the testing of this intuitive conclusion into the design of the study. In committing to a design that would distinguish between qualitatively different types of educational experiences abroad, the GCP adopted, with minor changes, the program classification system that Lilli Engle and John Engle have developed (Engle and Engle, 2003). This system would provide a better basis for outcomes measurement than the traditional classification of study abroad programs.[5]

The traditional classification system on which study abroad professionals have long relied provides descriptive labels for program types: "island," "direct enrollment," "faculty led," "exchange," "short term," "experiential" and other program "types." In practice these types frequently overlap. A "faculty led" program may directly enroll students in a part or all of their courses at a host university, may provide for reciprocal exchange of students between home and host institutions, may incorporate internships, service-learning or some other type of experiential work, and so on. This conventional approach fails to take into account certain salient program characteristics which arguably influence or even shape student learning, like the type of student housing or

the availability of experiential activities. A direct enrollment program that houses students in apartments with other students from the sending institution presumably provides different learning opportunities than a direct enrollment program that houses students with host families, or requires them to participate in community-based volunteer projects or internships with local firms or organizations.

In Engle and Engle's classification system, individual programs are defined with reference to the way that the program responds to seven pre-identified program characteristics. The variables of a program's design establish the parameters within which students learn in that program. This approach to classification allows outcomes to be measured through providing a common basis of comparison across programs. The seven key elements are:

1. Duration of program
2. Amount of earlier target language study for each student
3. Extent to which target language is used in course work on site
4. Context of academic work (students are enrolled in courses in one of four ways: home-institution courses taught by accompanying home-institution faculty, specially-arranged host-university courses taught by host-institution faculty, regular university courses taught by host-institution faculty, or a combination of the above)
5. Type of housing for each student
6. Whether students participate in structured cultural/experiential learning
7. Whether students receive on-site mentoring that asks them to reflect on learning

Our research design provides for these seven key elements to operate as independent variables. We have collected enough data to be able to begin to measure the impact of each of these variables, as well as the impact of different groupings of the variables (for example, the combined impact of housing, duration and use of the target language in course work), on the study's three learning domains—second-language acquisition, gains in intercultural sensitivity, and learning in a disciplinary context. In order to assure a significant range of variability across program types, we are testing student learning at some fifty different sites abroad, having selected the individual programs according to the varying ways that each responds to the seven key elements.

Specific Research Questions

The GCP research activities have been guided by a number of specific research questions that we developed for each of the three learning domains we had selected as

focal points of the study. These research questions focus on student learning within, as well as across, the three learning domains:

In the Second-Language Domain:

1. Are there significant differences in second-language oral proficiency gains between students who went abroad and those who learn on the home campus?
2. What demographic and program design features (the seven key elements identified by Engle and Engle in addition to student major, year in school, gender and a few other demographic variables) significantly improve second-language oral proficiency gains?
3. Are there significant differences in second-language oral proficiency gains between students who participate in more traditional compared to less traditional study abroad sites?

In the Intercultural Proficiency Domain:

1. Are there significant differences in intercultural-proficiency gains between those students who study abroad and those who remain on the home campus?
2. What demographic and program design features significantly improve intercultural proficiency?
3. Are there significant differences in gains in intercultural-proficiency of students enrolled in different types of study abroad experiences?
4. Do gains in intercultural proficiency correlate with gains in second-language oral proficiency?
5. Are there significantly different levels of intercultural proficiency of students at the completion of their course of study abroad and over time (five to six months later)?
6. Are there significant differences in intercultural proficiency gains among students who study the same language in study abroad programs in developed and non-developed countries?
7. Are there significant differences in intercultural-proficiency gains among students who study in countries where English is the language of communication and students who study in countries where a second language is the dominant language of communication?

In the Domain of Disciplinary Learning:

1. What do Business, Engineering, Science and Humanities faculty want and expect students in their particular disciplines to learn abroad?
2. What demographic and program design factors significantly improve learning in particular disciplinary contexts?
3. What is the relationship of second-language proficiency and/or intercultural proficiency to learning in particular disciplinary contexts?

Critical Project Support

We realized early in our discussions that while we held strong views about the goals of the study, and each had years of experience managing education abroad programs, we did not collectively have the resources that needed to begin, much less to complete, this level of comprehensive study. We thus sought three different types of assistance, each of which has proved significant in the research design, instrument development and testing to date. First, we did not have access to financial resources at our own institutions that would allow us to proceed with a study of the scope we envisioned. The funding we have received through the Department of Education's Title VI International Research and Studies Program has allowed the extensive research activities we have accomplished to date.

