

Study Abroad Pre-Departure and Post-Program Survey Data Analysis

Casey Quintanilla, UNC Department of Statistics and Operations Research

Client: Emily Marlton and Chris Hill, UNC Study Abroad Office

4/26/2019

Abstract

Current research in study abroad programming focuses on fostering resilience in students by allowing them to experience independence in a new environment and interactions with other cultures. To measure this quantitatively, program coordinators in the UNC Study Abroad Office sent surveys to participating students before and after their time abroad. This analysis examines students' answers to three questions relating to resilience. It uses variable selection and ordinal regression methods to find the drivers of positive answers as well as increases in answers from before studying abroad to after. The main drivers of a students' responses in these resilience-focused areas were found to be related to the students' use of foreign language while abroad and their efforts to interact with the local culture and people.

1. Introduction

At UNC, approximately 400 students study abroad each fall and spring semester, and about 800 in the summer. Participants come from various majors and can participate in programs all over the world, including non-English speaking countries. The programs might include a service-learning aspect or an internship. Students pursue different types of academic credit and come in with personal goals as well. The pre-departure and post-program surveys ask the students to reflect and document questions relating these factors and more to their overall experience and skills gained while abroad. The Study Abroad Office hopes to use this analysis to determine which aspects of programs and students enhance a student's experience and foster skills of resilience.

2. Data

Three data sets were used in the analysis:

- Spring 2018 post-program survey responses
- Summer 2018 pre-departure survey responses
- Summer 2018 post-program survey responses

The summer pre- and post- sets were merged together on the students' PID. Then, the following pre-processing steps were taken on both datasets (spring and summer):

- Creating an anonymous ID variable to replace the students' PID

- Transforming ordinal variables with 5 levels (such as Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree) to a -2/-1/0/1/2 scale
- Filling missing values in these variables with 0 – the neutral answer (detail below)
- Transforming categorical variables to dummy variables
- Removing dummy variables with fewer than 4 observations
- Dividing each column by its standard deviation
- Removing rows that were still missing data

In the end, Spring became a dataframe of 322 observations on 102 variables, and Summer had 567 observations on 155 variables. Each observation (row) represented a student and each variable (column) a survey question. Each column was numerical with a variance of 1 and no missing data from the pre-processing steps mentioned above.

2.1 Missing Data

There were certain sections of the surveys which only became available if the student answered one question in a particular way. For example, one question asked if the student studied abroad in a non-English speaking country. If they answered yes, several questions about their foreign language use would pop up for them to answer. If they answered no, they did not answer those questions. That left many blank entries in certain columns of the survey data. Simply deleting these rows resulted in too few observations left to aid an analysis. If the data were categorical, another category could have been created. However, it was important to preserve the ordinality of the data. So, an answer of 0 or “Neutral” was assigned to each empty element. This means that responses of 0 can actually represent neutral, or likely N/A or don’t know.

2.2 Variables

The pre-departure survey had the following 5 sections of questions for the student:

- S1: Basic information such as Major, Program, Region
- S2: How they found, chose, and prepared for the program
- S3: Future plans for after the program
- S4: Self-reflection
- S5: Personal goals

The post-program survey had 5 sections of questions as well:

- S1: Basic information
- S2: Personal goals and if they were satisfied
- S3: Experience while abroad including housing, language, cultural immersion
- S4: Self-reflection
- S5: Resources and Preparation

Three questions from S4 that indicated aspects of resilience were chosen as response variables of interest (all answered on the scale from Strongly Disagree to Strongly Agree):

- Q10: “I am confident in my ability to communicate with people of other languages and cultures.

- Q13: “I adapt well to new experiences and changing circumstances.”
- For spring: Q66: “I have noticed a shift in my perspective of other cultures.”
- For summer: Q7: “I am independent and self-confident.”

Because those three questions were chosen as representatives of the self-reflection section, other S4 questions were not used as predictors in any models. The full list of questions used as variables can be found in the appendix.

3. Spring

As the spring cohort of students did not receive a pre-departure survey, the analysis was done on the spring post-program data to determine drivers of a high score in the three resilience-related questions of interest: Q10, Q13, and Q66.

3.1 Clustering and Visualization

To explore the data, K-means clustering was run to create 2 clusters. The optimal number of clusters was chosen based on the silhouette score – an aggregate measure of each point’s similarity to other points in its cluster and dissimilarity to points outside its cluster, seen in figure 1.

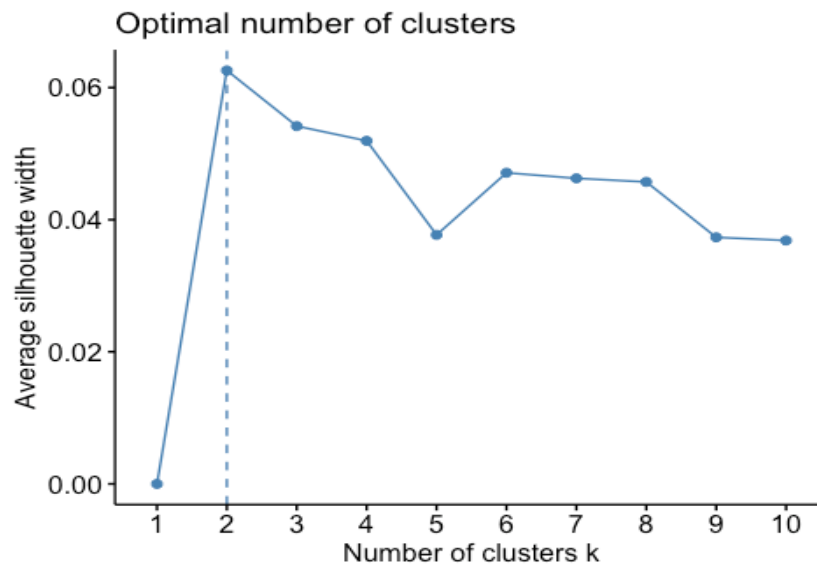


Figure 1. The average silhouette width is plotted against the number of clusters, the maximum of which is obtained at 2 clusters.

PCA was fit on the spring data to reduce dimensionality for visualization purposes. However, a very small proportion of the variance in the data was explained by the first two principal components.

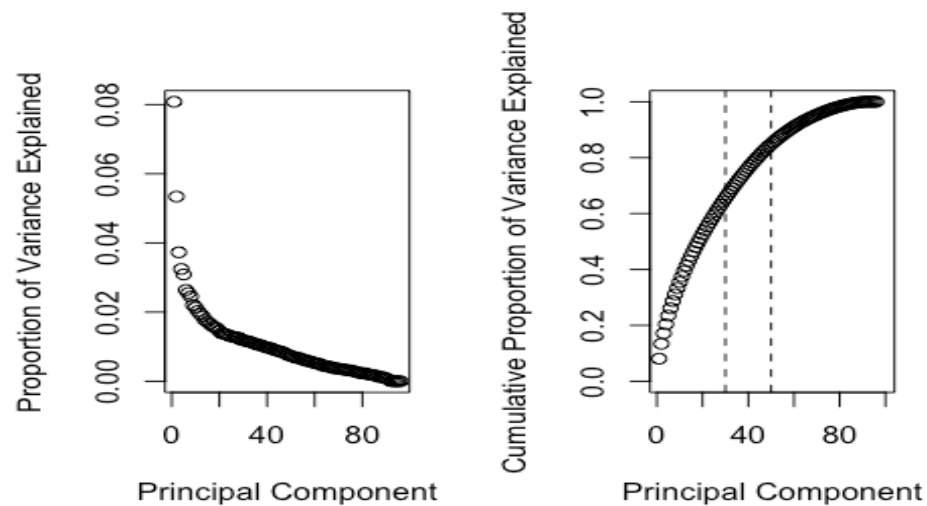


Figure 2. Plots of the proportion of variance explained by each PCA and the cumulative proportion of the same. 30 to 50 principal components are required to explain 60-80% of the variance.

As figure 2 shows, the first principal component explains just over 8% of the variation in the data, and the second under 6%. Despite this, the two minimally separated clusters are apparent when the data is plotted on 2 dimensions.

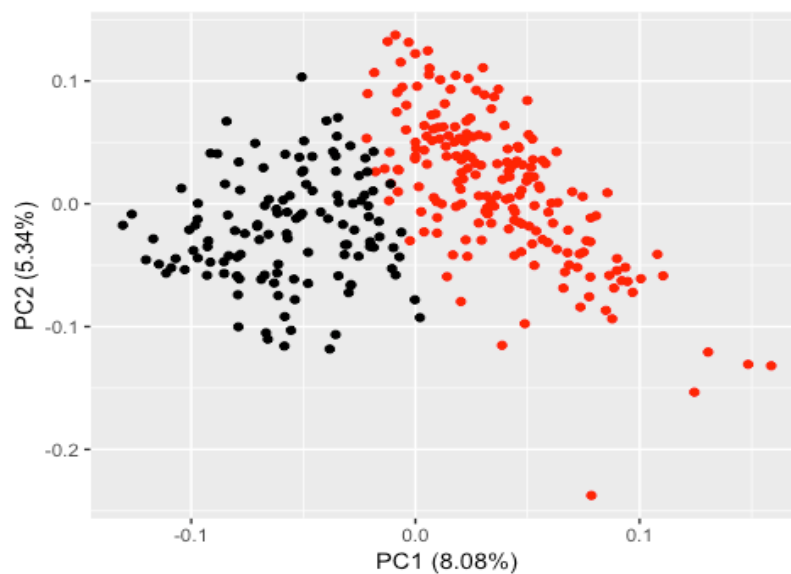


Figure 3. Spring data reduced to two dimensions and colored

Investigating the variables with distributions that change heavily from cluster 1 to 2 lead to the identification a few factors that might be behind the clustering. Below, these variables (Q5_1, Q53, Q56_1, and Q54) are plotted against the cluster number to highlight the difference in their distributions from cluster 1 to 2. Other variables did not show such dramatic differences in distributions from cluster 1 to 2.

