



Experiences of Non-Native English Speakers Learning Computer Science in a US University

Carmen Nayeli Guzman*
UC San Diego
cnguzman@ucsd.edu

Anne Xu*
UC San Diego
anxu@ucsd.edu

Adalbert Gerald Soosai Raj
UC San Diego
gerald@eng.ucsd.edu

ABSTRACT

Prior work has shown that the adjustment challenges faced by international students in English-speaking universities are primarily attributable to English language proficiency and culture. Although studies in computing education have documented the barriers that non-native English speakers face while learning CS using online programming platforms, little research has been done on their experiences in CS courses within a university setting. In this study, we seek to understand the experiences of non-native English speakers learning CS at a English-speaking university in the United States. Through weekly surveys and course grades from four CS courses, we investigate the differences between non-native and native English speakers with respect to the time spent studying outside of class, confidence levels at the start of a course, types of learning resources used, and expected and actual overall grades. We found that although non-native speakers require a higher minimum grade to be satisfied and often spend a longer time on average studying outside of class compared to native speakers, there was no difference in the actual grade received between the two groups. While there was no significant difference in the type of resources used (online vs in-person), there was a significant difference in the confidence level between native and non-native English speakers.

CCS CONCEPTS

• Social and professional topics → Computing education programs.

KEYWORDS

Non-native English Speakers; International Students; English-speaking University; Computer Science Education; Confidence Level; Studying Habits; Academic Performance; Academic Expectation

ACM Reference Format:

Carmen Nayeli Guzman*, Anne Xu*, and Adalbert Gerald Soosai Raj. 2021. Experiences of Non-Native English Speakers Learning Computer Science in a US University. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)*, March 13–20, 2021, Virtual Event, USA. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3408877.3432437>

*These authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution International 4.0 License.

SIGCSE '21, March 13–20, 2021, Virtual Event, USA.

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8062-1/21/03.

<https://doi.org/10.1145/3408877.3432437>

1 INTRODUCTION

Computer Science (CS) has become an essential skill taught in universities around the world but the most popular programming languages are in English as well as the majority of their documentation (e.g., Java docs) and online resources (e.g., Stack Overflow) [3, 16]. Consequently, this has made CS an English-centric discipline where the varying levels of English proficiency that exists among non-native English speakers become more pronounced [18, 25].

U.S. universities benefit educationally, economically, and culturally, from their international students, many of them being non-native English speakers [2]. In U.S. institutions of higher education, 5.3% of all students are international, while 18% of those international students are studying math/computer science [35]. Math/computer science is also the second most popular field among international students in the U.S. [35]. Research has shown that non-native English-speaking students generally face cultural and language barriers that impede their education and report higher levels of depression and anxiety, while some also face racial barriers that impact their experience studying in the U.S. [11, 14, 14, 41, 43, 49].

Although the barriers faced by non-native English speakers studying CS in countries where English is not the native language have been well studied [27, 31, 36–39, 46], there is a lack of research addressing the learning experiences of non-native English speakers learning CS in English-speaking countries around the world. This study addresses this gap and identifies differences between **Native English Speakers (NES)** and **Non-Native English Speakers (NNES)** in a U.S. university in order to further efforts to create a more inclusive and diverse computing education environment. Our research tries to answer the following questions:

- (1) *Is there a difference in starting confidence level between NNES and NES?*
 - (a) *Is there a difference in the starting confidence level between female and non-binary vs male students?*
 - (b) *Is there a difference in the prior CS experience between NNES and NES?*
- (2) *Is there a difference in the average hours spent studying outside of class in a CS course between NNES and NES?*
- (3) *Is there a difference in the minimum grade needed to be satisfied in a CS course between NNES and NES?*
- (4) *Is there a difference in the overall grade received between NNES and NES?*
- (5) *Is there a difference in the types of resources (online vs. in-person) used by NNES and NES?*

This study examines prior work by investigating the desired academic grades, personal time-investment in education, and CS academic self-confidence of NNES as they embark upon introductory level CS courses at an English-speaking university in the U.S. The following are our study's contributions to computing education research: 1) we show how NNES perform academically, and their

use of CS learning resources compared to native English speakers, 2) we present a refined understanding of the study habits of NNES, with some analyses performed on a weekly basis as students progress through introductory CS courses.

