

Japanese EFL University Students and the Study Abroad Experience: Examining L2 Development and Program Satisfaction after Three Weeks in North America

Pino Cutrone*

Nagasaki University, Japan

Brien Robert Datzman

Nagasaki University, Japan

Abstract

This study is designed to shed light on the study abroad experiences of Japanese EFL university students with two goals in minds: to measure the effects of short-term study abroad on L2 proficiency in this context, and to provide EFL/ESL professionals with information that can help them develop practices to improve their study abroad programs. 98 university student participants were divided into five groups consisting of four experimental groups who studied abroad and one control group who stayed at home. Each group was given L2 proficiency tests at two points in time: prior to four of the groups embarking on three-week study abroad programs, and after the four groups returned from their time abroad. In addition, post-study abroad questionnaires and interviews were administered to assess students' attitudes and activities during the treatment period. The findings demonstrate that short-term study abroad did not have a great impact on students' TOEFL PBT scores and also provide detailed information to help study abroad professionals around the world better understand the wants and needs of Japanese EFL university students. The authors discuss the implications of these findings and offer some suggestions for study abroad professionals and researchers to consider moving forward.

Key Words: study abroad, program development, second language acquisition, L2 proficiency

Introduction

“The real voyage of discovery is not in seeing new lands but in seeing with new eyes.”

– Marcel Proust, French novelist –

As the profound and well-known quote by Marcel Proust touches upon, sojourns to new and unfamiliar lands can have such a deep and lasting impact on a person's life. For university students, who are quite often only in the early stages of developing their identities and value systems, the experience of studying abroad can forever change and shape the way they view the world around them. While the value of the personal development aspect of study abroad (SA) is widely recognized and appreciated (O'Callagan, 2006), there seems to be some question as to what else students gain from their time abroad. In particular, there appears to be some debate surrounding the effects of SA on L2 (i.e., second language) proficiency. Until recently, there has existed a widespread belief that SA contexts, which provide natural settings and informal learning through out-of-class contact with the target language, lead to higher levels of proficiency than more formal in-class educational contexts where form-focused instruction is given (Tanaka & Ellis, 2003). However, in light of mounting evidence, this assumption may require a great deal of qualification.

As the next section will explore in greater detail, SA may affect gains in certain language-specific domains (such as fluency, pragmatics and discursive abilities); however, it does not necessarily affect development in all aspects of learner proficiency (and most notably in the area of grammatical competence).

* Tel. +81(0)95-819-2927; Email: k-pino@nagasaki-u.ac.jp; School of Global Humanities and Social Sciences, Nagasaki University, Nagasaki City, Nagasaki Prefecture, 852-8521, Japan

Ironically, the grammatical domain of language development tends to be the main area that program administrators assess to determine whether students have benefited from SA (Collentine, 2009). This is particularly true in Japan where non-communicative examinations, and the washback cycle that comes with them, tend to define the curriculum (Cook, 2010, 2011). Concerning junior and senior high school contexts in Japan, much has been written about how the non-communicative nature of high stakes entrance examinations results in EFL instruction that stresses rote learning and grammar translation over fluency and communicative competence (Caine, 2005; Reesor, 2002; Sakui, 2004).

Similarly, though instructional methods will certainly vary according to institutional goals, a great number of tertiary institutions in Japan employ non-communicative tests such as the TOEFL PBT or the TOEIC to track their students' progress over time (Lee, Yoshizawa & Shimabayashi, 2006; Vongpumivitch, 2013). Since these tests focus on listening, grammar and reading, so, too, do the foci of many administrators and instructors in these settings. However, this is not to imply that the focus on these tests is misplaced and that they should not be important to Japanese people. To the contrary, obtaining high scores on these tests can be invaluable to Japanese people, as many Japanese companies still place a high value on TOEIC scores when hiring new employees (Takahashi, 2012), and TOEFL PBT scores can still be used to demonstrate a certain threshold in English proficiency (which differs across universities) to be able to study at many universities around the world (Hagerman, 2009). Rather, relating specifically to the scope of this paper, the writers question (and seek to examine more closely) whether such tests are really the most useful tools to assess the effects of SA on L2 proficiency, particularly in short-term programs.

One of the aims of this study is to help fill in the gap that appears to exist in the SA research literature. That is, various studies have looked at the gains learners make abroad in isolation (i.e., by simply looking at the pre- and post-SA abilities of one group); however, as Collentine (2009) points out, very few have actually compared learners' gains abroad with those of their student counterparts back home. Without a control group, it is difficult to know if the pre-post difference is a result of the intervention (SA) or not. Consequently, this study seeks to compare the pre- and post-TOEFL PBT scores of Japanese university EFL students who studied abroad (SA) with students from the same faculty who stayed at home (AH). Further, in addition to investigating the educational aspect of SA, this study also serves the practical managerial purpose of helping the writers improve upon the short-term SA packages that they offer to their students in the future. That is, by examining students' post-SA assessments, teachers will be able to gain a deeper understanding of the wants and needs of learners in this context and, thus, be able to make recommendations for future excursions.

Background to the study

Regarding the Effects of SA on L2 Proficiency

In surveying the SA research literature, Coleman (1997) provides a list of twenty parameters for SA research, as follows: academic context, learning outcomes, age, program at home university, previous language learning, proficiency prior to departure, preparation, duration, outgoing/incoming group dynamics and structure, L1, L2, type of accommodation, who accommodation is shared with, program of non-language courses followed, who the program was taught by, which language the program was taught in, professional content, institutional support, and assessment. Dekeyser (1991) and Polanyi (1995) include personality and gender respectively to their list of individual differences that can affect SA outcomes. Adhering to the central themes of this study, the review of the literature that follows will humbly focus on two aspects of SA: the effects of SA on L2 proficiency and the organizational element of SA program development.

First, regarding research into the effects of SA on L2 proficiency, Carroll's (1967) original study of the relationship between the language proficiency of 2,782 American students majoring in French, German, Italian, and Russian and their SA experiences is often noted as the starting point. Carroll's study examined the L2 linguistic competence of college seniors, finding that even a short duration abroad resulted in higher levels of proficiency. Based on this solid claim of proof (and perhaps also influenced by anecdotal evidence and personal speculation), language professionals, school administrators, and students (and their parents) have traditionally assumed that SA contexts provide the best environment in which to acquire a foreign language. Over time however, researchers began to suspect that such early SA research lacked an overall systematic assessment of learners' gains, and concern with general proficiency shifted towards studies that focused on particular aspects of language competence (such as reading, writing, listening, and speaking). By narrowing the focus of each investigation, incorporating more diverse methods and adopting a wide variety of theoretical approaches, researchers began to shed new light on language learning in SA contexts. First, to the surprise of many, empirical studies consistently supported the notion that SA contexts did not necessarily bring about

higher proficiency; educational contexts often predicted higher proficiency, particularly in the area of grammatical competence. Second, in SA contexts, the amount of contact with the L2 seemed to have less influence on language learning than the type of the contact, which differed depending on learners' initial levels of proficiency.