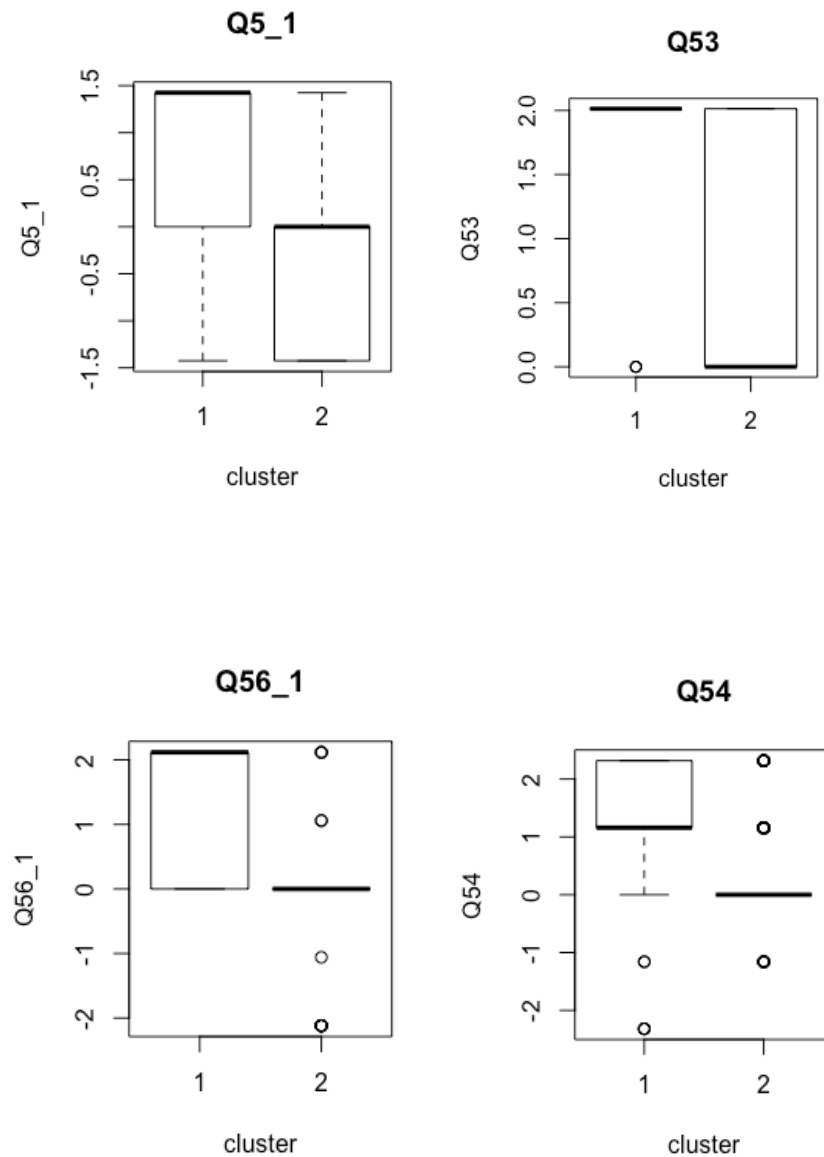


Figure 4. Boxplots of potential drivers of clustering.

To make sense of the boxplots in figure 4, note that the thick black bar represents the median survey response to that question. Half of the students responded higher than the black bar and the other half lower. If there is no box around the bar (such as Q53 cluster 1), then it represents the response of 100% of the students in that cluster, besides outliers. For example, in Q56_1, 50% of students in cluster 1

responded “Strongly Agree”, and 50% “Neutral” to “Strongly Agree”. In cluster 2 of the same question, 100% of students responded “Neutral” (besides outliers).

In Q5_1, which refers to students’ plans to take foreign language classes after studying abroad, the more positive answers of “Neutral” to “Strongly Agree” are found mostly in cluster 1, whereas cluster 2 is filled with mostly neutral and below answers. In Q53, which asks if the student studied abroad in a non-English speaking country, cluster 1 is almost entirely “yes” answers. Q56_1 asks how often the student spoke the language of the host country with their host family, and cluster 1 contains only answers of neutral (“Sometimes”) and above, mostly of “Always”. The second cluster is almost entirely comprised of neutral answers which represent both “Sometimes” answers and N/A if the student didn’t have a host family or went to an English-speaking country. Similarly, Q54 asks if the student felt their foreign language level was appropriate for the program, and cluster 2 is nearly entirely comprised of the neutral or N/A answer. It appears that use of a foreign language is the main driver between the separation of the two clusters – cluster 1 including students that studied abroad in a non-English speaking country, that utilized that host language in their daily life while abroad, felt comfortable in their ability to do so, and plan to continue to practice it upon return.

To assess how the drivers of the clusters might influence the response variables, each was plotted against the clusters.

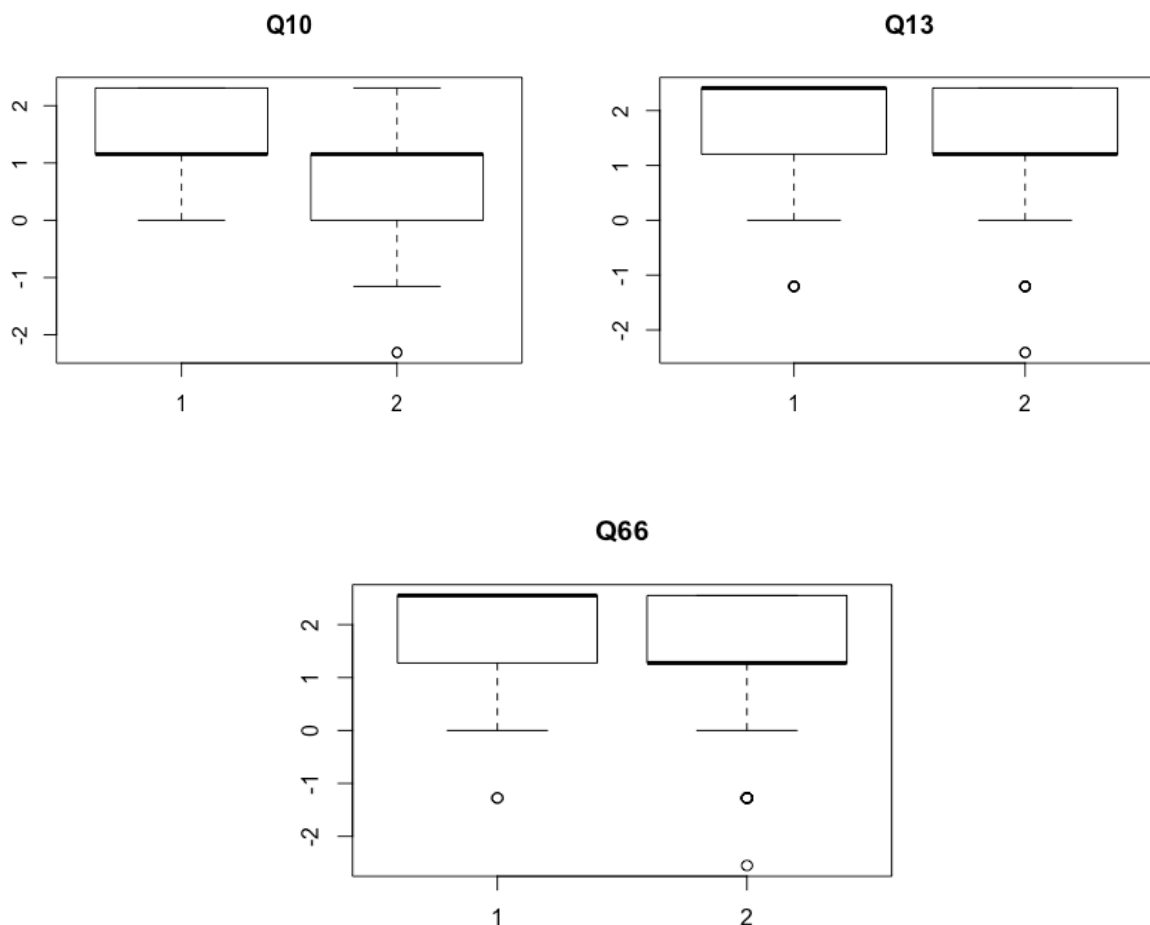


Figure 5. For each of the response variables, the distribution of the first cluster is aligned with higher responses to the question.

All three response questions have higher distributions of answers in cluster 1 than cluster 2. This might suggest that the factors behind cluster 1 indicate higher answers given for the response questions. As use of a foreign language while studying abroad was found to be a strong factor in separating cluster 1 from cluster 2, it would be reasonable to presume that the use of a foreign language increased students' ability to communicate with people of other languages and cultures, adapt to new environments and changing circumstances, and experience a shift in their perspective of other cultures – as seen through their answers to Q10, Q13, and Q66 respectively.