2 RELATED WORK

Phillip Guo conducted an extensive qualitative study to investigate the barriers and desires of non-native English speakers world-wide, learning programming via an online platform [15]. Based on the findings, several suggestions are made to help improve CS learning resources, including the move towards a universal framework. Non-native English speakers who were not fluent in English faced difficulties in many aspects of learning programming and producing written code. In Guo's study, non-native English speakers suggest improvements to existing CS resources to be more accessible to people from different language and cultural backgrounds. In our study, surveys were conducted at a U.S. university's introductory CS courses which were taught in-person.

Sümer et al. examined several variables including gender, English fluency, and social support, as predictors of depression and anxiety among international students [41]. The study included 440 self-selected international students including undergraduates, graduate, and doctoral students, from two college campuses in the eastern part of the U.S. This study reported that international students with lower levels of English proficiency reported higher levels of depression and anxiety and that increased social support resulted in decreased levels of depression and anxiety. It was also reported that among this group of international students, depression and anxiety had no relation to gender. In our study, we examine the factors, including prior experience, gender, and English fluency, that contribute to self-confidence among undergraduate students at a U.S. university in the western part of the U.S.

Soosai Raj et. al. conducted a series of experiments to measure the impact of bilingual CS education on student learning in CS courses for non-native English speakers in Tamil Nadu, India [36–39]. They found that although the use of students' native language increased classroom engagement and participation, teaching programming using Tamil and English was no different than teaching programming using only English with respect to student learning measured using test scores. We believe our study adds further insights on the relationship between academic performance and English proficiency by comparing the CS course grades of native and non-native English speakers over 10 weeks at a research-intensive U.S. university.

3 METHODOLOGY

Survey data used in this study was acquired from four lower division computer sciences courses taught at a research-intensive university in the United States. The courses are 1) CS1 – the first course in a two part sequence of an introduction to CS, 2) CS1.5 – the second course in the two part intro sequence, 3) CS1A – an accelerated introduction to CS, and 4) CS2 – an introduction to basic data structures. All these courses are taught using Java. The CS1 and CS1.5 courses combined consist of the same material covered in CS1A. Students taking CS1 are typically students with little to no prior CS background, while students taking CS1A are typically more comfortable with CS material and might already have some prior background in CS.

3.1 Collection of Student Survey Data

Informed consent was obtained from all students who participated in this study. Students enrolled in each course completed a pre-survey (<http://bit.do/nnespresurvey>) at the beginning of the quarter, providing information related to their native language status, demographics, and prior CS experience. Students also completed an online weekly reflection survey (<http://bit.do/weeklyreflection>) which asked them to reflect upon their learning experience for that week, resources used to assist them in their learning, and hours spent studying outside of class. To encourage student participation in the study, each weekly survey response was worth a small fraction of the grade for a weekly programming assignment. A total of 8-10 weekly reflection surveys were completed over the course of a quarter, depending on the specific course.

Native Language Status. In the pre-survey, students were asked, "What is your fluency with English?" The options were (1) I learned English *as a child* in a bilingual or multilingual setting. I am equally comfortable in English and another language, (2) I learned English *as a child* as my primary language, (3) I learned English *after* learning another primary language and I am equally comfortable in both languages, and (4) I learned English *after* learning another primary language and I am not as comfortable in English as I am in another language. Students who learned English *as a child* (i.e., selected either (1) or (2)) were categorized as native English speakers, and those who learned English *after* learning a different language (i.e., selected either (3) or (4)) were categorized as non-Native English speakers. Students whose native language status was undetermined, as in the case where students left this response blank, were removed from the study.

Confidence Level. The survey question related to academic self-confidence was phrased in the following manner: "On a scale of 1-5, how confident are you in your ability to do well in this course?" Self-efficacy is a judgement of capability of achieving a certain outcome [5]. Although our question is related to self-efficacy in CS [20, 24], we will refer to it as a question of self-confidence because "self-confidence" is a more familiar term [12].