Combining the surveys of the research literature in this area conducted by Freed (1990, 1993, 1995, 1998) and Coleman (1997), Tanaka and Ellis (2003) provide us with the following summary concerning the effects of SA on language learning:

1. Accuracy and complexity, measured in terms of frequency of mistakes, sentence length or syntactic complexity in oral production, did not change in any noticeable way.
2. Gains in fluency, in terms of the speaking rate (syllables per minute) or phonation/time ratio (percentage of total time spent speaking), were strong.
3. Overall oral proficiency scores, measured by the ACTFL Oral Proficiency Interview (OPI), were higher in learners in study-abroad programs than in learners who did not participate.
4. Gains in overall oral proficiency scores were stronger than gains in test scores on grammar, listening, and reading.
5. Vocabulary gains, measured by vocabulary tests, were stronger than those of comparable students who did not participate in a study-abroad program.
6. The higher the students' initial level of proficiency, the lower the gains in proficiency as a result of studying abroad. (p. 67)

To summarize, much of the research to date seems to suggest that, overall, SA does not greatly impact grammar, listening and reading (i.e., in terms of accuracy and complexity of language use), particularly in higher-level students. Rather, an increase in natural exposure to the target language in a SA context appears to contribute more to fluency and naturalness of speech (i.e., higher speech rate, and fewer disfluency features such as false starts, repetitions and corrections). Accordingly, the findings presented above would seem to challenge some of the initial (and, in many cases, current) assumptions of the effects of SA on language learning.

Moreover, concerning levels of proficiency specifically, several studies have demonstrated the great impact that SA can have on learners with lower levels of proficiency. With this in mind however, the research in this particular area is in its infancy and a great deal of work remains. In terms that have yet to be precisely defined, researchers have begun to speculate that there exists a certain threshold that learners must reach to benefit fully from the SA context of learning (Lafford & Collentine, 2006). According to Collentine (2009), "there are most likely specific domains that require a particular developmental threshold for overall gains to occur" (p. 221).

In addition, another area of SA research that requires more attention is duration, i.e., how varying lengths of time abroad affect SA-related outcomes. Presently, it is unclear what an ideal duration of a SA program might be, as results published in this area thus far are often conflicting and inconclusive, reflecting insufficient sample sizes, the heterogeneity of the studied populations and the idiosyncratic nature of the SA experience. One of the areas in which researchers appear to be reaching a consensus is concerning grammatical development, i.e., gains in grammatical competence made by SA students are not thought to outpace those made by AH students, at least within the timeframe of a semester to a year abroad (Collentine, 2004, 2009; DeKeyser, 1990, 1991).

Furthermore, in broader terms, the conventional wisdom in the field of SA that longer is better is supported by the findings of Dwyer (2004). In her landmark study, Dwyer surveyed 3,723 former SA students in the US from the previous 50 years. In measuring the longitudinal effects of SA on academic, career, intercultural and personal development, Dwyer (2004) found that "study abroad has a significant impact on students in the areas of continued language use, academic attainment measures, intercultural and personal development, and career choices" (p. 161). While various researchers have questioned the impact of short-term SA programs on student motivation (Freed, 1990; Sasaki, 2011), there is evidence demonstrating the benefits of short-term SA programs in various language domains. For instance, Llanes and Muñoz (2009) report on how 24 Spanish students of English were able to improve their oral fluency after 3-4 weeks studying abroad in an English-speaking country. In another study, Evans and Fisher (2005) tracked the development of 68 British pupils after only 6-11 days studying abroad in France and found dramatic and sustained improvements in their L2 listening and writing skills.

Finally, it is important to keep in mind that individual differences can have a great impact on SA-related outcomes. In fact, as Huebner (1995) has pointed out, SA tends to accentuate individual differences, as “the overseas experience seems to result in a much wider variety of performances and behaviors than does study at home” (p. 191). This was certainly the case in studies conducted by Brecht and Davidson (1991) and Brecht, Davidson and Ginsberg (1995) who examined 668 American learners’ acquisition of Russian in a SA context and found a great deal of individual variation on students’ attitudes and L2 performance in this context. Undoubtedly, second language acquisition can become quite complicated by the socio-cognitive and socio-cultural pressures that learners face in the SA context, which Collentine (2009) describes as “a situation that sends many more messages to learners than does the AH context as to the complete repertoire of skills and behaviors one needs to be communicatively functional” (p. 226).

Concerning the Development of SA Programs for JEFLs

In addition to investigating the educational aspect of SA, this paper examines administrative and organizational aspects of running SA programs for Japanese EFL learners. First of all, it should be noted that SA in this context is currently facing a time of crisis, as Japanese EFL students are not studying abroad like they used to. After peaking in 2004 with 82,945 students, the number of Japanese studying abroad has continued to decline every year, dropping to a mere 58,060 in 2010 (Nagata, 2013). This 30 % decrease in less than 10 years has been attributed largely to the shrinking number of students and the fact that they perhaps tended to be more inward looking than students in the past. In light of this recent trend, the Japanese government, fully recognizing the benefits of the SA experience, has initiated several programs and incentives, which included doubling the number of SA scholarships available in the 2014 academic year, to encourage more Japanese university students to pursue SA (Torres, 2013). Accordingly, in this state of flux, it is of increasing importance that SA administrators in Japan work to understand the wants and needs of their learners and provide them with the best SA experience possible.

Given that many SA administrators serve primarily as academic staff in Japanese educational institutions, they often have larger programmatic goals in mind and may be especially concerned with the results of department-driven test scores (Hidasi, 2004). Thus, it is not surprising that few studies have been published concerning student attitudes and satisfaction relative to the SA experience. Nonetheless, even within this limited body of work, the positive effect of SA on Japanese EFL learners’ affective dimensions is evident. For instance, in a study involving 24 Japanese university students on a six-week summer study program in the US, Geis and Fukushima (1997) observed increased levels of motivation in the students’ classroom behavior upon returning to Japan and resuming their studies. In another study, which involved 139 Japanese high school students on a one-year study program in the US, Yashima (1999) found that students who tended to be more extroverted and less Japanese-centered were better able to adjust to their SA environment. Similarly, in a study involving 60 Japanese high school students on a SA program in the US, Yashima, Zenuk-Nishide and Shimizu (2004) examined the effect of willingness to communicate (WTC) on students’ perceived satisfaction of their SA experience through a questionnaire and found that pre-program WTC and self-confidence led to greater levels of satisfaction and opportunities to communicate.

While various positive aspects of student attitudes and satisfaction have been duly noted, two fairly recent studies offer some insights into how administrators can improve their students’ SA experiences. First, based on qualitative and quantitative data collected from 28 Japanese students (ranging from 19-25 years old) returning from a short-term SA abroad program conducted in an Australian university, Furmanovsky (2005) reveals some of his participants’ perceived weaknesses, which include the inability to participate fully in class discussions, weak presentation skills, the inability to explain aspects of Japanese culture in English, and a poor grasp of contemporary world affairs. Furthermore, the participants in Furmanovsky’s (2009) study expressed “a clear preference for smaller classes, in which relaxed discussion or an exchange of views with other students within groups is possible” (p. 9).

Second, another study that may help shed light on SA program reform was conducted by McIntyre (2007), who sought to determine a set of needs and goals that could be used to develop a preparation program for long-term overseas study by Japanese EFL university students. Based on interview data collected from five Japanese university students upon returning from a SA program in the US, McIntyre found that the needs and goals of students could be grouped into three main areas: cognitive/academic needs, linguistic needs, and socio-cultural needs. Educationally, participants placed a high premium on academic reading, writing, presentation and discussion skills. From a personal perspective, participants expressed how important the personal relationships they develop while abroad are to their overall SA experience. In particular, they valued

being attended to and taken care of by their host-culture teachers and were especially concerned with their ability to forge long-lasting friendships with members of the host culture.

Research Questions

Research into the area of SA is multifaceted, and this study, which modestly seeks to determine the effects of short-term SA on students' L2 proficiency and attitudes, is designed with the following objectives in mind. First, by drawing on the research to date, this study reassesses the general assumption that SA leads to higher L2 proficiency levels and, more specifically, reexamines whether standardized tests, such as the TOEFL, are adequate measurements to reflect success in (and from) SA. Second, by examining students' post-SA assessments, this study aims to provide EFL/ESL professionals with information that can help them develop practices to improve their SA programs, and, more specifically, to offer insights designed to help them better understand the wants and needs of Japanese EFL university students. To shed light on these areas of SA, the following research questions (RQs) were formulated:

RQ 1: What impact did the SA experience have on students' TOEFL PBT scores?