3.2 Variable Selection

Before running ordinal regressions to examine drivers of high scores of the response questions, the issue of collinearity in the predictor variables needed to be addressed. Often, collinearity is reduced by creating linear combinations of the variables such as PC scores. However, this would not be ideal for a study abroad advisor looking for interpretable drivers of a students' resilience. So as a preliminary assessment of which predictors drove high or low responses, binomial Elastic Net was used to produce a sparse matrix of variables determined significant to a student's answer to the three response questions. Those coefficients were then used as predictors in the proportional odds regression model to predict more accurately the Q10, Q13, and Q66 responses.

3.2.1 Elastic Net – Background

Elastic Net is a combination of two penalized linear models: Ridge and LASSO. Ridge regression addresses the issue of collinearity in the predictor variables by creating linear combinations. LASSO penalizes the addition of too many variables, producing a much shorter list of non-zero coefficients which are easy to interpret. Elastic Net tunes a parameter called alpha which determines the amount of each type of penalty to use. An alpha of 1 induces the LASSO penalty and creates sparse coefficients – very interpretable but may not include enough variables to fit the data well. An alpha of 0 induces the Ridge penalty which reduces collinearity but doesn't aid in variable selection.

3.2.2 Elastic Net – Use

Elastic Net is much less ambiguous for binomial responses than multinomial, so a binary variable was created representing if the student answered “Strongly Agree” or “Agree” to each response question (1), or “Neutral”, “Disagree”, or “Strongly Disagree” (0). So, an Elastic Net regression was fit to each of these temporary responses, producing three sets of coefficients.

The alpha level for this data had to be set to accomplish two conflicting goals:

- Be high enough to reduce the number of non-zero coefficients produced so the significant variables could be used in the ordinal regression model. If too many coefficients were used, the proportional odds algorithm did not converge.
- Be low enough to not cut out too many variables. If the coefficients were too sparse, variables with potentially useful information were not included in the ordinal regression.

Therefore, the alpha was adjusted several times until a balance was achieved between enough variables to predict the response but sparse enough for the ordinal regression algorithm to converge and the model to

be interpretable. The table below shows various alpha levels tested and the number of non-zero coefficients it produced. In this case, the alpha was set to 0.3 for all three logistic regressions.

Alpha	Response	# of Non-zero Coefficients
0.3	Q10	26
	Q13	6
	Q66	12
0.5	Q10	8
	Q13	2
	Q66	1
0.2	Q10	39 *Ordinal regression did not converge
	Q13	13 *Ordinal regression did not converge
	Q66	29 *Ordinal regression did not converge

Table 1. Number of non-zero coefficients produced by elastic net for each response and several alpha levels.

3.3 Ordinal Regression

Although Elastic Net could have been used to model the data given, it did not take advantage of the ordinal nature of the response variables. Thus, it was used only as a variable selection technique to yield potential predictors for a more appropriate model: proportional odds logistic regression models (POLR). POLR models were fit for each of the three response variables using the nonzero coefficients produced by the Elastic Net regressions. The Lipsitz goodness of fit test and the multinomial Hosmer and Lemeshow (HL) test were both run to make sure the POLR models were a good fit. Then stepwise selection was run to reduce the number of variables further, and the goodness of fit tests were run again to make sure the smaller model still fit the data. Both goodness of fit p-values for each model are shown in the table below.

Response	POLR predictors	Lipsitz test p-value	HL test p-value	Predictors after stepwise selection	Lipsitz test p-value	HL test p-value
Q10	Q14ForeignLanguage Q14SatisfyDegree Q5_1 Q5_6 Q5_11 Q2Language Q2GenEd Q42 Q53 Q54 Q56_3 Q56_4 Q60_2 Q60_5 Q61_2 Q61_3	0.1979	0.995	Q5_1 Q5_11 Q2Language Q2GenEd Q42 Q54 Q56_4 Q60_5 Q61_2 Q61_3 Q48_1 Q82FamilyContributions Q71 Q73 Major1_Computer Science Major1_Biostatistics	0.7874	0.926

	Q48_1 Q82FamilyContributions Q82PersonalSavings Q71 Q73 Major1_ComputerScience Region_North America Q90_Internship Q47_Program-provided housing					
Q13	Q5_6 Q54 Q61_2 Q61_3 Major1_Economics Major1_Business administration	0.9649	0.9442	Q5_6 Q54 Q61_2 Q61_3	0.8315	0.4391
Q66	Q14ChallengeMyself Q14PersonalGrowth Q45 Q61_1 Q61_2 Q48_2 Q48_4 Q71 Major1_Biology Major1_Media and journalism Major1_Political science Region_North America	0.5742	0.9492	Q14ChallengeMyself Q61_1 Q61_2 Q71 Region_North America	0.3309	0.72

Table 2. Predictors used in first a more saturated POLR model and its corresponding goodness of fit p-values, then the same for a POLR model selected by stepwise selection.

The POLR Predictors column in Table 1 lists the predictors selected from the Elastic Net logistic regressions and included in the first POLR model. The moderate and high Lipsitz and HL p-values indicate that these POLR models have no evidence of being misfit. The Stepwise Predictors column displays the predictors retained after performing stepwise variable selection on those POLR models. Their corresponding goodness of fit test p-values are shown in the last two columns. Again, the tests indicate no evidence of lack of fit in the models.

The summaries of the stepwise coefficients are displayed in the three tables below. The most significant variables – those with p-values less than 0.01 – are highlighted, although there are many more with p-values below 0.05 and 0.1, indicating they are also significant even if less so.

Response: Q10 I am confident in my ability to communicate with people of other languages and cultures.			
Coefficient	Question	Value	p-value
Q5_1	Because of your study abroad experience, will you take a foreign language beyond the minimum requirement?	0.3624289	0.003
Q5_11	Because of your study abroad experience, will you change your career choice?	0.42726	0.000

Q2Language	What kind of credit did you pursue abroad? Language	-0.2542133	0.028
Q2GenEd	What kind of credit did you pursue abroad? General Education requirement	-0.2611657	0.021
Q42	Was this program a good match for the credits you pursued?	0.2569663	0.027
Q54	The language level I had prior to going abroad was an appropriate level for this program	0.4171964	0.001
Q56_4	While abroad how often did you utilize the language of your host country with community members?	0.2649259	0.039
Q60_5	How often did you interact with other Americans (not from UNC)?	-0.2243378	0.046
Q61_2	I am satisfied with my effort to integrate myself with the local culture.	0.2495895	0.095
Q61_3	I am satisfied with the quality of contact I had with the local culture.	0.3187857	0.048
Q48_1	I felt safe in my housing.	0.2836675	0.020
Q82FamilyContributions	How did you pay for the program? Family Contributions	0.2860403	0.014
Q71	The Heels Abroad Handbook helped me prepare for studying abroad.	-0.1688229	0.152
Q73	The Study Abroad Pre-Departure Conference/Orientation helped me prepare for studying abroad.	-0.2285856	0.052
Major1_Computer Science	Computer Science Major	0.3468672	0.006

Table 3. Coefficient estimates from stepwise selection on a proportional odds logistic regression model with Q10 as the response and their corresponding and p-value.

After both preliminary variable selection and stepwise model selection, any variable that has been retained through this point should be taken note of. However, those variables with very low coefficients (highlighted) do indicate the most significant predictors of a student's confidence in their ability to communicate with people of other languages and cultures after studying abroad, and include:

- Deciding to take a foreign language beyond the minimum requirement after studying abroad
- Deciding to change their career choice because of studying abroad
- Feeling that the language level they had was appropriate for the program
- Being a computer science major

Response: Q13 I adapt well to new experiences and changing circumstances.			
Coefficient	Question	Value	p value
Q5_6	Because of your study abroad experience, will you apply for global jobs (international or domestic)?	0.2303431	0.041
Q54	The language level I had prior to going abroad was an appropriate level for this program.	0.2726116	0.016
Q61_2	I am satisfied with my effort to integrate myself with the local culture	0.5151729	0.001
Q61_3	I am satisfied with the quality of contact I had with the local culture.	0.2667947	0.096

Table 4. Coefficients from stepwise selection on a proportional odds logistic regression model with Q13 as the response.

The most significant predictor of a student's ability to adapt well to new experiences and changing circumstances after studying abroad is:

- Feeling satisfied with their effort to integrate with the local culture

Response: Q66 I have noticed a shift in my perspective of other cultures.			
Coefficient	Question	Value	p value
Q14ChallengeMyself	What was your primary goal for studying abroad? Challenge myself	0.2648123	0.018
Q61_1	I am satisfied with the amount of cultural interaction made available by the program.	0.3510617	0.018
Q61_2	I am satisfied with my effort to integrate myself with the local culture.	0.2746126	0.055
Q71	The Heels Abroad Handbook helped me prepare for studying abroad.	0.2301317	0.047
Region_North America	Region – North America	-0.5398837	0.000

Table 5. Coefficients from stepwise selection on a proportional odds logistic regression model with Q66 as the response.

The most significant predictor of a student noticing a shift in their perspective of other cultures was studying in North America. However, with a negative coefficient estimate, studying abroad in north America would reduce the log-odds of the student noticing a shift in their perspective of other cultures.

3.4 Conclusions

Given data only from after a student returns from studying abroad, it is apparent that key drivers of gaining resilience-related skills and confidence include students feeling satisfied with their cultural interaction and integration, their language ability in relation to the program, and making decisions about their future because of their experience studying abroad.