Second, we recognized that we would need to seek expertise if we were to carry out a valid and reliable assessment study. The Title VI funding allowed us to contract as consultants three educators and researchers, each of whom continues to provide important expertise and direction in research design, instrument selection and development, training, workshop facilitation and data analysis. Dr. Mitchell Hammer (Professor of International Peace and Conflict Resolution at American University and co-creator of the Intercultural Development Inventory) and Dr. Michael Paige (Professor of International and Intercultural Education at the University of Minnesota) provide expertise in intercultural learning and research design. Dr. Margaret Malone (Senior Testing Associate for the Language Testing Division of the Center for Applied Linguistics and former Director of Peace Corps' worldwide testing program), brings to the study her expertise in second language acquisition and assessment design.

Third, we realized as we were selecting the programs abroad that would serve as our testing sites that we were not going to end up with a sufficiently broad range of programs if we confined ourselves to those operated by our four institutions. We therefore approached several study abroad providers and asked whether they would be interested in participating in the study. They generously agreed; our testing sites thus include a number of programs operated by The Council on International Educational Exchange

(CIEE), The Institute for the International Education of Students (IES), the University of Pittsburgh and the American University Center of Provence. The collaboration of these program providers, in providing access to students at a number of their sites abroad, allows the GCP to collect data across a wider range of program types than would have been possible had the research been restricted only to programs run by the four institutions.

Research Instruments and Procedures: Oral Proficiency and Intercultural Sensitivity

Oral Proficiency

The Simulated Oral Proficiency Interview (SOPI), an instrument developed in the mid-1980s at the Center for Applied Linguistics as an alternative to the Oral Proficiency Interview (OPI), is used to measure the relative proficiency gains of both the students abroad and the student control groups on the home campuses. The OPI requires a certified tester to conduct a face-to-face interview with, and to rate the oral proficiency of, each subject. The SOPI, a tape-mediated instrument whose validity and reliability are well established (Stansfield 1991, 1996), is better suited to the needs of our study since it requires no trained human interviewers on site: instead, it uses a test booklet, a master test tape, and an examinee test tape. SOPI administration at the sites abroad requires no special training or expertise. The local Resident Director or another administrator at each program site does the testing: he or she simply pushes the "play" button on one tape recorder, the "record" button on another, and then allows the student to complete tasks that are assigned both on the master tape and in the test booklet. Students complete the SOPI in about forty-five minutes. At most sites, the test is administered either in small groups of up to three students, with two tape recorders used for each student, or in larger groups in language laboratories. Once the testing is completed at test sites abroad, raters at Georgetown University's Office of International Programs, trained to evaluate SOPI performances using ACTFL proficiency guidelines (1999), assess the relative proficiency of each student.

Student oral proficiency gains across eight languages are being measured: in Arabic, French, German, Hausa, Japanese, Mandarin, Russian and Spanish. Knowledge gained through the SOPI testing will serve to establish standards for oral-proficiency gains across eight languages in a wide variety of study abroad programs. While a significant number of studies have examined second-language gains in a study abroad context, our research aims to go beyond these earlier studies, none of which has documented gains in a number of languages while simultaneously examining the impact of the various program design and demographic variables.

Intercultural Sensitivity

The GCP is utilizing the Intercultural Development Inventory (IDI) to measure student gains in intercultural sensitivity. Theoretically grounded in Milton Bennett's "Developmental Model of Intercultural Sensitivity," a frequently-cited model that describes several developmental stages that human beings go through in adapting interculturally (Bennett, 1993; Hammer, Bennett & Wiseman, 2003), the IDI is a well-respected, cross-culturally generalizable paper-and-pencil instrument that allows individual subjects to be positioned along a developmental continuum that identifies the level of intercultural competence of each respondent. Extensive field-testing and analysis of the IDI have shown it to be reliable and valid. The IDI has been developed using rigorous psychometric scale development methods (DeVellis, 1991); the extensive testing used in developing the IDI (Hammer, Bennett & Wiseman, 2003; Hammer and Bennett, 1998/2004) supports its use in measuring student gains in intercultural proficiency during study abroad. It typically takes respondents no more than twenty-five minutes to complete the test.