RQ 2: How did students feel about their study abroad experiences?

Methodology

Design

This study examines the L2 proficiency and attitudes of five groups of Japanese EFL university students, comprised of four experimental groups who studied abroad (in four separate SA programs that lasted approximately three weeks) and one control group who stayed at home. Each group was given TOEFL PBT tests at two points in time: prior to four of the groups going abroad, and after the four groups returned from their time abroad. Additionally, post-SA questionnaires and interviews were conducted to determine how the SA students felt about their time abroad as well as what type of activities the AH students engaged in while their counterparts studied abroad.

Participants

The 98 participants were all freshmen students at a national university in Japan (71 females and 27 males), who were enrolled in a faculty that focuses on the study of global humanities and social sciences and that emphasizes the study of English. When this study began, participants were on average at an intermediate level of English proficiency (as reflected by their TOEFL PBT scores), between 18 and 20 years old and had studied English for 6.5 years on average (including a collective six years in junior and senior high school). The students had been enrolled at the university for six months and were all taking courses related to English study. Participating of their free will and understanding the nature of the study, all participants were given explicit instructions (i.e., verbal and written, in both English and Japanese) regarding this study and their role in it.

As all of the participants of this study were comprised of students taught by the researchers, the JEFL learners in this study constituted an opportunistic sample. That is, all first-year students in the above mentioned faculty (who happened to be accessible to the researchers) were approached to participate in the study. The researchers had no control in choosing which students studied abroad and which ones stayed at home at the time of this study, as such decisions were made by the students themselves within the parameters of the SA options made available to them by their faculty and university. Most, if not all, of the AH students chose not to SA due to financial restraints and/or because they held out hope that they would have other opportunities in the future to SA in destinations that were more desirable to them.

Study Abroad Contexts

As shown in Table 1 below, this study comprises five groups: four experimental groups who studied abroad and one control group who stayed at home. The four SA programs involved in this study contained several similarities: they were all based in major universities in North America, and the ESL classroom instruction each program provided focused on the development of the four major skills of language competence (i.e. Reading, Writing, Listening and Speaking). In addition, each program provided a range of socio-cultural (SC) activities, which included educational field trips to museums and historical and cultural landmarks, sightseeing excursions, group shopping trips, parties with teachers, local students, and other international students, etc. The duration of each program was also similar in that Groups A, B and D studied abroad for three weeks, while

Group C studied abroad for 3.5 weeks.

In some ways however, the programs contained some important differences. First, the groups who attended the Canadian universities (Groups A, B and C) stayed with host families, and the average number of hours of ESL instruction each group received each week was 24. In comparison, the group who attended an American university (Group D) stayed in a student dormitory, which they shared predominantly with their classmates from Japan, and received an average of 18 hours of ESL instruction per week. Of the three groups that stayed with host families (Groups A, B, and C), members of Groups of A and B were often grouped together with one or more international students. In several cases these roommates were Japanese students from other universities, and a few students in Group B were paired together with their classmates from the same university in Japan. Further, the make-up of class members in each program was somewhat different. Whereas members of Groups A and B were put in classes that were mixed with students from other countries to varying degrees, Groups C and D were in customized classes that contained only Japanese students from the home institution.

Table 1
The Five Groups Used in this Study

Groups (* Control)	Number in each Group	Location	Length	Hours of Study / Week	Content of Study	Make-up of Class Members	Accommodation Settings (alone or with other students; if with others, monolingual or mixed nationalities)
A	13	Central Canada	3 weeks	24	4 skills (+ SC activities)	partially mixed nationalities	Homestay (with other students; partially mixed nationalities)
B	16	Central Canada	3 weeks	24	4 skills (+ SC activities)	mixed nationalities	Homestay (with other students; partially mixed nationalities)
C	13	Western Canada	3.5 weeks	24	4 skills (+ SC activities)	monolingual (all Japanese)	Homestay (mostly alone with host family)
D	15	Western United States	3 weeks	18	4 skills (+ SC activities)	monolingual (all Japanese)	Dormitory (with other students; monolingual - all Japanese)
* E	41	Japan	N/A	Variable	Variable	independent	N/A

At Home Context

Concerning the control group used in this study, members of Group E, who stayed at home (AH) did not receive any formal EFL instruction while the other students studied abroad. However, it is important to note that after the first TOEFL test (i.e., the pre-test that all participants took), all participants in this study (i.e., experimental and control groups) received similar instruction in the final month and a half of the semester before the SA students departed for North America. All students were enrolled in classes consisting of General English, English Communication, Reading and Discussion, Reading and Writing, English Pronunciation and Phonetics, and a First-Year Seminar class.

Instruments

Three instruments were used in this study: questionnaires, interviews and language proficiency tests. The two questionnaires used in this study were designed respectively (1) to assess the efficacy of each SA program in order to make improvements for future excursions and (2) to monitor the activities of the AH students while their SA counterparts were abroad. First, the post-SA questionnaire for SA students consisted of ten questions, which comprised a 27-item inventory (see Appendix A). Seven questions were open-ended and were general in nature (with the exception of Question 1 which was a more focused open-ended question, as it asked specifically for identifying information). The remaining three questions were comprised of 20 items that were divided into three categories of assessment (i.e., regarding academic experience, homestay experience, and cultural experience). The three categories comprised 12, 10 and 8 closed questions respectively, with statements on a Likert-scale ranging in perceptions from *poor* to *excellent* or ranging in degree from *a little* to *a great deal*. Second, the questionnaire for AH students consisted of five items, three of which were closed questions and two of which were open-ended questions (see Appendix B). Both questionnaires were administered in English; however, Japanese support (and translation) was provided to participants who requested clarification. The questionnaire was piloted on a group of five Japanese EFL university students

(who did not participate in this study), which resulted in the researchers modifying the wordings of two of the items on the post-SA questionnaire.

Another data collection instrument used in this study was interviews. Each interview was conducted in the primary researcher's office, with only the interviewee (i.e., study participant) and the interviewer (i.e., the primary researcher) present at the time of the interview. The aims of the interview were twofold: to follow up on any ambiguities and/or noteworthy responses from the participant's questionnaire, and to dig deeper into how students felt about each aspect of their SA experience. The interviews were semi-structured in that the interviewer had a general plan for the interviews, but did not enter with a predetermined set of questions, as some of the questions were guided by the context of the conversations and the responses of the interviewee. The interviews were conducted in English; however, Japanese support was provided to participants who requested clarification.

Lastly, in order to measure overall English proficiency, a paper-based version of the Test of English as a Foreign Language (TOEFL) was administered. Specifically, the paper-based version of the TOEFL used in this study refers to the TOEFL Institutional Testing Program (ITP), which is an internal testing program used in the participants' home university to monitor students' English progress over time. The TOEFL ITP Level 1 adheres to the same academic standard, composition and scoring system of the regular TOEFL Paper-Based Test (PBT). All questions are multiple choice and students answer questions by filling in an answer sheet. The tests evaluate skills in the following three areas:

- Section 1 (Listening Comprehension) measures the ability to understand spoken English as it is used in colleges and universities,
- Section 2 (Grammar: Structure and Written Expression) measures recognition of selected structural and grammatical points in standard written English, and
- Section 3 (Reading Comprehension) measures the ability to read and understand academic reading material in English.