4. Summer

Unlike in the spring analysis set, students that studied abroad in the summer of 2018 were given a pre-departure survey as well as post-program. This creates the potential for finding additional drivers of the response questions of interest as well as drivers of the differences between those responses from before studying abroad and after. What factors might predict a student will gain resilience while studying abroad?

4.1 Clustering and Visualization

In the same fashion as the spring dataset, the silhouette width was used to determine that creating 2 clusters by K-means was optimal for the merged (pre- and post-) summer dataset.

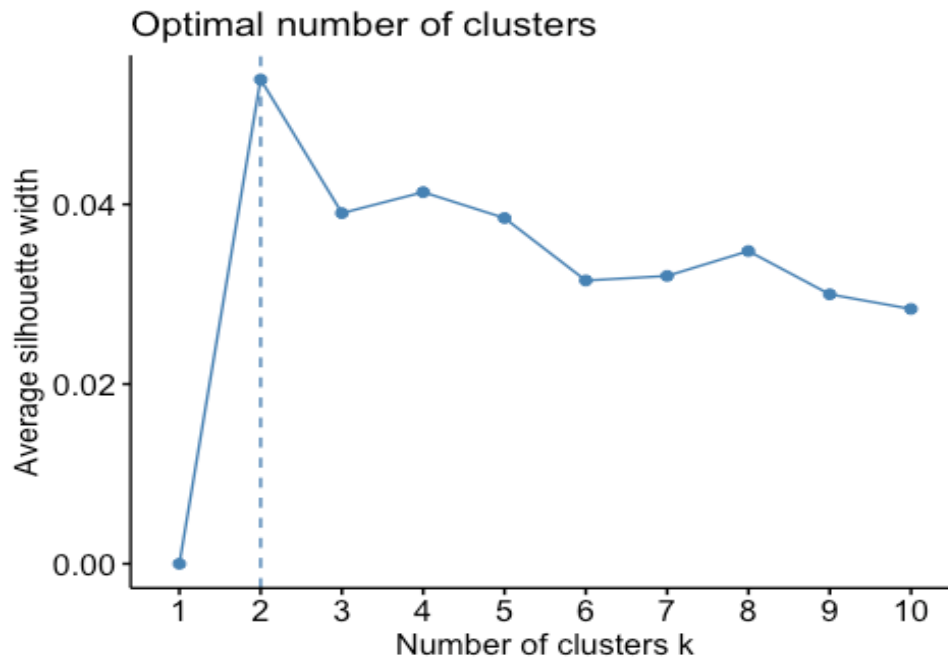


Figure 6. The average silhouette width plotted against the number of clusters, the maximum of which is obtained at 2 clusters.

PCA was used to reduce the dimensionality of the data in order to visualize these clusters, although again the first two principal components explain minimal amounts of the variation in the data (6% and 4%, respectively).

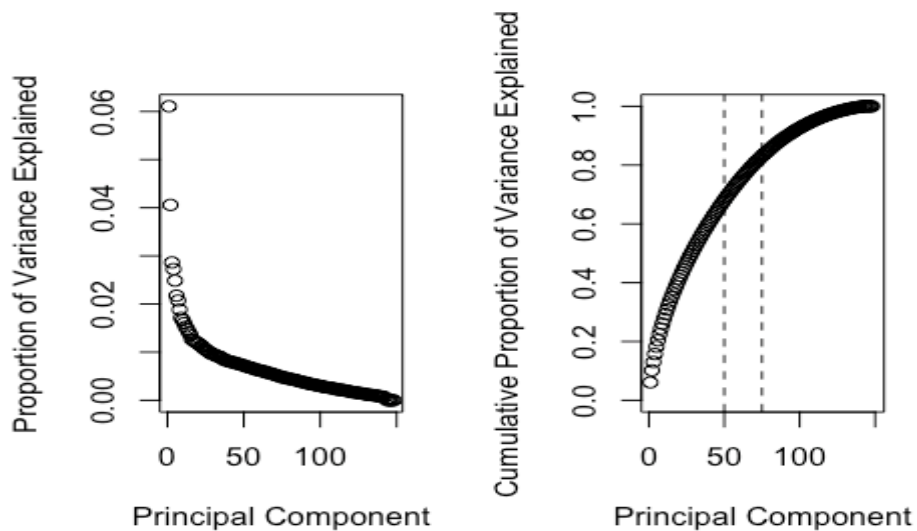


Figure 7. Variance explained by each principal component.

However, two dimensions is still enough to see separation in the clusters.



Figure 8. The summer data reduced to two dimensions from PCA and colored by cluster.

Some of the same foreign language-related variables that emerged as drivers of the clustering in the spring data also emerged in the summer.

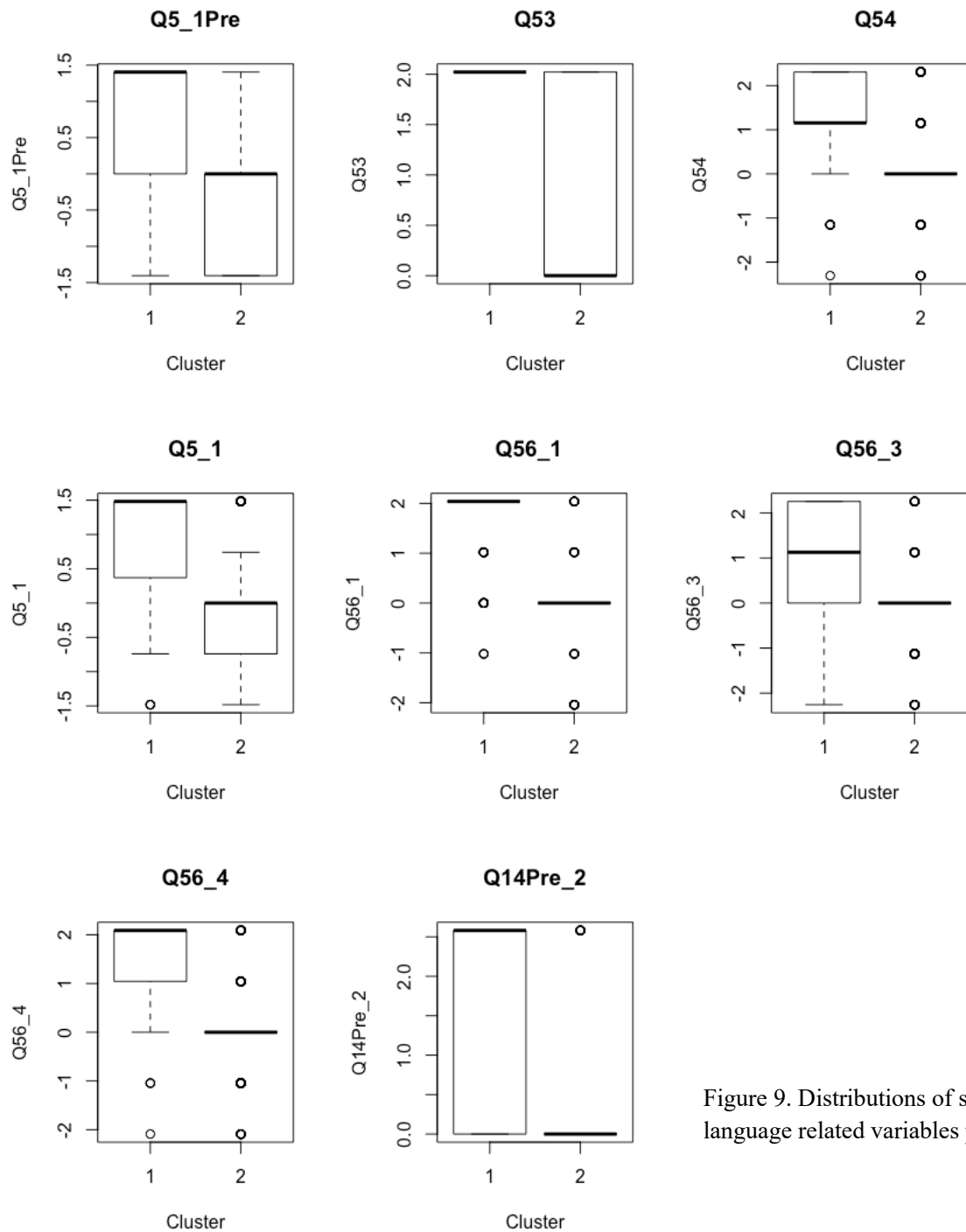


Figure 9. Distributions of some foreign language related variables plotted by cluster.

Cluster 1 is comprised entirely of students that studied in non-English speaking countries (Q53). Almost all of them utilized the host language often with their host family (Q56_1). They generally planned to continue studying a foreign language (Q5_1), felt comfortable in their foreign language skill level (Q54), and utilized the language with local students and community members while abroad (Q56_3 and Q56_4).

For many, their primary goal in studying abroad was to improve their foreign language proficiency (Q14Pre_2). In contrast, cluster 2 contains more students that studied abroad in English-speaking countries, and therefore gave neutral or N/A answers to questions relating to utilizing the host language.

In addition to these variables, clusters arose from differences in why students chose their program: cluster 1 includes more students that pursued minor credit (Q2Minor) and chose their program based on relevant coursework available (Q1Pre_2). However, this might be a testament to the popularity of the Spanish for the professions minor summer program in Costa Rica.