Prior Experience. In the pre-survey, students reported their relevant prior programming experience as well as their prior CS experience. The prior programming experience question asked them to rate their programming experience on a scale of 1-5, while the prior CS experience question asked them to select previous CS and math related course work, including extracurricular CS related activities. In prior CS experiences, AP CS A [9] and college level CS courses taken by the students in college or high school, were considered to be prior CS experience. These were the only courses considered for prior CS experience because it is known that AP CS A is taught in Java, and obviously advantageous for students taking the four courses in our study, which are also taught in Java. Any other college level CS course experience was considered rigorous, and thus, also considered to be prior CS experience.

Resources. To determine if there was a correlation between the students' native language status and any particular learning resource (i.e., discussion board, tutor, TA, etc.), we used responses to a question from the weekly reflection survey that asked students, which resources, if any, did they seek help from to complete work

for their CS course during that week. To determine if there is a difference in the type of resources used by NES and NNES, the resources used by students were placed into two categories – online and in-person. The online resources consist of discussion board (e.g., Piazza [33]), online programming resources such as Java documentation [29], Q&A websites like Stack Overflow [4], etc. The in-person resources are those requiring live verbal communication in English with teaching assistants (TA), tutors, peers in the same course, and the instructor.

3.2 Data Analysis

We used the following packages for data analysis: Pandas [26], NumPy [45], SciPy [47], Matplotlib [17], Seaborn [48], and Pingouin [44]. Our data analysis is available at <https://bit.ly/nnesdata>.

Confidence Level Data. For each class, students who did not respond to the confidence level question were excluded from the analysis. Similarly, students who did not report their gender or prior experience were dropped from our analysis of the effect of gender and prior experience on confidence level (see Section 4.1).

Hours Data. In each weekly survey, students were asked ‘In the last week, approximately how many hours did you spend outside of class time working on work for this course?’. As these surveys were self-reported data, sometimes students seemed to report an inaccurate number of hours for a certain week (e.g., 200 hours). To address this, we identified outliers for every week by identifying any data point more than 1.5 Inter-Quartile Range (IQR) below the first quartile or above the third quartile. We found that outliers were usually above 25 hours based on the IQR definition for most weeks and so to establish a baseline, after discussion among the research team, we decided to label reported hours above 25 as outliers, and replaced every outlier with the median value of the hours reported for that week.

Overall Grades Data. For each course, we used grades of students who filled out the pre-survey and completed the course. We dropped the data of students that did not do either of these things. For students who dropped out of a course at any point during the quarter, we dropped them out of the dataset when calculating overall grades for native and non-native English speakers. Although for certain courses the assumption of normality of data was not fulfilled, we still used t-tests for inferential statistics as t-tests do not require assumptions of normal distribution in sufficiently large samples [23].

Resources Data. About 30% of the students per week left the resources question blank, and this data was missing at random (MAR) [32]. In order to reduce bias in the data set that would occur if we dropped the missing data, we choose to use probabilistic imputation drawing from the empirical distribution of the resources used per week to lessen bias while also keeping variance of the dataset [7].

Inferential Statistics. In our data analysis, to find if there were any statistically significant difference between native and non-native English speakers, we used t-test and Mann-Whitney U-test as appropriate.

For each analysis where a 2-sided independent samples t-test was used, the means were determined with an alpha value of 0.05 for

statistical significance. The following assumptions were checked: (1) The Central Limit Theorem (CLT) applies to each sample individually, as sample sizes are both greater than 30 and QQ plots showed normality. (2) The means of both the groups are random samples that are independent of each other. (3) The variances of the confidence level values are approximately equal (for each course, a Levene’s test showed the variances were not significantly different ($p > .05$), indicating the two variances are approximately equal). For each analysis where a 2-sided Mann-Whitney U test was used, the assumptions of 1) a dependent variable measured on an ordinal or continuous scale, 2) an independent variable consisting of two categorical variables, 3) independent groups, and 4) observations that have the same shape were fulfilled. We measured the effect size of t-test using Cohen’s d and Mann-Whitney U-test using rank-biserial correlation (RBC). For the analysis where the multiple linear regression model was produced, the assumptions of the model were tested.

4 RESULTS

4.1 RQ 1: Is there a difference in confidence level between NNES and NES?

The difference in the confidence level of non-native English speakers (NNES) and native English speakers (NES) was determined for each of the four courses. Since the data is ordinal, a Mann-Whitney U-test was performed to determine if the difference in confidence level is statistically significant between the two groups. The results are summarized in Table 1 and includes all students who filled out the pre-survey at the beginning of the course.