Data Collection Procedures

Table 2 shows the schedule for the administration of the pre- and post-SA TOEFL tests, and the post-SA satisfaction questionnaires and interviews. The pre-SA TOEFL was administered some two months before the majority of the students began the SA program, while the post-SA TOEFL took place two weeks after the end of the SA program. The post-SA satisfaction questionnaire and the AH questionnaire were administered approximately one week after students returned, and the post-SA interviews were conducted approximately one to two weeks after students returned.

Table 2

Schedule of the Study

Instrument	Time Administered
TOEFL 1	2 months before study abroad
TOEFL 2	2 weeks after study abroad
Post-SA Satisfaction and AH Questionnaires	1 week after study abroad
Interviews	1-2 weeks after study abroad

Data Analysis

This section describes how the data produced by the tests, questionnaires and interviews were analyzed in this study. First, concerning the TOEFL tests, paired samples t-tests were used to determine whether the differences between the means of pre-SA and post-SA TOEFL scores were statistically significant for each group. The Statistical Package for the Social Sciences (SPSS), version 21.0, was used to analyse such data in this study. Two-tailed tests were used, and alpha levels (α) were set at 0.05. However, as the groups in this study were relatively small, probability statistics have to be viewed with caution. Thus, considering the possibility of Type 1 errors (i.e., the false rejection of the null hypothesis) occurring, probabilities less than the more stringent 0.01 and 0.001 levels will also be shown and discussed.

Second, the post-SA questionnaires and interviews were designed to shed light on how participants from each group felt about their SA experience as well as to discover what type of activities the AH students engaged in while their SA counterparts studied abroad. Responses to the closed questions on the questionnaire produced quantitative data and will, thus, be presented using descriptive statistics that denote the general

characteristics of the sample. Responses to the open-ended questions on the questionnaire and interview produced qualitative data, which will be presented in two ways: based on key aspects of SA that the researchers have chosen for further inquiry and exploration, and other previously untargeted themes that have emerged throughout the course of this study (i.e., which often came to light when the interviewer delved deeper into particular student responses and/or certain trends that began to emerge in the data).

Results

TOEFL Scores

This section has been divided into three parts. The first part will report the results concerning student TOEFL scores before and after SA, while the second and third parts will present the quantitative and qualitative aspects of students' post-SA assessments respectively. Corresponding to each of the four experimental groups (A, B, C and D) respectively, Tables 3, 4, 5 and 6 show the results of pre- and post-SA TOEFL scores. It should be noted however that TOEFL scores were only available for 71 of the 98 participants, as many of the participants were absent on one (or both) of the testing days.

As shown below, the four experimental groups exhibited similar improvements in TOEFL scores over time. The mean total TOEFL scores for Groups A, B, C and D improved by approximately 22, 20, 18 and 21 points respectively. The difference between Pre-test and Post-test scores was statistically significant for all groups, which reflected an improvement of 2.6, 4.1, 3.7 and 4.4 % respectively. As the breakdown of the sections in each table demonstrates, the largest gain in proficiency was seen in the Reading Comprehension Section of the TOEFL for all groups (10.7, 11, 6.5 and 6.2 % respectively) and the smallest in the Grammar Section (0, 1, .9 and 1 % respectively).

Table 3

Group A's TOEFL Scores Before and After Studying Abroad

N = 9	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	p
1. Listening	48.44 (3.21)	49 (3.24)	.56	1.16	-.679	.516
2. Grammar	46.89 (1.76)	46.89 (2.15)	0	0	0	1
3. Reading	47 (1.73)	52 (3.71)	5	10.64	-4.29	.003**
Total	470.78 (19.55)	492.78 (24.23)	22	2.55	-4.08	.004**

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .01$ level = ***)

Table 4

Group B's TOEFL Scores Before and After Studying Abroad

N = 13	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	p
1. Listening	51.92 (3.38)	51.77 (3)	.15	.29	.273	.790
2. Grammar	48.62 (3.53)	49.62 (4.21)	1	2.06	-.952	.360
3. Reading	48.23 (2.49)	53.54 (5.03)	5.31	11.01	-3.28	.007**
Total	496 (18.44)	516.39 (18.51)	20.39	4.11	-3.65	.003**

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .01$ level = ***)

Table 5

Group C's TOEFL Scores Before and After Studying Abroad

N = 12	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	p
1. Listening	48.75 (3.91)	50.83 (2.62)	2.08	4.27	-2.29	.043*
2. Grammar	48.42 (3.94)	48.5 (4.52)	.08	.17	-.05	.961
3. Reading	47.42 (3.83)	50.5 (3)	3.08	6.5	-3.6	.004**
Total	481.92 (29.91)	499.50 (21.76)	17.58	3.65	-2.21	.049*

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .001$ level = ***)

Table 6

Group D's TOEFL Scores Before and After Studying Abroad

N = 10	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	p
1. Listening	50.3 (3.37)	51.9 (3.47)	1.6	3.18	-3.21	.011*
2. Grammar	49 (4.45)	50 (3.74)	1	2.04	-.72	.488
3. Reading	48.2 (5.07)	51.2 (4.05)	3	6.22	-2.5	.034*
Total	491.8 (35.19)	513.2 (30.97)	21.4	4.35	-4.44	.002**

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .001$ level = ***)

Table 7 presents the TOEFL score results of the experimental groups (Groups A-D: all SA students) together, while Table 8 reports on the results of the control group (Group E: the AH students). Combining the findings shown in Tables 7 and 8, Figure 1 collectively compares the pre- and post-SA TOEFL PBT scores of SA students with those of AH students. As Figure 1 illustrates, the students who did not go abroad exhibited greater improvements in TOEFL scores overall; however, the scores of both groups generally followed the same path. The mean total TOEFL score for SA students improved by approximately 20.2 points, while that of AH students improved by 28 points. The difference between pre-test and post-test scores was statistically significant for both groups, which reflected an improvement of 4.2 and 5.8 % respectively. As the breakdown of the sections in Tables 7 and 8 demonstrate, the largest gain in proficiency was observed in the Reading Comprehension Section of the TOEFL for both groups (8.6 and 10.7 % respectively) and the smallest in the Grammar Section (1.1 and 3.2 % respectively).

Table 7

Pre- and Post-SA TOEFL Scores of all SA Students

N = 44	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	P
1. Listening	49.98 (3.7)	50.98 (3.15)	1	2	-2.67	.011*
2. Grammar	48.3 (3.57)	48.84 (3.91)	.55	1.12	-.857	.396
3. Reading	47.75 (3.42)	51.86 (4.1)	4.11	8.61	-6.39	.000***
Total	486.05 (27.3)	506.23 (24.9)	20.18	4.15	-6.63	.000***

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .001$ level = ***)

Table 8

The TOEFL Scores of AH Students at Two Points in Time

N = 27	Pre-test Mean (SD)	Post-test Mean (SD)	Mean Diff.	(%)	t	p
1. Listening	49.89 (4.47)	51.37 (4.86)	1.48	2.97	-2.3	.03*
2. Grammar	47.74 (3.47)	49.25 (5.04)	1.52	3.16	-1.9	.068
3. Reading	46.19 (5.55)	51.15 (4.31)	4.96	10.74	-4.22	.000***
Total	479.33 (35.51)	507.3 (37.88)	27.96	5.84	-4.86	.000***

(Mean difference of Pre-test → Post-test significant at $p < .05$ level = *; significant at $p < .01$ level = **, and significant at $p < .001$ level = ***)