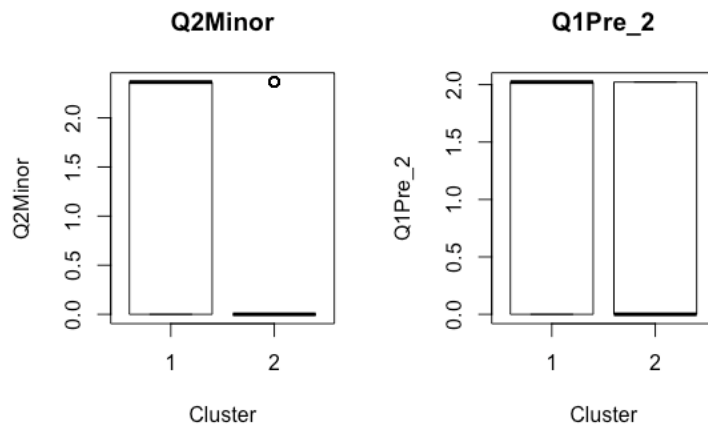


Figure 10. The distributions of answers to Q2Minor and Q1Pre_2 plotted by cluster.

Q60_1 asks how often the student interacted with their host family, so cluster 1 includes those who answered often and cluster 2 includes neutral and N/A answer – likely those students that did not do a home-stay. A similar separation can be seen in Q47, where nearly all students in cluster 1 lived in a home-stay and nearly all in cluster 2 did not, but the opposite classification emerged for students in program-provided housing.

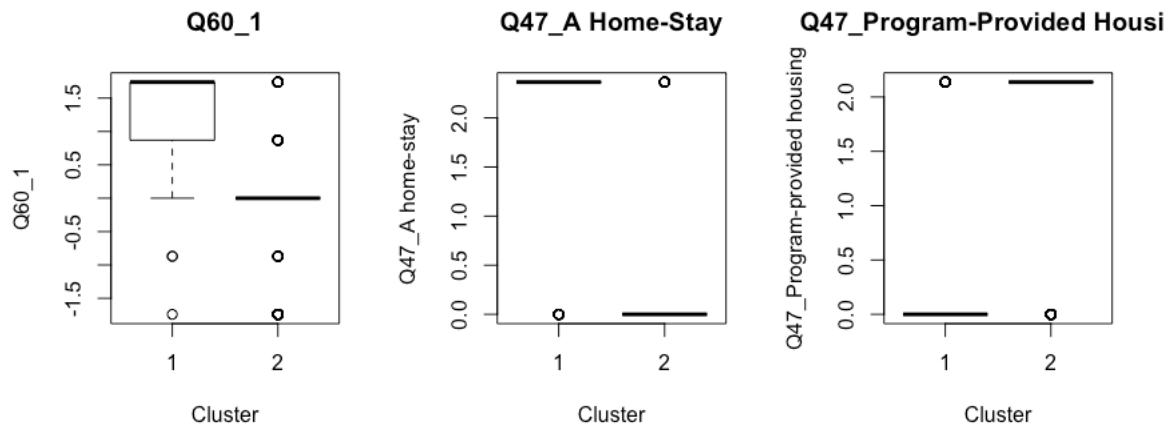


Figure 11. The distributions of the answers to some home-stay related questions plotted by cluster.

The relationship between the clusters and the response variables of interest may give indication of what drives the response – presumably, if the clustering drives the response then the drivers of the cluster would as well. Below, the distributions of the response variables are plotted against the clustering the visualize this possible relationship. In the summer dataset, each question of interest created two response variables: the raw score from the post-program survey, and the difference between post-program score and pre-departure score.

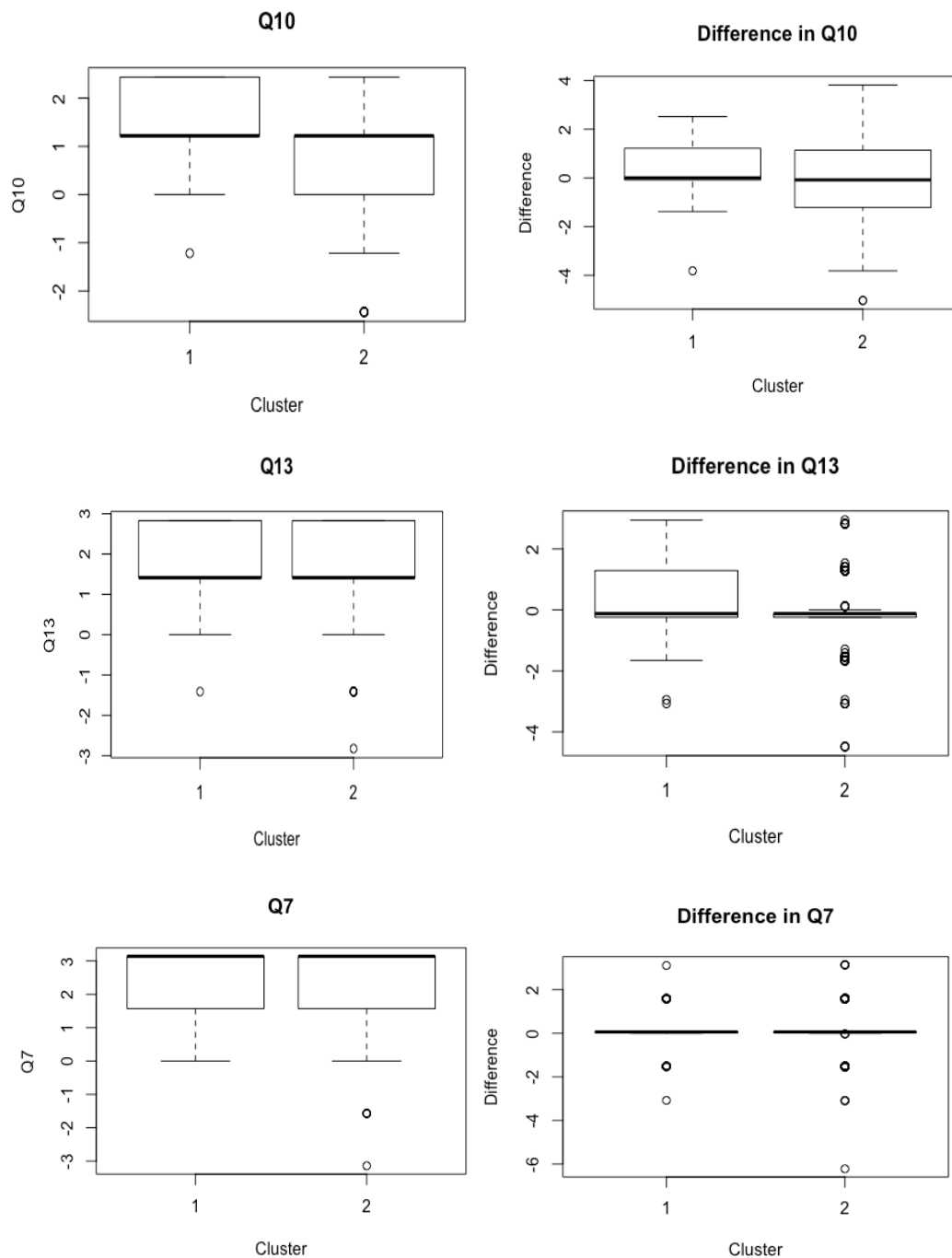


Figure 12. Distributions of the response variables plotted by cluster.

Not much variation in the response variables can be extracted from the clustering. The only differences noted were that in cluster 1, more students reported an answer of Agree or Strongly Agree for Q10, versus most students answering Neutral or Agree in cluster 2. Cluster 1 also contained fewer students that reported a decrease in their answer to Q10 from pre- to post- studying abroad. Lastly, cluster 2 contained almost entirely students that reported no change in their answer to Q13 from pre- to post- studying abroad, whereas more students in cluster 1 reported some improvement.

As cluster 2 included students that did not study abroad in a non-English speaking country and did not do a home-stay, a reasonable inference could be made that missing out on those experiences are factors to students not improving their ability to communicate with people of other languages and cultures or adapt well to new environments and changing circumstances.

4.2 Testing Pre- and Post- Differences

To compare the means of response questions by pre- or post-study abroad, a two-sample t-test was run to test the hypothesis that there is no difference between students' responses to the questions before or after studying abroad.

Response	Mean of Pre-	Mean of Post-	p-value
Q10	0.939	1.024	0.1515
Q13	1.908	1.804	0.0811
Q7	2.139	2.339	0.000781

Table 6. Testing the null hypothesis that the difference in means between each pre- and post-response equals 0 against the alternative that the difference in means is not equal to 0.

The results of the two-sample t-test indicate evidence that the true difference of means between Q7 responses before and after studying abroad is not equal to zero. However, the same cannot necessarily be said for Q10 and Q13.

4.3 Testing the Effect of Pre- or Post-

Another method used to test significant changes in responses to the questions of interest from before and after a student studied abroad was to test the effect of whether the question was answered on the pre-departure or post-program study. To accomplish this, a new dataframe was created with a column for each response question and a column indicating if the response came from the pre- or post- survey. Then a one-way ANOVA and a non-parametric Kruskal-Wallis test model were performed on each response question, the results of which are shown below.

Response	One-way ANOVA p-value
Q10	0.152
Q13	0.0811
Q7	0.000781

Table 7. P-values for each one-way ANOVA

The one-way ANOVA tests the significance of when the survey was taken on the response question. The p-values shown indicate that Q7 might be the only response where whether a student has gone abroad yet has a significant effect on their answer to the response question.