Course	Pop.	N	M	Mdn	SD	U	p	r
CS1	NNES	259	3.2	3.0	1.0	37619	.001	< -.2
	NES	366	3.6	4.0	0.9			
CSE1.5	NNES	50	3.14	3.0	0.9	2040	.3019	0.05
	NES	86	3.07	3.0	1.2			
CS1A	NNES	137	3.7	4.0	1.0	25694.5	.034	-0.12
	NES	336	3.9	4.0	0.9			
CS2	NNES	110	1.0	3.3	3.0	7742.5	.407	-0.06
	NES	133	1.0	3.4	4.0			

Table 1: Confidence levels of NES and NNES at the start of each course.

Pop. - Population; N - Number of students who took the pre-survey in each group; M - Mean; Mdn - Median; SD - Standard Deviation; U - Mann-Whitney U statistic; p - p-value; r - Rank-biserial correlation effect size

In both CS1 and CS1A, the mean confidence level of NNES was lower than NES and a Mann-Whitney U-test showed that the difference was significant. In CS 1.5, a U test showed that the distribution of the confidence level among each group was not statistically significant, although the mean confidence level of NNES was higher. In CS2, the mean confidence level of NNES was lower than NES but the difference was not statistically significant.

While a focus of this work is to determine the influence of native language status on the observed differences in confidence level

between NES and NNES, we recognize and address other important factors that contribute to confidence level. Prior studies have shown that gender and prior academic achievement are factors correlated with academic self-efficacy and confidence [22, 28, 34]. Below we show our results on the effects of native language, gender, and prior experience in CS1 and CS1A (two courses where NNES and NES reported significantly different confidence levels - see Table 1).

To determine the influence of native language status on the confidence levels by controlling for the influence of gender or self-reported prior experience, an ANCOVA analysis was attempted [6]. In both CS1 and CS1A, the homogeneity of regression slopes assumption for the ANCOVA analysis was not fulfilled. Therefore, a multiple linear regression model was produced. In our regression analysis, although the assumption of homoscedasticity for multiple linear regression was moderately violated, the predictor variables may still be evaluated for significance [21].

In CS1 the regression model with a coefficient of determination, $R^2 = 0.19$ shows that native language status ($\beta = 0.25$, $t(627) = 3.47$, $p < .001$), prior experience ($\beta = 0.31$, $t(627) = 8.48$, $p < .001$), and gender ($\beta = 0.40$, $t(627) = 5.52$, $p < .001$) are significant predictors of confidence level. In CS1A the regression model with a coefficient of determination, $R^2 = 0.29$ shows the effect of native language status on confidence level was not significant ($\beta = 0.04$, $t(469) = 0.48$, $p = .63$), while prior-experience ($\beta = 0.42$, $t(469) = 11.85$, $p < .001$), and gender ($\beta = 0.25$, $t(469) = 2.96$, $p = .003$) were significant predictors of confidence level.

In the CS1 and CS1A courses, the mean confidence level was higher for NES, and a U-test showed that the distribution of the confidence level among each group was significantly different. The multiple linear regression models showed that gender and prior experience contributed significantly to the confidence levels of students in CS1 and CS1A, while native language status was found to be a significant predictor of confidence level in CS1 only. This raises the question of whether there is a significant difference in the prior CS experience of NNES and NES and warrants further analysis of the confidence levels of female and non-binary vs. male students in the courses.

To assess the difference in overall prior CS-related experience of NNES vs NES, we compared their self-reported prior experience, and their calculated prior CS experience. The calculated prior CS experience was based on having previously taken AP CS A [9], and any other college level CS course in the past. The results of the self-reported prior experience analysis for both courses is summarized in Table 2.

Table 2: Self-reported prior experience of NES and NNES in CS1 and CS1A.