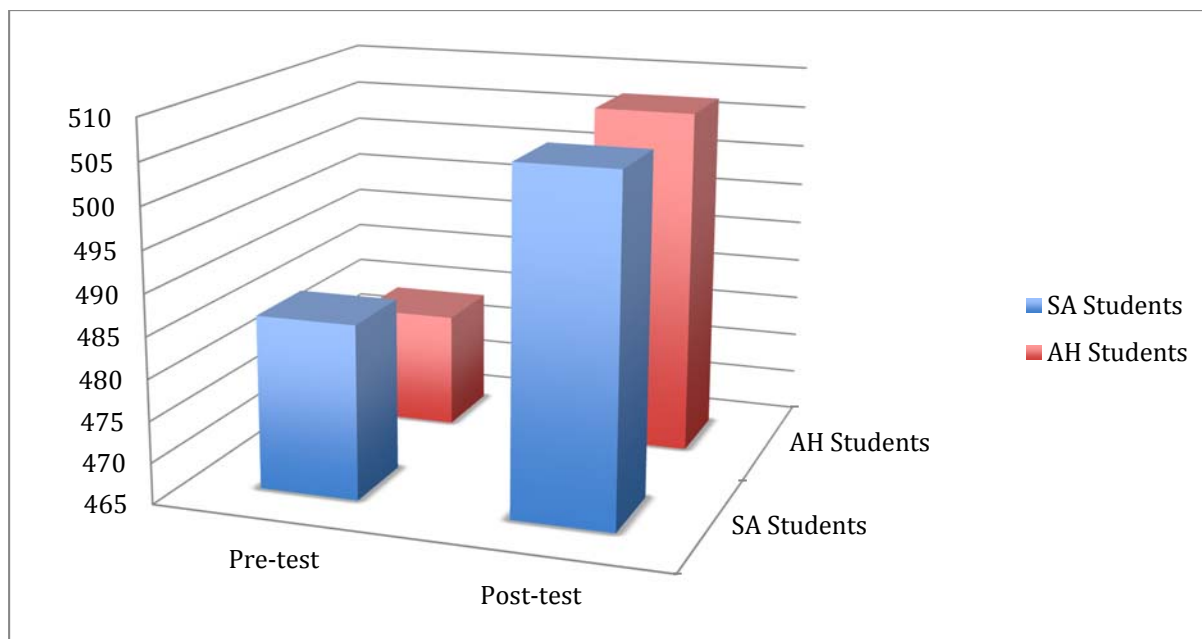


Figure 1. Comparing TOEFL Scores of SA and AH Students at Two Points in Time

Lastly, concerning the control group (Group E), 28 of 41 AH students responded to the questionnaire designed to monitor the activities of the students who did not go abroad. Although the AH students did not receive any formal EFL instruction during this time, several of them engaged in independent study to varying degrees: while no student reported to using English *a great deal*, 2 students responded to using English *much*, 9 students reported to using it *somewhat*, 12 said that they used it *a little* and 5 answered *not at all*. Some of the activities students took part in were watching English movies, communicating with foreign friends, reading graded readers and studying for the TOEFL test. Regarding how much time students spent studying for the TOEFL test specifically, 4 students reported *somewhat*, 14 responded *a little* and 10 said *not at all*. The main activities student mentioned was studying from TOEFL textbooks and attempting practice tests.

Quantitative Data Reporting on Students' Post-SA Assessments

Tables 9, 10 and 11 report on the following three facets of students' post-assessment questionnaires respectively: the academic aspect of SA, information pertaining to student accommodations, and the cultural experience. Each of these tables presents the average score for each group (A, B, C and D) corresponding to the items on the questionnaire that were closed questions. Therefore, corresponding to ratings on a Likert-scale ranging in perceptions from *poor* (1) to *excellent* (5) or in degree from *a little* (1) to *a great deal* (5), larger scores in the tables below would convey a greater sense of satisfaction or achievement. As these tables demonstrate, ratings across groups, as well as across categories, were generally quite high overall; however, there were a few notable exceptions in which groups expressed less satisfaction (i.e., noted when a group's mean rating on a particular item was ≤ 3). For instance, as shown in Table 9 (concerning the Academic aspect of SA), Group D's mean rating of 2.8 on Item B suggests that may have not been wholly satisfied with the preparation they received from their home university pre-departure. In comparison, Groups A, B and C mean ratings were 3.46, 3.5 and 3.31 respectively.

Moreover, as shown in Table 10 (concerning students' SA accommodations), the mean ratings for Group D (who, unlike the other groups, stayed in a student dormitory) were below 3 on 8 of 10 items and consistently lower than the other three groups on all 10 items. Finally, as shown in Table 11 (concerning the cultural aspect of SA), Group C and D's mean ratings of 2.92 and 3 on Item H indicates that the members in these groups may have not been wholly satisfied with the degree to which they were able to make friends with people in the local community off campus. Following this same theme, it appears that members in all groups may have not been wholly satisfied with the amount of contact they had with locals off-campus, as demonstrated by the fact the mean ratings were ≤ 3.5 for all groups on all items (F, G and H) concerning contact with local people. The issue and themes brought to light by these results warrant further enquiry and explanation and were, thus, explored in the follow-up interviews.

Table 9

Students' Post-SA Assessments Concerning the Academic Aspects of SA

Groups	A N = 13	B N = 16	C N = 13	D N = 15	Total Avg.
A. The preparation you received from your home university pre-departure	3.46	3.5	3.31	2.8	3.27
B. The orientation and support you received from the host university upon arrival	4.15	3.94	3.92	3.33	3.84
C. The accessibility and helpfulness of the host university's on-site director	4	4.5	4	3.87	4.09
D. The degree to which the study abroad program met your expectations overall	3.62	4.56	3.77	3.73	3.92
E. The safety of the host university's campus	4.77	4.75	4.69	4.60	4.70
F. The overall classroom instruction you received at the host university	4.15	3.94	3.92	4.57	4.15
G. Your instructors' course preparation and delivery at the host university	4.08	4.38	4.08	4.47	4.25
H. The reasonableness of assignments at the host university	3.92	4.06	3.85	3.93	3.94
I. The usefulness of assignments at the host university	3.85	4.31	4.15	4.27	4.15
J. The degree to which classroom instruction at the host university matched your English proficiency level	3.77	4	3.54	3.93	3.81
K. The degree to which you think your English communication skills improved overall	3.85	4.33	3.92	3.67	3.94
L. The degree to which you think your TOEFL score improved	3.08	3.19	3.08	3	3.09

Table 10

Students' Post-SA Assessments Concerning their SA Accommodation

Groups	A N = 13	B N = 16	C N = 13	D N = 15	Total Avg.
A. The degree to which the homestay/dormitory experience met your expectations	4.15	4.25	4	2.69	3.77
B. The cleanliness of the accommodation you stayed in	4.38	4.31	4.23	3.62	4.14
C. The friendliness of your host family (or other inhabitants of dormitory)	4.38	4.63	4.23	3.23	4.12
D. The food your host family (or dormitory food services) prepared for you	4.15	4.31	4.15	3	3.90
E. The safety of the home/dormitory and area that you stayed in	4.46	4.81	4.46	3.85	4.4
F. The support you received from your host family/dormitory regarding your commute to and from your host university	4.23	4.25	4.15	3.62	4.06
G. The degree to which your host family (or other inhabitants of dormitory) was interested in Japan and Japanese culture	4.15	4.38	3.69	2.69	3.73
H. The degree to which you were able to communicate in English with your host family (or other inhabitants of dormitory)	3.92	4.25	4.15	2.69	3.75
I. The degree of confidence that you will keep in touch with your host family (or other inhabitants of dormitory) in the future	3.77	4.06	3.69	2.46	3.5
J. The degree of support you received from your host university regarding any issues you had with your accommodations	4	4.11	4.13	3	3.81