Participating students take the IDI three times: first as a pre-test, shortly after the students arrive at their program sites abroad; second as a post-test, shortly before the end of their programs; and finally on a post/post-test basis, some five to six months after they return to their home campuses. We hypothesize that administering the IDI a third time, as a post/post-test six months after the students' return, will shed light on an important question: how much of the intercultural learning that takes place through participation in various sorts of education abroad experiences occurs during the months following the return to campus, when students have had time to reflect upon and otherwise process their experiences abroad? Our use of the IDI to document the intercultural learning of students, including the third administration of the IDI to measure learning after re-entry, breaks new ground. Very few studies have in fact even attempted to measure gains in intercultural proficiency of students abroad, as measured through a valid and reliable instrument like the IDI (see Paige, 2003). No previous study has attempted to correlate gains in intercultural proficiency with the key program and demographic variables whose impact we will be measuring. Finally, each institution faces challenges in attempting to integrate returning students back into the home campus; we anticipate that the IDI data will provide significant information about student learning occurring during this critical re-entry period.

Developing a New Instrument: The Third Domain

Interview Guide

Our research design was significantly informed by faculty perceptions and concerns about student learning abroad. The decision to include a third domain, learning within a disciplinary context, was motivated by our sense that the focus of our research should proceed beyond second-language and intercultural learning if the study were to respond effectively to the interests of faculty in business, engineering and science, disciplines historically less associated with study abroad than the humanities and social sciences. Faculty members designing and leading programs abroad whose objective is to provide professional or technical learning opportunities for their students would be better convinced by research that argues for the value of study abroad from within their disciplines. In order to incorporate into the research design learning outcomes valued by faculty in business, engineering and science, as well as faculty members in the humanities, we organized four day-and-a-half faculty workshops during the first year of the study.

We convened separate workshops for business faculty and humanities faculty at Georgetown University in December 2002, a third workshop for engineering faculty at Rice University in February 2003 and a fourth for science faculty at the University of Minnesota in April of that same year. Ten faculty members were invited to attend each of the workshops: four from the hosting institution and two from each of the other three consortial institutions. (The engineering workshop at Rice was an exception: since among the consortial institutions only Rice and Minnesota offer engineering degrees, a total of only six engineering faculty members attended, four from Rice and two from Minnesota). Two education abroad administrators from each institution, the three outside consultants and, in the case of the workshop for science faculty at Minnesota, the study's external evaluator[6], also participated.

In addition to providing overviews of the SOPI and IDI for participating faculty, the study's three consultants facilitated discussions that allowed the faculty to identify learning competencies—the specific knowledge, skills, awareness or values—that they believed that students in their disciplines acquired, developed or enhanced while studying abroad. Faculty members from each of the disciplinary groups were told that their goal was, by the end of the workshop, to identify and produce a list of specific competencies, and each group did produce such lists. When the last of the four workshops was over, the study's three outside consultants analyzed the competencies that each of the four groups of faculty had identified, and they highlighted four "core competencies" that had been discussed at all four workshops. Faculty concluded, after

considerable discussion, that students majoring in business, engineering, the humanities and the sciences acquire, develop or enhance the following shared competencies while studying abroad:

- The ability to function on multicultural teams
- An understanding of ethical and professional responsibility
- An understanding of the impact of disciplinary solutions in a global and societal context
- An ability to apply disciplinary knowledge

The consultants subsequently developed an interview guide that was field tested at the end of spring semester 2004. This guide consists of three different hypothetical scenarios, each of which asks students to imagine how they will respond to a series of questions posed by an interviewer during either a job interview or an admissions interview for graduate school. After being trained at Georgetown by the consultants to conduct the interview, staff at the consortial institutions conducted these field tests by interviewing seniors who studied abroad during spring semester or summer of 2003. In responding to the questions posed during these hypothetical hiring or admissions scenarios, these seniors will, we hypothesize, respond in such a way as to demonstrate that they have acquired, enhanced or developed the four core competencies through their experiences abroad. Control groups of seniors who have not studied abroad will also be interviewed at the consortial institutions. All interviews will be taped; the consultants will train staff at Georgetown University to analyze the interviews.

Research Limitations and Implications

We are still some months away from being able to analyze and interpret test data. The study's formal testing began in Fall 2003, and Fall semester participants are not scheduled to take the post/post-test IDI until April of 2004; the field test of the third instrument, the field guide, occurred in May, 2004. However, even in the absence of hard data, our experience in designing the study, in carrying out the pilot and in undertaking formal testing for a semester and a half has led to several early conclusions that have implications for future research.

1. It is much more difficult to get students to volunteer for this sort of study than we had anticipated. The original grant proposal called for testing as many as 800 students abroad a year; we were able to test only 200 students during spring semester and summer of 2003, the pilot phase of the study. In fall 2003, the first semester of the formal study, some 200 students completed required pre- and post-tests (this represents

academic year students who completed the pre-test SOPIs and IDIs, as well as fall semester students); at this rate, we anticipate that in 2003-04 we will test between 500 and 550 students during the four testing periods (academic year, fall semester, spring semester and summer). The attrition rate has to date been significant: only about 55% of the students who volunteer to take pre-tests are to this point completing post-tests. We anticipate additional attrition—how much remains to be seen—when students who were abroad in Fall 2003 are asked to take the IDI post/post-tests at their home campuses.