Course	Pop.	N	M	Mdn	SD	U	p	r
CS1	NNES	259	1.9	2.0	0.9	54588.5	<.0001	-0.15
	NES	368	2.1	2.0	1.0			
CS1A	NNES	137	3.0	3.0	1.1	26509.5	.01	-0.14
	NES	338	3.3	3.0	1.0			

N - Number of students who took the pre-survey in each group

The results from Table 2 show that for both CS1 and CS1A, the mean prior experience was greater for NES compared to NNES, and a U-test showed that the distributions of the self-reported prior experience for each group was significantly different. In both CS1 and CS1A, the mean calculated CS experience was greater for NES than NNES, and a Mann-Whitney U-test showed that the calculated prior CS experience of NES and NNES is significantly different in both CS1 ($U = 43715.5$, $p = .02$) and CS1A ($U = 19076.5$, $p = < .001$).

The influence of gender on the confidence level of all students in each of the four courses was assessed with a Mann-Whitney U-test, and the rank-biserial correlation coefficient, r . The results of this analysis are summarized in Table 3.

Table 3: Confidence levels of males (ML) and females + non-binary (FL + NB) at the start of each course

Course	Pop.	N	M	Mdn	SD	U	p	r
CS1	ML	392	3.6	4.0	0.9	30066	<.001	.32
	FL+NB	230	3.1	3.0	1.0			
CSE1.5	ML	84	3.3	3.0	1.2	1547.5	.014	.23
	FL+NB	49	2.8	3.0	0.8			
CS1A	ML	359	4.0	4.0	0.9	11735.5	<.001	.37
	FL+NB	107	3.4	3.0	0.8			
CS2	ML	164	3.5	4.0	0.9	4497.5	<.001	.28
	FL+NB	78	3.0	3.0	1.0			

N - Number of students who took the pre-survey

From Table 3 we can see that the confidence level for males was higher than females and non-binary students, and the difference was statistically significant in all four courses. Using Kerby simple difference formula [19] for rank-biserial correlation coefficient r , we can see that in CS1A (as an example), around 69% of the data supports the hypothesis that males have a higher confidence level than females and non-binary students.

To determine if there was a difference in the confidence levels of NES and NNES who *completed* the courses CS1 and CS1A, a Mann-Whitney U-test was performed. In CS1, NES continued to have a higher mean confidence level ($N = 142$, $M = 3.6$, $SD = 0.9$) compared to NNES ($N = 331$, $M = 3.3$, $SD = 1.0$), and the distributions of confidence level in NES and NNES were significantly different ($U = 18492.0$, $p = < .001$). In CS1A, NES continued to have a higher mean confidence level ($N = 171$, $M = 3.9$, $SD = 0.8$) compared to NNES ($N = 263$, $M = 3.8$, $SD = 0.9$) but the difference was not statistically significant ($U = 13930.0$, $p = .08$).

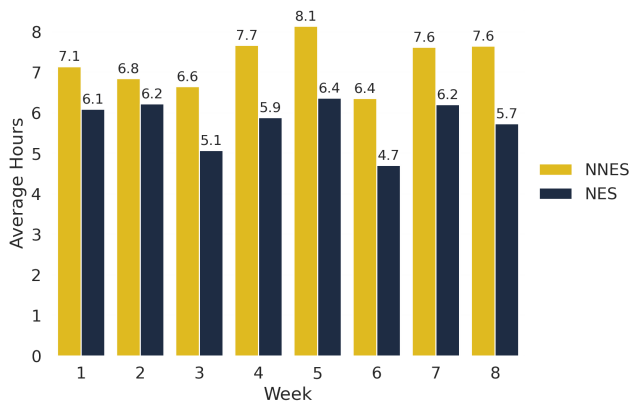
4.2 RQ 2: Is there a difference in the average hours spent studying outside of class between NNES and NES?

In all three classes that we had weekly survey data (except CS1.5 as it had only two weeks' data), non-native English speakers (NNES) reported a higher average number of hours spent outside the class when compared to native English speakers (NES). We performed an independent samples t-test and found that the difference between the two groups was statistically significant in all the three classes. We note that there is a small effect size for CS1, but the effect size is greater for CS1A and CS2. The results of average hours spent studying is reported in Table 4.