Table 11

Students' Post-SA Assessments Concerning the Cultural Aspects of SA

Groups	A N = 13	B N = 16	C N = 13	D N = 15	Total Avg.
A. The degree to which the overall cultural experience met your expectations	4.23	4.38	4.5	4.07	4.3
B. The sociocultural activities you participated in, as organized by your host university	4.23	4.5	4	3.87	4.15
C. The sociocultural activities you participated in, as organized independently and informally by you and/or your friends	4.31	4.19	4.15	3.64	4.07
D. The sociocultural activities you participated in, as organized by your host family	4.08	3.75	3.69	3.08	3.65
E. The degree to which you were able to learn about and experience the host culture	4.23	4.06	4.46	3.36	4.03
F. The amount of contact you had with people in the local community off-campus	3.31	3.63	3.31	3.29	3.39
G. The degree to which you were able to communicate successfully in English with people in the local community off-campus	3.46	3.5	3.46	3.21	3.41
H. The degree to which you were able to make friends with people in the local community off campus	3.23	3.38	2.92	3	3.13

Qualitative Data Reporting on Students' Post-SA Assessments

In general, the quantifiable aspects of students' post-SA assessments were borne out in the qualitative data. Due to space limitations however, it is not possible to include all responses from students who went abroad. Therefore, to represent the various patterns found in student responses, and in answer to questions pertaining to key concepts in this study (i.e., namely, RQ 2), the following general summarizations are provided:

- Most, if not all, of the participants indicated that their SA experience was an overwhelmingly positive one.
- Many of the participants admitted to feeling anxious at first about going abroad.
- Some of the participants acknowledged that their fears about going abroad could have been allayed somewhat if they had received more information (and much earlier) prior to departure.
- Many participants linked post-SA satisfaction with how much individual attention they received in their classes and in their home stay/dormitory contexts.
- Specifically, the perceived quality of participants' SA experience seemed to be largely dependent on personal relationships, and, in particular, on the friendships they formed with their teachers, host families and other international students.
- Participants tended to prefer staying with a host family to staying in a student dormitory.
- Most participants preferred not to share a host family with another student and were particularly disappointed when another Japanese student was placed with them.
- Most participants preferred their classes to contain as few Japanese students as possible.
- In general, participants preferred small and intimate classes where they had more opportunities to speak.
- Most, if not all, participants preferred socio-cultural activities and coursework that related to the practical usage of English over coursework that was more academic in nature (such as Academic Reading, Writing and TOEFL/TOEIC preparatory courses).
- Generally speaking, participants preferred classes taught by native speakers of English.
- Many of the participants felt that the length of their SA program was too short and expressed the desire to have it extended by another week or two.
- Several participants expressed disappointment at the abrupt end of their SA program, i.e., upon

returning from their SA program, they felt they had little opportunity to use English and/or no forum to follow up on (and discuss) their SA experiences.

Discussion

The good news for students in this context is that they were able to significantly improve their TOEFL PBT scores in a relatively short period of time; however, in answer to RQ 1, SA appears to have had very little impact on student TOEFL PBT scores. In fact, as shown by the skills measured on the TOEFL PBT, AH students actually showed greater improvements than SA students (28 points vis-à-vis 20 points). This was not surprising as members of the AH group were not distracted by the many communicative demands that members of the SA group had to face, not to mention the fact that some members of the AH group actually dedicated specific attention to the TOEFL test in their independent studies. Nonetheless, the fact that the standard deviation was quite high for the AH group would seem to suggest considerable variability within the performances of this group.

While the general findings of this study are consistent with those of recent studies in this area, it was somewhat interesting to see all groups in this study significantly improve their TOEFL PBT scores. There are several possible reasons to explain this. First, the increase in scores from the pre-test to the post-test may be simply attributed to the experience factor, i.e., the fact that students taking the post-test were much better prepared (and far less anxious) because they knew what to expect from their previous experience in the pre-test (which was the first time most of the students had ever taken the TOEFL). Second, as shown by the marked improvement in the Reading Section across the board ($p < .001$ level), improvements in this area may have had more to do with the home institution's curriculum, as pre-SA instruction seemed to emphasize the development of reading skills.

Moreover, while the final part of the Results Section would seem to provide a general answer to RQ 2, some of the summarizations made therein would seem to require a degree of qualification. In particular, a major theme in student responses was that they wanted as much individual attention as possible. This speaks to the lack of L2 practice opportunities afforded Japanese EFL learners in such a homogeneous nation as Japan. Simply put, the number one fear among many Japanese students taking part in SA programs seems to be that they will waste their time abroad speaking Japanese. Thus, it makes sense for SA professionals to avoid putting Japanese students in monolingual settings wherever possible and to make a concerted effort to provide such learners with practical contexts for authentic L2 use.

With these goals in mind however, active participation in the classroom may in some cases be difficult to attain, as the research is littered with studies and anecdotes about how shy and passive Japanese EFL learners supposedly are (Greer, 2000; Townsend & Danling, 1998). Researchers such as Anderson (1993) have disputed these types of characterizations, claiming them to mere reflections of Western ethnocentrism. That is, Japanese learners actually do speak, and sometimes they speak a lot, but the contexts in which they communicate are culturally sanctioned and do not always correspond to the cultural codes of the West. Adhering to the well-known Japanese proverb *the nail that sticks up gets hammered down*, many Japanese learners tend to be reluctant to speak in contexts where they will stand out in front of their peers. Thus, it is imperative that educators in SA contexts be aware of (and sensitive to) the social and cultural codes for language use that do indeed exist. Further, as one of the writers noted previously, one instructional approach that may help draw out Japanese EFL learners involves a combination of techniques that draw on the dynamics of the Japanese classroom, with strategies that promote a Western style of interaction (Cutrone, 2009).

Conclusion

Before any conclusions are presented, it is necessary to consider the limitations of this study. First, having focused mainly on the SA-AH distinction in this study, it is important to point out that there exist several phenomena not examined in this study that may affect student performances in certain contexts. Specifically, an individual's behaviour at any point in time will also be greatly influenced by numerous variables such as personality, mood, age, gender, class, the context of the situation, group dynamics, etc. As the section below mentions, more research concerning individual differences in SA contexts would be fruitful. Second, concerning research design, the aspect that dealt with students' L2 performances involved mainly quantitative data and analysis. Undoubtedly, additional insights into the impact of the SA could have been gained by probing a bit more, with a qualitative approach, into the participants' perceived gains in, or self-assessments of, aspects of their language. Presently, the authors are embarking on a larger scale study to build on the findings of this exploratory study.

This study has explored two important aspects of the SA experience: (1) the effects of short-term SA on L2 proficiency and (2) SA program development in the Japanese university EFL context. Combining the findings of this study with those of previous studies, the following conclusions and suggestions are provided:

1. Regarding the Effects of SA on L2 Proficiency:

- Widely-used tests such as TOEIC and TOEFL PBT do not seem to be adequate measures for success in (or from) short-term SA.
- Readily available tests that include Oral Proficiency Interviews (OPI) and focus on fluency in spoken discourse, such as those administered by ACTFL and (to a lesser degree) IELTS, seem to be more well-suited to measuring the effects of SA (i.e., even the TOEFL iBT which has a speaking component is not ideal, as it focuses on students producing mini speeches rather than engaging in and reacting to the unpredictable nature and flow of dynamic and emerging conversations with an interlocutor).
- Ideally, future measurements could be constructed to include more pragmatic and discursive elements of language competencies.
- Finally, as we have come to learn, the real value of short-term SA probably has very little to do with linguistic skills at all; SA administrators would, thus, perhaps benefit from shifting their focus to investigating the affective gains produced by SA, which are perhaps more likely to reflect the deep and far-reaching impact that SA truly does have on students' lives.