We made special efforts to motivate student participants, such as a lottery that would award prizes to a limited number of students, with first prize the reimbursement of the winning student's international round-trip ticket. However, during the pilot phase of testing it became clear that the lottery was doing little to encourage students to participate. We therefore adopted new tactics to increase participation. First, the focus of our motivational efforts moved from the students to the Resident Directors and other staff at the sites abroad who are responsible for administering the SOPI and IDI. Program staff at Georgetown made special efforts to communicate directly with these staff abroad, both to find out what particular challenges, if any, each faced in testing students, and to ask them to take special pains to encourage more students to participate. The administrators abroad were especially encouraged to speak to the advantages of the study, to emphasize that participating students will receive a certificate with their SOPI post-test score. Second, the IDI has been posted on-line so that students can complete the test at a time convenient to them.

Our experience to date suggests that there is an inverse proportion between design complexity and student participation: the more complex and time-consuming the testing, the more difficult it will be to get students to take the tests. In a similar vein, we also believe that the number of institutions involved in a study probably influences student participation: it will be relatively easier for an institution to motivate its own students to participate in a study than it will for an institution to get other institutions' students to participate. In other words, those involved in designing a complex study with programs abroad sponsored by multiple colleges and universities, one that additionally relies on provider programs abroad, need to be prepared to devote significant energies to motivating student participants.

2. In aiming to document the impact of seven key learning variables on student learning within and across three domains, our study's scope, relative to most earlier studies, is quite broad. Even so, it focuses largely on the characteristics of programs abroad than it does on participants. Only the second of the seven key variables—the amount of earlier target language study—focuses on the student. As we've noted, we're committed to this assessment study since we believe that the results of our research will provide information that will allow us to make changes in programs abroad in order to maximize

student learning. For example, if we learn that students who participate in structured experiential learning activities make greater oral-proficiency gains than those enrolling in programs with no experiential component, we will presumably want more of our students to do internships, community-based work, research and other forms of experiential education while abroad. However, we recognize that we would be wise to wait for the results of other studies that focus on learner characteristics before deciding that all of our students should enroll in programs with experiential components.

We believe that future research focusing on such learner characteristics as student motivation and student learning capacity will complement the research we're doing and will provide a more complete picture of learning abroad. It seems likely that there are students who are motivated to learn a second language in certain contexts, and for certain purposes, rather than others. A student who is socially reserved and motivated to learn well in an academic setting, for purposes of improving his or her writing skills, may, in other words, not automatically perform as well when placed in a community-based learning situation that requires him or her to interact with native speakers in natural conversations. We have also known students who seem to have an unusual capacity for learning second languages, naturally gifted language learners who rapidly and effortlessly acquire second languages, whether in a classroom on the home campus or in any of a variety of programs abroad. There is a clear need for research that will provide better insights about learner motivation and capacity—and we will do well, in the absence of such research, not to draw too many conclusions based on program-focused research of the sort that we are conducting.

3. We chose to use the SOPI to measure oral proficiency because it is a valid and reliable instrument that allowed testing of oral proficiency gains in eight languages at a large number of sites because the testing can be completed without having to send certified OPI interviewers abroad. The test will allow us to examine whether gains in second language oral proficiency correlate with gains in intercultural sensitivity. However, since the SOPI is limited to measuring gains in oral proficiency, this study will not reveal whether our students are making gains in writing or reading while abroad.

It will be important for future studies to provide more holistic measures of second-language learning abroad. We are interested in learning more about the effectiveness of the *Test d'Evaluation de Français*. The *TEF*, designed by the *Chambre de Commerce et d'Industrie de Paris* and used widely in France and Canada, measures oral comprehension, written comprehension and vocabulary and structure. Future research should seek to document the impact of the seven key variables on student language performance based on the TEF or other instruments that allow for more holistic measurement of language performance.

4. When we originally designed the study, we intended to include a longitudinal component that would collect data on the impact that study abroad has on post-study abroad performance in courses in the major, on cumulative GPAs, on time to graduation, employability after graduation, and on other data that faculty in business, engineering, the humanities and science regard as indicative of their students' success. This would have involved staying in touch with large numbers of participants for some number of years beyond graduation. We decided not to pursue this part of the study because the administration of other components was proving more time-consuming than anticipated.