Table 4: Comparison of hours studied outside of class between NES and NNES

Course	Pop.	N	Mean	SD	SE	t	p	d
CS1	NNES	226	8.1	2.7	0.2	2.4	.02	0.2
	NES	305	7.6	2.8	0.2			
CS1A	NNES	104	7.5	4.0	0.4	3.4	<0.001	0.4
	NES	306	6.0	3.4	0.2			
CS2	NNES	110	12.5	3.4	0.3	2.9	.004	0.4
	NES	132	11.1	3.8	0.3			

N - Number of students who took the pre-survey in each group and completed at least one weekly survey; SE - Standard Error; *d* - Cohen's *d* effect size

**Figure 1: Average Hours Spent Studying Outside of Class for NES and NNES in CS1A**

From Figure 1, we can see that for every week, NNES spend more hours studying outside of class compared to NES.

4.3 RQ 3: Is there a difference in the minimum grade needed to be satisfied between NNES and NES?

In every course, students were asked at the start of the quarter 'What is the MINIMUM grade you think you would be satisfied with in this course?'. Their available choices to select were from A+ to D. To get the mean for letter grades, we converted the ordinal letter data to their numerical GPA values (e.g. A+ is 4.3, A is 4.0, A- is 3.7, etc.). Although A+ is equivalent to A in terms of GPA values at our university, in this study we used different values to differentiate A+ and A in our analysis. Out of the four classes, NNES had a higher minimum grade needed to be satisfied than NES for every class and for three classes (except CS1) the difference between the two groups was found to be statistically significant using a Mann-Whitney U-test. The results for minimum grade required to be satisfied are summarized in Table 5.

Table 5: Comparison of min. grade required to be satisfied for NES and NNES

Course	Pop.	N	M	Mdn	SD	U	p	r
CS1	NNES	260	3.6	3.7	0.52	43493	.044	.09
	NES	368	3.5	3.7	0.48			
CS1.5	NNES	53	3.5	3.7	0.58	1726	.011	.25
	NES	87	3.2	3.3	0.63			
CS1A	NNES	137	3.7	3.7	0.4	18874.5	<.001	.18
	NES	338	3.6	3.7	0.5			
CS2	NNES	112	3.6	3.7	0.5	5886	.0025	.22
	NES	134	3.4	3.7	0.6			

N - Number of students who took the pre-survey in each group

4.4 RQ 4: Is there a difference in the overall grade received by NNES and NES?

In three out of the four classes, NNES had a higher overall actual grade. We ran an independent samples t-test to compare the mean overall grades between NES and NNES, and found that in three out of four classes (CS1, CS1.5, CS2), NNES had a higher overall grade than NES that was not statistically significant. In CS1A NNES scored lower than NES and the difference was statistically significant. The results are summarized in Table 6.

Table 6: Comparison of overall actual grades received between NES and NNES

Course	Pop.	N	Mean	SD	SE	t	p	d
CS1	NNES	177	90%	0.1	0.009	-1.5	.123	0.2
	NES	266	88%	0.1	0.008			
CS1.5	NNES	41	76%	0.3	0.04	-1.0	.341	0.2
	NES	63	71%	0.3	0.03			
CS1A	NNES	136	78%	0.3	0.03	2.9	.005	0.3
	NES	331	87%	0.2	0.01			
CS2	NNES	105	80%	0.16	0.05	-1.9	.053	0.2
	NES	122	76%	0.19	0.007			

N - Number of students who took the pre-survey in each group and finished the course

4.5 RQ 5: Is there a difference in the types of resources used by NNES and NES?

We investigated the hypothesis that NNES might be hesitant to use in-person resources in general, such as teaching assistants and tutors due to language barriers, and that they would favor the online learning resources, such as language documentation and discussion boards. Students reported their use of resources on a weekly basis, and a Chi-Square independence test was performed on a weekly basis. The results of this analysis show that among the students in each of the four courses, the use of online or in-person resources is not dependent on native language status. For example, for a randomly selected week from the accelerated programming

course, CS1A, a chi-square test of independence showed that there was no significant association between native language status and learning resource preference, $X^2(2, N = 476) = 0.673, p = .41$.

5 DISCUSSION

In this study, we found that non-native English speakers (NNES) had significantly lower average confidence levels than native English speakers (NES) in two out of four courses. Our findings confirm the claim that non-native English speakers experience reduced confidence in an English-speaking academic setting [30].