Response	Kruskal-Wallis p-value
Q10	3.078e-09
Q13	< 2.2e-16
Q7	< 2.2e-16

Table 8. P-values for each Kruskal-Wallis test.

The Kruskal-Wallis tests the null hypothesis that the distributions of each response variable are independent of when they took the survey. The extremely low p-values suggest very strong evidence that students' answers to the response questions differ depending on if they had studied abroad yet or not.

The results of the ANOVA model were quite consistent with the T-test, but the non-parametric test resulted in much more significant p-values. Both the T-tests and ANOVA model rely on the assumption that the data is normally distributed, so given that the survey data does not meet this assumption, the non-parametric test was trusted. The effect of a student studying abroad was significant in their response to the three questions of resilience.

4.4 Predicting Answer to Response – Post-Program

The same method used in the spring analysis was used on the summer for variable selection. The difference with the summer data was the opportunity to include questions from the pre-departure survey as predictors of the post-program response. A logistic regression was fit to predict if each response variable was answered as 1 or 2 (“Agree” or “Strongly Agree”), using Elastic Net to produce sparse yet sufficiently descriptive coefficients. In this case, the Elastic Net parameter alpha was set to 0.2 for Q10 and Q13 responses, but 0.175 for Q7. Cross-validation was used to find the optimal lambda penalty, and the non-zero coefficients were used as predictors in the proportional odds logistic regression (POLR) model.

4.5 Ordinal Regression (POLR)

First, all nonzero coefficients produced by the Elastic Net model were used as predictors of the response questions, then stepwise selection was used to further simplify the model. Goodness of fit was tested for each model with both Lipsitz and Hershorn and Lemeshow (HL) tests. The coefficients included and p-values for each model are displayed in the following table.

Response	POLR Coefficients	Lipsitz test p-value	HL test p-value	Stepwise coefficients	Lipsitz test p-value	HL test p-value
Q10	Q3FacultyPre Q3HostPre Q4WalkInPre Q5_1Pre Q5_5Pre Q5_6Pre Q5_1 Q5_6 Q5_11 Q2GenEd Q42 Q53 Q54 Q56_4 Q60_2 Q60_3 Q61_2 Q48_1 Q48_2 Q1Pre_5 Q1Pre_7 Q1Pre_4 Q14Pre_1 Major1_Economics Major1_Global studies Major1_Geological sciences Major1_Environmental science and studies Major1_Chemistry Major1_Archaeology Major1_Mathematics Major1_Public policy Major1_Statistics and analytics Major1_Health policy and management Major1_Neuroscience Major1_Nutrition Q90_Internship Q90_Service-learning Region_Australia/Pacific Islands Region_Middle East Q47_Program-provided housing Q47_A home-stay	0.4055	0.6015	Q3HostPre Q4WalkInPre Q5_5Pre Q5_6Pre Q5_1 Q5_6 Q54 Q60_2 Q60_3 Q61_2 Q48_2 Q1Pre_5 Q1Pre_7 Q14Pre_1 Major1_Economics Major1_Geological sciences Major1_Environmental science and studies Major1_Archaeology Major1_Mathematics Major1_Statistics and analytics Major1_Health policy management Major1_Neuroscience Region_Australia/Pacific Islands Region_Middle East Q47_A home-stay	0.2735	0.3666
Q13	Q5_2Pre Q5_6Pre Q5_6 Q2GenEd Q45 Q60_2 Q60_4 Q60_5 Q60_3 Q61_2 Q61_3 Q48_1 Q82PersonalSavings Q82FinancialAid	0.03411	0.9021	Q5_6Pre Q5_6 Q45 Q60_2 Q60_5 Q61_2 Q61_3 Q48_1 Q82PersonalSavings Q82FinancialAid	4.346e-05	0.3691

	Q1Pre_6 Q1Pre_5 Major1_Nursing Major1_English Major1_Computer Science Major1_Media and journalism Major1_Archaeology Major1_Biomedical and health sciences engineering Q90_Internship Q90_Service-learning Region_Australia/Pacific Islands Q47_Other			Q1Pre_6 Q1Pre_5 Major1_Media and journalism Major1_Biomedical and health sciences engineering Q90_Service-learning Region_Australia/Pacific Islands		
Q7	Q3AdvisorPre Q4MeetingPre Q4101Pre Q5_8Pre Q5_5 Q5_6 Q45 Q60_5 Q61_3 Q1Pre_6 Major1_Chemistry Major1_Archaeology Q47_Other	0.5061	0.5754	Q5_6 Q45 Q60_5 Q61_3 Q1Pre_6 Major1_Chemistry Major1_Archaeology Q47_Other	0.5288	0.8151

Table 9. Predictors included in each POLR regression – before and after stepwise selection, with corresponding goodness of fit test p-

Table 8 lists the coefficients included in both stages of the POLR model – all nonzero coefficients from the Elastic Net model and after using stepwise selection to reduce this. Despite one small p-value for the Lipsitz test on the Q13 model, the models appear to be good fits on the data as the tests are meant to be looked at holistically.

The summaries of the stepwise coefficients are displayed in the three tables below. The most significant variables – those with p-values less than 0.01 – are highlighted, although there are many more with p-values below 0.05 and 0.1, indicating they are also significant even if less so.

Response: Q10 I am confident in my ability to communicate with people of other languages and cultures			
Coefficient	Question	Value	p value
Q3HostPre	How did you find out about your specific study abroad program? Host/third party website	2.261	0.024
Q4WalkInPre	What resources have you utilized up to this point? Walk-in Wednesday	1.922	0.055
Q5_5Pre	After your study abroad experience will you highlight study abroad experience on resume and in interviews?	-2.239	0.025
Q5_6Pre	After your study abroad experience will you apply for international jobs?	2.633	0.008
Q5_1	Because of your study abroad experience, will you take a foreign language beyond the minimum requirement?	2.608	0.009
Q5_6	Because of your study abroad experience, will you apply for global jobs (international or domestic)?	2.135	0.033
Q54	The language level I had prior to going abroad was an appropriate level for this program	3.132	0.002
Q60_2	How often did you interact with local students/friends (from the host country)?	2.608	0.009
Q60_3	How often did you interact with international students/friends (from other countries)?	3.028	0.002
Q61_2	I am satisfied with my effort to integrate myself with the local culture.	1.734	0.083
Q48_2	The housing was appropriate based on local standards.	1.783	0.075
Q1Pre_5	Why did you choose this specific program? Recommended by faculty or staff	-1.518	0.129
Q1Pre_7	Why did you choose this specific program? Other	-1.628	0.104
Q14Pre_1	What is your primary goal for studying abroad? Satisfy degree requirements	1.794	0.073
Major1_Economics	Economics major	3.577	0.000
Major1_Geological sciences	Geological sciences major	-1.85	0.064
Major1_Environmental science and studies	Environmental sciences and studies major	-2.245	0.025
Major1_Archaeology	Archaeology major	-1.986	0.047
Major1_Mathematics	Mathematics major	1.737	0.082
Major1_Statistics and analytics	STAN major	1.625	0.104
Major1_Health policy and management	HPM major	1.959	0.050
Major1_Neuroscience	Neuroscience major	2.468	0.014
Region_Australia/Pacific Islands	Region – Australia/Pacific Islands	-2.018	0.044
Region_Middle East	Region – Middle East	-1.871	0.061
Q47_A home-stay	Housing – A home-stay	1.549	0.121

Table 10. Summary of coefficients from stepwise selection on POLR model with Q10 as the response.

The most significant drivers of a student feeling confident in their ability to interact with people of other languages and cultures after studying abroad are:

- Deciding to apply for international jobs after studying abroad
- Deciding to take a foreign language beyond the minimum requirement after studying abroad
- Feeling that the language level they had prior to the program was appropriate
- Interacting often with local students/friends
- Interacting often with international students/friends (from other countries)
- Being an Economics major

Response: Q13 I adapt well to new experiences and changing circumstances.			
Coefficient	Question	Value	p value
Q5_6Pre	After your study abroad experience will you apply for international jobs?	1.975	0.048
Q5_6	Because of your study abroad experience, will you apply for global jobs (international or domestic)?	3.702	0.00
Q45	The program-led excursions enhanced my classroom learning and/or contributed to my cultural learning.	2.315	0.021
Q60_2	How often did you interact with local students/friends (from the host country)?	1.749	0.080
Q60_5	How often did you interact with other Americans (not from UNC)?	2.039	0.041
Q61_2	I am satisfied with my effort to integrate myself with the local culture.	3.389	0.001
Q61_3	I am satisfied with the quality of contact I had with the local culture.	1.776	0.076
Q48_1	I felt safe in my housing.	1.53	0.126
Q82PersonalSavings	How did you pay for your study abroad program?	3.276	0.001
Q82FinancialAid	How did you pay for your study abroad program?	-2.049	0.040
Q1Pre_6	Why did you choose this specific program? Program Length	-3.021	0.003
Q1Pre_5	Why did you choose this specific program? Recommended by faculty or staff	-2.148	0.032
Major1_Media and journalism	Media and journalism major	-2.048	0.041
Major1_Biomedical and health sciences and engineering	BME major	-2.612	0.009
Q90_Service-learning	Did your program include a service-learning component?	-2.832	0.005
Region_Australia/Pacific Islands	Region – Australia/Pacific Islands	1.789	0.074

Table 11. Summary of coefficients from stepwise selection on POLR model with Q13 as the response.