2. Concerning the Development of SA Programs for JEFLLs:

- SA administrators must clearly identify program goals from the start. Ideally, home universities and host universities would work in close conjunction to meet the needs and goals of a particular group of students.
- To assess the efficacy of their programs (and determine whether certain objectives have been met), SA professionals need to continually monitor the development of students who SA (i.e., through observation and by soliciting feedback from students).
- To maximize the SA experience for students, SA professionals should work to develop sufficient pre-SA preparation courses and post-SA follow-up sessions.
- Pre-SA preparation courses should go well beyond initial orientation sessions (that help students obtain their passports, buy their plane tickets, process forms, etc) to involve expectations and cultural awareness training, language training specific to the SA context, and practical and logistical information (Amelsvoort, 2009).
- Post-SA sessions can help students reflect on their SA experience and provide them with opportunities to reinforce what they learned while abroad. Activities such as surveys and presentations would enable students to discuss (and deconstruct) aspects of their SA experience with their peers (Minehane, 2012). Furthermore, for students who have been abroad for extensive periods of time in particular, re-entry support would also be helpful (Chapell, Inaldo, White & Pirani, 2008).
- Wherever possible (i.e., concerning both accommodation and classroom contexts), SA program organizers should try to avoid putting Japanese students in monolingual settings.
- Ideally, EFL instruction in short-term SA contexts would do well to focus on practical language usage and experiential learning rather than test preparation and lecture-based methods.
- Lastly, when dealing with Japanese EFL learners, it really helps to have SA professionals who are familiar with (and sensitive to) the general wants and needs of this particular group of learners. As professionals in this context may very well know, some Japanese EFL learners, regardless of their English proficiency level, can have an especially difficult time conveying feelings of dissatisfaction and unhappiness directly (Cutrone, 2013); thus, it is often left to the instructor and/or program director (and the receptive skills they possess) to intuitively pick up on what the students are truly feeling and to take action accordingly.

3. In terms of future research:

- To better understand the communicative needs and demands of learners in the SA context, more

research comparing the L2 development of SA students with that of their AH counterparts needs to be carried out.

- In addition to controls (AH groups), which are needed to help eliminate alternative explanations of experimental results, researchers would be well served to include delayed post-tests to detect whether the effects of SA are indeed sustainable over time.
- Furthermore, researchers need to begin to assess the effects of particular teaching strategies and curriculum design on learner development in (and for) SA contexts. According to Doughty and Long (2003), Task-Based Language Teaching (TBLT) provides one such curricular framework worthy of further inquiry, as the types of interactions in which learners can engage in are particularly well-suited to SA contexts.
- Finally, concerning individual differences, little is known about how variables such as gender, age, proficiency levels, and personality affect SA outcomes and, thus, more research in such areas would be useful.

In conclusion, the present study contributes to our understanding of the SA experiences of Japanese EFL university students. In the broader context of language pedagogy, the findings of this study seem to confirm that short-term SA has little effect on students' L2 proficiency where grammar and reading are concerned. The writers hope that some of the suggestions offered in this paper will serve to assist SA professionals; however, they also fully appreciate that a great many of the recommendations provided may be beyond the control of many SA professionals, who often have to contend with various programmatic considerations as well as the practical constraints of their respective contexts. With this in mind, the writers believe that the first and most important step is to dedicate more research attention to SA and, thus, modestly hope that this study will serve as a platform for future investigation and diagnosis into this area.

References

- Amelsvoort, M. V. (1999) A pre-departure program for students who study abroad. *The Internet TESL Journal*, 5(9). Retrieved from <http://iteslj.org/Lessons/Amelsvoort-PreDeparture.html>
- Anderson, F. (1993). The enigma of the college classroom: Nails that don't stick up. In P. Wadden (Ed.) *A handbook for teaching English at Japanese colleges and universities* (pp. 101-110). New York, NY: Oxford University Press.
- Brecht, R., & Davidson, D. (1991). Language acquisition gains in study abroad: Program assessment and modification. Paper presented at the NFLC conference on language testing, Washington D.C.
- Brecht, R., Davidson, D., & Ginsberg, R. (1995). Predicting and Measuring Language Gains in Study Abroad Settings. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37-66). Amsterdam, Holland: John Benjamins Publishing Company.
- Caine, N. A. (2005). EFL examination washback in Japan: Investigating the effects of oral assessment on teaching and learning. Unpublished Master's Thesis, University of Manchester, Manchester.
- Carroll, J. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals* 1, 131-151.
- Chapell, L., Inaldo, J., White, M., & Pirani, C. (2008, May). Providing post-study abroad support – It's easier than you think! Paper presented at the 2008 NAFSA National Conference, Washington, DC.
- Coleman, J. (1997). Residence abroad within language study. *Language Teaching*, 30(1), 1-20.
- Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition*, 26(2): 227-248.
- Collentine, J. (2009). Study abroad research: Findings, implications and future directions. In M. H. Long & C. J. Catherine (Eds.), *The handbook of language teaching* (pp. 218-233). New York, NY: Wiley.
- Cook, M. (2010). Offshore outsourcing teacher in-service education: The long-term effects of a four-month pedagogical program on Japanese teachers of English. *TESL Canada Journal*, 28(1), 60-76.
- Cook, M. (2011, January 16). Improve teacher training at home. *The Japan Times*, p. 14.
- Cutrone, P. (2009). Overcoming Japanese EFL learners' fear of speaking. *University of Reading's Language Studies Working Papers*, 1(1), 55-63.
- Cutrone, P. (2013). *Assessing pragmatic competence in the Japanese EFL context*. New Castle Upon Tyne, UK: Cambridge Scholars Publishing.
- DeKeyser, R. (1990). From learning to acquisition? Monitoring in the classroom and abroad. *Hispania*, 73, 238-247.