Hopefully, the results of testing with the interview guide that the GCP consultants have developed will provide evidence that study abroad allows students, especially those in non-traditional disciplines, to enhance their professional prospects. However, more research needs to be done to convince skeptical faculty that their majors will be better prepared for having studied abroad. There is value in designing a longitudinal study of the type that we originally envisioned, as there is in undertaking retrospective analyses of data already collected from past participants.[7]

What's Next

This is an ambitious study that has in several respects proved more challenging than we had anticipated. We have, however, already learned a significant amount about research design and are confident that the study's data will offer important insights about the extent to which each of the study's seven key independent variables enhances student learning. We look forward to sharing the results of our study in a future *Frontiers* issue.

Notes

[1] The authors would like to thank the members of their staffs who are responsible for the day-to-day administration of the research program: Maia Stierankova, Georgetown University, Project Coordinator, Sophie Gladding, University of Minnesota; Dr. Cheryl Matherly, Rice University; and Annmarie Whalen, Dickinson College.

[2] The study is funded by two U.S. Department of Education Title VI International Research and Studies Program grants. A three-year grant was awarded in 2002 to Georgetown University for the study designed at that point including Georgetown University, the University of Minnesota and Rice University. That original consortium subsequently invited Dickinson College to join the study; Dickinson in 2003 was awarded a separate two-year Department of Education Title VI grant to fund its participation in the study.

[3] Vande Berg, Balkcum and Scheid are founding Board members of the Forum, and Whalen is a member of the Forum's Advisory Council.

[4] The Fall 1998 special issue of *Frontiers*, "Language Learning in a Study Abroad Context," (Guest Editor: Barbara F. Freed), and Freed's earlier *Second Language Acquisition in a Study Abroad Context*, provide comprehensive collections of and commentaries on research on second language learning abroad.

[5] The commentary here draws on the discussion about the Engle and Engle's classification system in Vande Berg, M. (2003) A Research-based approach to education abroad classification. www.forumea.org/research.html.

[6] Carl Herrin (incoming chair of the Section on U.S. Students Abroad [SECUSSA] of NAFSA: Association of International Educators, and a Founding Board member of the Forum on Education Abroad) is the study's external evaluator.

[7] See, in this regard, the article in this special *Frontiers* issue by Dr. Mary Dwyer, "More is Better: the Impact of Study Abroad Program Duration," which focuses on data collected through a retrospective longitudinal study carried out by the Institute for the International Education of Students (IES).

References

American Council on the Teaching of Foreign Languages, (1999). ACTFL Proficiency *Guidelines* – Speaking, Hastings-on-Hudson, NY: ACTFL.

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

Chin, K. K. (2003). *Open Doors: Report on International Educational Exchange*. New York: Institute of International Education.

DeVellis, R.F. (1991). *Scale development*. Thousand Oaks, CA: Sage.

Engle, L. and J. Engle. (2003). "Study abroad levels: toward a classification of program types." *Frontiers: the Interdisciplinary Journal of Study Abroad*, IX, 1-20.

Forum on Education Abroad. (2001) *Prospectus* .[print brochure](1-6)

Freed, B. F. (1995). *Second language acquisition in a Study Abroad Context. Philadelphia*: John Benjamins Publishing Company.

Freed, B. F. (Guest Editor) (1998). *Frontiers: the Interdisciplinary Journal of Study Abroad. Special Issue: Language learning in a study abroad context, IV.*

Hammer, M.R., Bennett, M.J. (1998/2004). *The intercultural development inventory (IDI) manual*. Portland, OR: Intercultural Communication Institute.

Hammer, M.R., Bennett, M.J. (2003). Measuring intercultural sensitivity: The Intercultural development inventory. *International Journal of Intercultural Relations, 27*, 421-443.

Paige, M. (2003) (Ed.). *Special issue of the International Journal of Intercultural Relations, 27*.

Stansfield, C.W. (1991). A comparative analysis of simulated and direct oral proficiency interviews. In Saranee Anivan (Ed.). *Current Developments in Language Testing* (pp 199-209). Singapore: Regional English Language Center.

Stansfield, C.W. (1996). *Test Development Handbook: Simulated Oral Proficiency Interview (SOPI)*. Washington, DC: Center for Applied Linguistics.

Stansfield, C.W. & Kenyon, D.M. (1992). Research on the comparability of the oral proficiency interview and the simulated oral proficiency interview. *System, 20*, 347-64.

Vande Berg, M. (2003). A Research-based approach to Education Abroad Classification. *Forum on Education Abroad*:www.forumea.org/goalsactivities.html.