The most significant drivers of a student having the ability to adapt to new experiences and changing circumstances after studying abroad are:

- Deciding to apply for international jobs because of studying abroad
- Feeling satisfied with their effort to integrate into the local culture
- Paying for the program with personal savings
- Choosing the program based on its length
- Being a BME major
- Participating in a program with a service-learning component

Response: Q7 I am independent and self-confident.			
Coefficient	Question	Value	p-value
Q5_6	Because of your study abroad experience, will you apply for global jobs (international or domestic)?	2.36	0.018
Q45	The program-led excursions enhanced my classroom learning and/or contributed to my cultural learning.	3.417	0.001
Q60_5	How often did you interact with other Americans (not from UNC)?	3.124	0.002
Q61_3	I am satisfied with the quality of contact I had with the local culture.	2.283	0.022
Q1Pre_6	Why did you choose this specific program? Program Length	-2.464	0.014
Major1_Chemistry	Chemistry major	-1.726	0.084
Major1_Archaeology	Archaeology major	-2.96	0.003
Q47_Other	Housing: other	-1.86	0.063

Table 12. Summary of coefficients from stepwise selection on POLR model with Q7 as the response.

The most significant drivers of a student feeling a high sense of independence and self-confidence after studying abroad are:

- Feeling that the program-led excursions enhanced their learning
- Interacting with non-UNC Americans often
- Being an archaeology major (negative)

4.6 Predicting Difference in Score

The final response variables of interest were the differences in Q10, Q13, and Q7 from before studying abroad to after. Proportional odds logistic regression was again used to investigate the drivers of such differences, with a similar variable selection process as in the raw score cases.

Variable Selection

Examining the distributions of the responses led to a slightly different setup for the variable selection classification model. Histograms of the three responses are shown below.

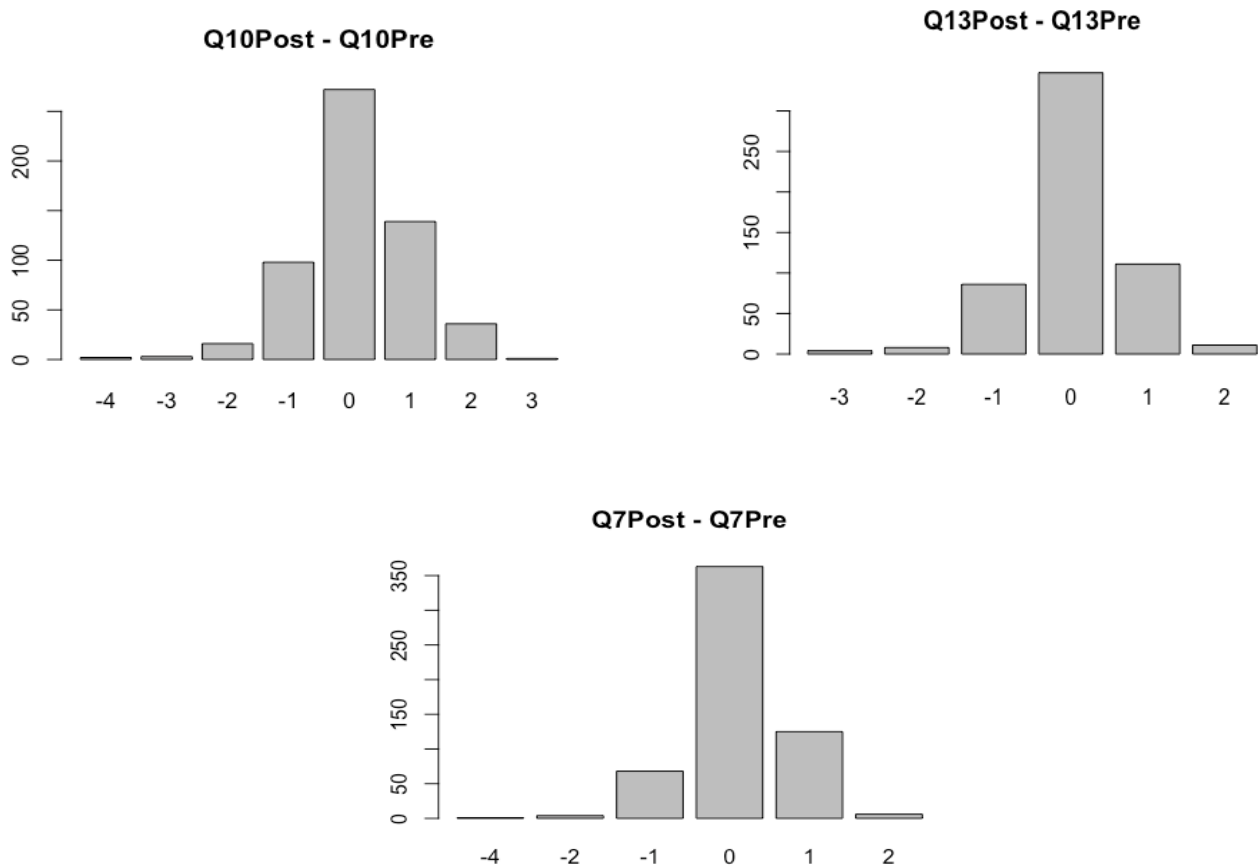


Figure 13. Distributions of the differences of Q10, Q13, and Q7 from before to after studying abroad.

The histograms of the differences in answers to the response questions show that there are few students that experienced a change of answer by more than 1 point, and most students in fact reported the same answer both pre- and post- studying abroad. Thus, to investigate drivers of positive or negative difference, a multinomial classification regression was fit to the response of a negative difference, 0 difference, or a positive difference. Cross-validation was used to determine the optimal lambda penalty parameter, and the alpha Elastic Net parameter was set to 0.5. In order to produce the same non-zero coefficients for each class of the response, the argument “grouped” was used in the model. This, however, produced less sparse coefficients, and a decision had to be made regarding how many and which coefficients to use as predictors in the ordinal regression model, the trade-off being between a well fit model (more predictors) and a converging, interpretable model (fewer). The predictors ultimately chosen are displayed in the following table, along with the large model’s goodness of fit p-values, the variables selected by stepwise selection, and that model’s goodness of fit p-values.

Response	POLR Predictors	Lipsitz test p-value	HL test p-value	Stepwise predictors	Lipsitz test p-value	HL test p-value
Difference in Q10	Q5_1Pre Q53 Major1_Public policy Q1Pre_1 Q5_11 Q2GenEdPre Q3WebsitePre Q54 Q2ElectivePre Q56_2 Q60_6 Q3AdvisorPre Region_South America Q1Pre_7 Q42 Major1_Management and society Q14Pre_7 Q4MeetingPre Region_Africa Major1_Economics Q56_1 Major1_Archaeology Q2Minor Q5_2 Major1_Statistics and analytics Major1_Philosophy Q47_Other Major1_Psychology Q40 Q5_8Pre Region_Middle East Major1_Anthropology Q73 Major1_Biology Q4PeerPre Major1_Nursing Q5_5 Q4101Pre Major1_Exercise and sportscience Q3PrintMaterialsPre Q14Pre_3 Major1_Peace, war and defense	0.8058	0.4181	Q5_1Pre Q53 Q1Pre_1 Q5_11 Q3WebsitePre Q2ElectivePre Q56_2 Q3AdvisorPre Q1Pre_7 Q42 Major1_Management and society Q14Pre_7 Major1_Economics Q56_1 Q5_2 Major1_Statistics and analytics Major1_Philosophy Q47_Other Q40 Region_Middle East Q5_5 Major1_Exercise and sport science	0.9765	0.2891
Difference in Q13	Q4WebsitePre Q5_1Pre Q90_Internship Q61_2 Q14Pre_4 Q45 Major1_Russian language and culture Major1_Biomedical and health sciences engineering Major1_Health policy and management Major1_Exercise and sport science Q56_2	0.9527	0.6623	Q4WebsitePre Q90_Internship Q61_2 Q45 Major1_Russian language and culture Major1_Biomedical and health sciences engineering Major1_Health policy and management	0.6327	0.9651

	Major1_Anthropology Q47_Independently arranged housing Q5_2Pre Q82Scholarship Region_Asia Q2GenEd Q2MinorPre Q61_3 Q83 Q47_A home-stay Q73 Major1_Communication studies Q71 Q3PrintMaterialsPre Q2LanguagePre Region_Middle East			Q56_2 Q2GenEd Q2MinorPre Q83 Major1_Communication studies		
Difference in Q7	Q14Pre_4 Major1_Chemistry Major1_Nursing Q45 Q60_2 Q14Pre_0 Q60_5 Q5_5Pre Q5_2Pre Q47_Other Major1_Nutrition Q5_11 Q4PersonalResearchPre Major1_Computer Science Q82Scholarship Q90_Major1_Communication studies Q1Pre_5 Q90_Internship Q1Pre_7 Q3AdvisorPre Q48_2 Major1_Biomedical and health sciences engineering Q5_11 Q4HandbookPre Q54 Q48_1 Q3FairPre Major1_Neuroscience	0.2509	0.4009	Q14Pre_4 Major1_Chemistry Q45 Q5_2Pre Q47_Other Major1_Nutrition Q5_11 Q4HandbookPre Q54 Major1_Neuroscience	0.5617	0.6426

Table 13. The predictors, used in the POLR model and selected after stepwise selection, with their corresponding goodness of fit p-values, for each response.

The summaries of the stepwise coefficients are displayed in the three tables below. The most significant variables – those with p-values less than 0.01 – are highlighted, although there are many more with p-values below 0.05 and 0.1, indicating they are also significant even if less so.