- DeKeyser, R. (1991). Foreign language development during a semester abroad. In B. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 104 – 199). Lexington, MA: D.C. Heath.
- Doughty, C. J., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning and Technology*, 7(3), 50-80.
- Dwyer, M. M. (2004). More is better: The impact of study abroad program duration . *Frontiers: The Interdisciplinary Journal of Study Abroad*, 10(1), 151-163
- Evans, M., & Fisher, L. (2005). Measuring gains in pupils' foreign language competence as a result of participation in a school exchange visit: The case of Y9 pupils at three comprehensive schools in the UK. *Language Teaching Research* 9(2), 173–192.
- Freed, B. (1990). Language learning in a study abroad context: The effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency. In J. Atlatis (Ed.), *Georgetown University round table on languages and linguistics 1990* (pp. 459-477). Washington D.C.: Georgetown University Press.
- Freed, B. (1993). Assessing the linguistic impact of study-abroad: What we currently know – what we need to learn. *Journal of Asian Pacific Communication*, 4(1), 151-161.
- Freed, B. (1995). Language learning and study-abroad. In B. Freed (Ed.), *Second language acquisition in a study-abroad context* (pp. 3-33). Amsterdam, Holland: John Benjamins.
- Freed, B. (1998). An overview of issues and research in language learning in a study-abroad setting. *Frontiers: The Interdisciplinary Journal of Study-abroad*, 4(1), 21-60.
- Furmanovsky, M. (2005). Japanese students' reflections on a short-term language program. *The Language Teacher*, 29(12), 3-9.
- Geis, L., & Fukushima, C. (1997). Overview of a study abroad course. *The Language Teacher*, 21(11), 15-20.
- Greer, D. (2000). The eyes of Hito: A Japanese cultural monitor of behavior in the communicative classroom. *JALT Journal*, 22(1), 183-195.
- Hagerman, C. (2009). English language policy and practice in Japan. *Osaka Jogakuin College Kiyo Journal*, 6, 47-64.
- Hidasi, J. (2004). The impact of culture on second language acquisition. Paper presented at the Comparative Pragmatics Association 2nd Conference, Tokyo (3 April). Retrieved from: <http://www.childresearch.net/RESOURCE/RESEARCH/2006/exfile/HIDASI.pdf>
- Huebner, T. (1995). The Effects of Overseas Language Programs: Report on a Case Study of an Intensive Japanese Course. In B. Freed (Ed.), *Second language acquisition in a study-abroad context* (pp. 171-193). Amsterdam, Holland: John Benjamins.
- Lafford, B., & Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language: From research to application. In M. R. Salaberry & B. Lafford (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp.103-126). Washington, D.C.: Georgetown University Press
- Lee, S. I., Yoshizawa, K., & Shimabayashi, S. (2006). The content analysis of the TOEIC and its relevancy to language curricula in EFL contexts in Japan. *JLTA (Japan Language Testing Association) Journal*, 9(1), 154-173.
- Llanes, À., & Muñoz, C. (2009) A short stay abroad: Does it make a difference? *System*, 37(3), 353-365.
- McIntyre, D. J. (2007). Preparation for long-term overseas study: Toward an integrated approach. *Asian EFL Journal*, 9(3). Retrieved from http://www.asian-efl-journal.com/Sept_2007_djm.php
- Minehane, G. (2012). Door to door. Nagoya, Japan: Perceptia Press.
- Nagata, K. (2013, August 28). Forum stresses career benefits of study abroad. *The Japan Times*. Retrieved from <http://www.japantimes.co.jp/news/2013/08/28/national/forum-stresses-career-benefits-of-study-abroad/#.VO-4tEIVfHg>
- O'Callagan, J. A. (2006). How a one semester study abroad experience affects an undergraduate college student's identity development. Unpublished Master's Thesis, Montana State University, Bozeman, Montana.
- Polanyi, L. (1995). Language learning and living abroad: Stories from the field. In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 271–292). Amsterdam, Holland: John Benjamins.
- Reesor, M. (2002). The bear and the honeycomb: A history of Japanese English language policy. *NUCB JLCC (Nagoya University of Commerce and Business Journal of Language, Culture and Communication)*, 4(1), 41-52.
- Sakui, K. (2004). Wearing two pairs of shoes: language teaching in Japan. *ELT J*, 58(2), 155-163.

- Sasaki, M. (2011). Effects of various lengths of study-abroad experience on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly*, 45(1), 81-105.
- Takahashi, J. (2013). An Overview of the Issues on Incorporating the TOEIC test into the University English Curricula in Japan. *Bulletin of Tama University, School of Global Studies*, 4(1), 127-138.
- Tanaka, K., & Ellis, R. (2003). Study-abroad, language proficiency, and learner beliefs about language learning. *JALT Journal*, 25(1), 63-84.
- Torres, I. (2013, May 7). Government to encourage more Japanese college students to study abroad. *Japan Daily Press*. Retrieved from <http://japandailypress.com/government-to-encourage-more-japanese-college-students-to-study-abroad-0728362/>
- Townsend, J., & Danling, F. (1998). Quiet students across cultures contexts. *English Education*, 31(1), 4-25.
- Vongpumivitch, V. (2012). Assessing English in East Asia. In J. A. Kunnan (Ed.), *The companion to language assessment*, Volume IV, Part 13 (pp. 1671-1680). Boston, MA: Wiley.
- Yashima, T. (1999). Influence of personality, L2 proficiency, and attitudes on Japanese adolescents' inter-cultural adjustment. *JALT Journal*, 21(1), 66-86.
- Yashima, T., Zenuck-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54(1), 119-152.

Appendix A

Post-Study Abroad Questionnaire

Please answer the questions below and let us know if you have any questions.

1. Which Study Abroad Program did you attend?

2. How would you describe your overall experience abroad?

3. In retrospect, was there any information you wished you had received earlier? If so, what?

4. Please rate the academic (and on-campus) aspects of the study abroad program by circling one of the words below each sentence.

A. The preparation you received from your home university pre-departure

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

B. The orientation and support you received from the host university upon arrival

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

C. The accessibility and helpfulness of the host university's on-site director (and support staff) throughout the study abroad program

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

D. The degree to which the study abroad program met your expectations overall

Not at all-----A little-----Somewhat-----Much-----A great deal

E. The safety of the host university's campus

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

F. The overall classroom instruction you received at the host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

G. Your instructors' course preparation and delivery at the host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

H. The reasonableness of assignments at the host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

I. The usefulness of assignments at the host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

J. The degree to which classroom instruction at your host university matched your English proficiency level

Not at all-----A little-----Somewhat-----Much-----A great deal

K. The degree to which you think your English communication skills improved overall

Not at all-----A little-----Somewhat-----Much-----A great deal

L. The degree to which you think your TOEFL score improved

Not at all-----A little-----Somewhat-----Much-----A great deal

5. Please rate the following aspects of your experience staying with a host family (Note that since members of Group D did not stay with a host family, the term "host family" was substituted with "student dormitory" in this section of their questionnaire).

A. The degree to which the homestay experience met your expectations

Not at all-----A little-----Somewhat-----Much-----A great deal

B. . The cleanliness of the home you stayed in

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

C. The friendliness of your host family

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

D. The food your host family prepared for you

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

E. The safety of the home and area that you stayed in

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

F. The support you received from your host family regarding your commute to and from your host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

G. The degree to which your host family was interested in Japan and Japanese culture

Not at all-----A little-----Somewhat-----Much-----A great deal

H. The degree to which you were able to communicate in English with your host family

Not at all-----A little-----Somewhat-----Much-----A great deal

I. The degree of confidence that you will keep in touch with your host family in the future

Not at all-----A little-----Somewhat-----Much-----A great deal

J. The degree of support you received from your host university regarding any issues you had with your accommodations

Not at all-----A little-----Somewhat-----Much-----A great deal

6. Please rate the following aspects of your cultural experience while studying abroad:

A. The degree to which the overall cultural experience met your expectations

Not at all-----A little-----Somewhat-----Much-----A great deal

B. The sociocultural activities you participated in, as organized by your host university

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

C. The sociocultural activities you participated in, as organized independently and informally by you and/or your friends

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

D. The sociocultural activities you participated in, as organized by your host family

Poor-----Not so Good-----Satisfactory-----Good-----Excellent

E. The degree to which you were able to learn about and experience the host culture

Not at all-----A little-----Somewhat-----Much-----A great deal

F. The amount of contact you had with people in the local community off-campus (not including host family)

Not at all-----A little-----Somewhat-----Much-----A great deal

G. The degree to which you were able to communicate successfully in English with people in the local community off-campus (not including host family)

Not at all-----A little-----Somewhat-----Much-----A great deal

H. The degree to which you were able to make friends with people in the local community off campus (not including host family)

Not at all-----A little-----Somewhat-----Much-----A great deal

7. What aspects of your host university's study abroad experience did you enjoy the most? Please explain your answer.
8. What aspects of the study abroad experience did you dislike and/or struggle with? Please explain your answer.
9. Would you recommend the Study Abroad Program you participated in to other students? Why or why not?
10. Generally speaking: If you could change anything about the study abroad options offered in your home university, what would it be? Please explain your answer.

Appendix B

Questionnaire for AH Students

Please answer the questions below and let us know if you have any questions.

1. Why did you not study abroad this summer?
2. How much did you study and/or use English this summer?

Not at all-----A little-----Somewhat-----Much-----A great deal

3. Relating to your answer to Question 2 above, please describe specifically what type of activities you did this summer to improve your English?
4. How much did you study specifically for the TOEFL test this summer?

Not at all-----A little-----Somewhat-----Much-----A great deal

5. Relating to your answer to Question 4 above, please describe specifically what type of activities you did this summer to improve your TOEFL score?