Response: Difference in Q10 I am confident in my ability to communicate with people of other languages and cultures.			
Coefficient	Question	Value	p value
Q5_1Pre	After your study abroad experience will you continue taking a foreign language beyond the minimum requirement?	1.823	0.068
Q53	Did you study in a non-English speaking country?	4.634	0.000
Q1Pre_1	Why did you choose this specific program? Location	-1.925	0.054
Q5_11	Because of your study abroad experience will you change your career choice?	3.796	0.000
Q3WebsitePre	How did you find out about your specific study abroad program? UNC study abroad website	2.934	0.003
Q2ElectivePre	What kind of credit do you intend to pursue? Elective	1.945	0.052
Q56_2	While abroad how often did you utilize the language of your host country with program participants?	1.558	0.119
Q3AdvisorPre	How did you find out about your specific study abroad program? Study Abroad Advisor	-2.168	0.030
Q1Pre_7	Why did you choose this specific program? Other	-2.179	0.029
Q42	Was this program a good match for the credits you pursued?	2.32	0.020
Major1_Management and society	Management and society major	-1.958	0.050
Q14Pre_7	What is your primary goal for studying abroad? Other	-2.189	0.029
Major1_Economics	Economics major	1.433	0.152
Q56_1	While abroad how often did you utilize the language of your host country with host family members?	2.662	0.008
Q5_2	Because of your study abroad experience will you use this experience for Honors Thesis or Capstone?	-1.521	0.128
Major1_Statistics and analytics	STAN major	1.45	0.147
Major1_Philosophy	Philosophy major	-2.757	0.006
Q47_Other	Housing Other	-2.657	0.008
Q40	Did you satisfy your goals?	2.011	0.044
Region_Middle East	Region – Middle East	-2.533	0.011
Q5_5	Because of your study abroad experience will you highlight study abroad experience on applications and in interviews?	2.326	0.02
Major1_Exercise and sport science	EXSS major	1.774	0.076

Table 14. The predictors retained after stepwise selection on the POLR regression and their t-values and p-values for response Q10Post – Q10Pre.

The most significant drivers of a student increasing their confidence of their ability to communicate with people of other languages and cultures are:

- Studying abroad in a non-English speaking country
- Changing their career choice because of their study abroad experience
- Finding their program on the UNC Study Abroad website
- Utilizing the language of the host country with the host family
- Being a philosophy major
- Living in a housing option other than those included on the survey while abroad

Response: Difference in Q13 I adapt well to new experiences and changing circumstances.			
Coefficient	Question	Value	p-value
Q4WebsitePre	What resources have you utilized up to this point? UNC study abroad website	-2.197	0.028
Q90_Internship	Did your study abroad program include any of the following components? Internship	3.721	0.000
Q61_2	I am satisfied with my effort to integrate myself with the local culture.	2.792	0.005
Q45	The program-led excursions enhanced my classroom learning and/or contributed to my cultural learning.	3.263	0.001
Major1_Russian language and culture	Russian language and culture major	2.853	0.004
Major1_Biomedical and health sciences engineering	BME major	-2.054	0.040
Major1_Health policy and management	HPM major	3.112	0.002
Q56_2	While abroad how often did you utilize the language of your host country with program participants?	1.608	0.108
Q2GenEd	What kind of credit did you pursue abroad? General Education requirement	-1.417	0.157
Q2MinorPre	What kind of credit did you pursue abroad? Minor	1.684	0.092
Q83	My program's UNC Study Abroad budget sheet was a helpful guide for planning my study abroad-related finances.	1.818	0.069
Major1_Communication studies	Communications major	-2.402	0.016

Table 15. The predictors retained after stepwise selection on the POLR regression and their t-values and p-values for response Q13Post – Q13Pre.

The most significant drivers in students increasing their ability to adapt to new and changing circumstances are:

- Participating in an internship while abroad
- Feeling satisfied with their effort to integrate with the local culture
- Feeling that the program-led excursions enhanced their learning
- Being a Russian language and culture major
- Being a health policy and management major

Response: Difference of Q7 I am independent and self-confident.			
Coefficient	Question	Value	p value
Q14Pre_4	Primary goal for studying abroad is to take courses not available at UNC or different from UNC	-1.421	0.155
Major1_Chemistry		-3.676	0.000
Q45	The program-led excursions enhanced my classroom learning and/or contributed to my cultural learning.	3.098	0.002

Q5_2Pre	After your study abroad experience will you use this experience for honors thesis or capstone?	-1.889	0.059
Q47_Other	Housing Other	-2.243	0.025
Major1_Nutrition	Nutrition major	2.762	0.006
Q5_11	Because of your study abroad experience will you change your career choice?	2.457	0.014
Q4HandbookPre	What resources have you utilized up to this point? Heels Abroad Handbook	-2.309	0.021
Q54	The language level I had prior to going abroad was an appropriate level for this program	-2.517	0.012
Major1_Neuroscience	Neuroscience major	2.453	0.014

Table 16. The predictors retained after stepwise selection on the POLR regression and their t-values and p-values for response Q7Post – Q7Pre.

The most significant drivers of a student increasing their feeling of independence and self-confidence from studying abroad are:

- Being a chemistry major
- Feeling that the program-led excursions enhanced their learning
- Being a nutrition major

5. Conclusion

Both sets of data showed that the experiences students have while abroad regarding immersion in the local culture, speaking a foreign language, and staying with a host family have significant impact on their confidence in their ability to communicate with people of other cultures, to adapt well to new and changing circumstances, and to notice a shift in their perspective of their own culture. However, other factors which might not be as significant are still important to note.

Additionally, the Summer cohort showed that studying abroad was a significant effect on students' answers to the self-reflection questions.

However, many variables showed evidence of being statistically significant drivers of the response questions, and many factors must come into play when determining a student's growth while abroad.

Appendix - List of survey questions used as potential predictor/response variables.

1. Post-Program (Spring and Summer)

Q58 Major

Region Region

Q1 Why did you choose this specific program?

Q2 What kind of credit do you intend to pursue?

Q3 How did you find out about your specific study abroad program?

Q4 What resources have you utilized up to this point?

Q5 After your study abroad experience, will you...?

Continue taking a foreign language beyond the minimum requirement (1)

Use this experience for honors thesis or capstone (2)

Highlight study abroad experience on resume and in interviews (5)

Highlight study abroad experience on essays for graduate or professional school applications (7)

Apply for international jobs (6)

Apply for domestic jobs with a global component (8)

Q7 I am independent and self-confident

Q8 I am confident in my problem-solving skills.

Q9 I am comfortable navigating a predominantly foreign language environment.

Q10 I am confident in my ability to communicate with people of other languages and cultures.

Q11 I am comfortable with the level of my foreign language skills for my study abroad destination.

Q12 I am confident with my interpersonal communication skills.

Q13 I adapt well to new experiences and changing circumstances.

Q14 What are your goals for studying abroad?

2. Pre-Departure (Summer only)

Q58 Major

Q90 Did your study abroad program include any of the following components?

Region Region

Q14 What were your goals for studying abroad?

Q40 Did you satisfy your goals?

Q5 Because of your study abroad experience, will you...?

Take a foreign language beyond the minimum requirement (1)

Use this experience for Honors Thesis or Capstone (2)

Change or add major/minor (9)

Highlight study abroad experience on

applications and in interviews (5)

Apply for global jobs (international or domestic) (6)

Change your career choice (11)

Q2 What kind of credit did you pursue abroad?

Q42 Was this program a good match for the credits you pursued?

Q45 The program-led excursions enhanced my classroom learning and/or contributed to my cultural learning.

Q53 Did you study in a non-English speaking country?

Q54 The language level I had prior to going abroad was an appropriate level for this program?

Q56 While abroad how often did you utilize the language of your host country with the following people?

Host family members (1)

Program participants (2)

Local students (from the host country) (3)

Community members (4)

Q60 How often did you interact with the following people while abroad?

Host family members (1)

Local students/friends (from the host country) (2)

International students/friends (from other countries) (4)

Other UNC students (6)

Other Americans (not from UNC) (5)

Local program staff (3)

Q61 Using the scale below, respond to the following statements about the intercultural interactions you had abroad.

I am satisfied with the amount of cultural interaction made available by the program. (1)

I am satisfied with my effort to integrate myself with the local culture. (2)

I am satisfied with the quality of contact I had with the local culture. (3)

Q7 I am independent and self-confident.

Q8 I am confident in my problem-solving skills.

Q9 I am comfortable navigating a predominantly foreign language environment.

Q10 I am confident in my ability to communicate with people of other languages and cultures.

Q12 I am confident with my interpersonal communication skills.

Q13 I adapt well to new experiences and changing circumstances.

Q65 I have noticed a shift in my perspective of my own culture.

Q66 I have noticed a shift in my perspective of other cultures

Q69 While abroad, did you encounter any challenges related to your identity?

Q47 While abroad, I lived in :

Q48 Using the scale provided, respond to the following statements about your housing abroad.

I felt safe in my housing. (1)

The housing was appropriate based on local standards. (2)

The housing provided the basic needs to learn and study. (3)

I would recommend this housing to future participants. (4)

Q82 How did you pay for your study abroad program?

Q83 My program's UNC Study Abroad budget sheet was a helpful guide for planning my study abroad-related finances.

Q71 The Heels Abroad Handbook helped me prepare for studying abroad.

Q73 The Study Abroad Pre-Departure Conference/Orientation helped me prepare for studying abroad.