Moreover, the differences in gender contributed significantly to confidence level, with male students reporting significantly higher confidence levels compared to female and non-binary students. Our findings match with the results of a previous study that found female students report much lower confidence that their male counterparts in Computer Science [8].

The difference in prior experience also contributed significantly to confidence level where more experienced students reported higher confidence levels when compared to less experienced students. This agrees with previous work which claims that academic self-efficacy is attributed to prior academic achievement [22].

When comparing the confidence levels of students who *completed* the course, we found that there was no significant difference in the confidence levels of NNES and NES in CS1A. But the confidence level of NNES was significantly lower than NES at the start of the course. Since CS1A is an accelerated course which requires some prior CS experience, one possible explanation for this could be that the students who did not have the necessary prior experience to succeed in the course, and might have had lower confidence levels, dropped the course.

In all four courses, NNES responded that they required a higher minimum grade to be satisfied compared to NES. This may be due to cultural differences of international students who are NNES. Our findings match with previous studies in which Chinese international students often have high motivations to achieve academic success [49].

For three courses, non-native English speakers spent significantly more time every week studying outside of class. A possible interpretation for this could be that the expectations of NNES to achieve a higher minimum grade causes them to spend more time studying. This adds values to the findings of Sümer et al. [41] which found that international students with lower levels of English proficiency reported higher levels of depression and anxiety. Another interpretation may be that due to language barriers, NNES must spend more time outside of class to read and understand programming assignments, Q&A websites like Stack Overflow [15] and have difficulty acquiring technical vocabulary [10]. Our results confirm the findings of Airey that Swedish students usually just took notes during English lectures which required them to spend more time outside of class in order to understand the content better [1].

In three of the four courses, NNES had higher overall grades than NES but the differences were not significant. In CS1A, NES performed significantly better than NNES. This may be because CS1A used programming assignments based on cultural references (e.g., fast and furious movies) that non-native English speakers may not be able to relate to. Our study adds value to Guo's study [15]

which recommends avoiding culturally-specific slang and to use more culturally-agnostic code examples in instructional materials.

Although NNES did not prefer any particular type of resources in our study, we believe there is a need to develop resources specific to NNES learning CS as suggested in previous studies [13, 15]. More work must be done to understand the needs of NNES related to learning resources that will help them overcome language barriers.

5.1 Limitations and Future Work

Due to the limitation of language proficiency being self-reported in this study, generalizations were made with regards to the native language status of the students. Students who reported 'learning English after learning another primary language and are equally comfortable in both languages' were grouped with the students who were 'not as comfortable in English' because it was unknown at what point in their lives they learned English, and it was unclear if they could be considered native English speakers. Future studies should focus specifically on non-native English speakers who are not as comfortable in English (preferably measured using a standardized English proficiency test). In our future studies, we will ask explicit questions about students' native language to overcome this limitation.

Furthermore, we did not take the cultural background of students into account although culture influences students' communication and performance in programming projects [42]. Although the gender and prior experience factors were addressed as they relate to self-efficacy, there were many other demographic details relevant to a public university that were neglected (e.g., transfer-student status and race). As a result, future work should go more in depth and consider the sub-groups based on language and/or demographic backgrounds. Factors like culture, languages known, race, country of origin, etc. may be important to have a holistic understanding of non-native English speakers learning CS. Another limitation is that for the CS1.5 course, weekly survey data was only available for two weeks, which prevented a more holistic view of the four courses.

Future work should focus on learning why NNES are spending more time studying, by interviewing students to gain insights on their time-management skills and resources they use. Also, we should seek to understand the disparities that exist among NNES in relation to their access to AP CS A [9] and college level CS courses, and advocate that these opportunities become more broadly available in places where they are currently lacking. Future work could ask questions from a validated survey instrument such as the Computer Programming Self-Efficacy Scale [40] to measure students' aptitudes and/or self-efficacy in CS.

6 CONCLUSION

Overall, we found that non-native English speakers began courses with lower confidence levels but also higher expectations for themselves, and worked longer than native English speakers throughout the course, and had similar overall grades at the end of the term (in most courses). We conclude that more studies are needed to better understand the barriers that non-native English speakers face in computing to create more inclusive spaces for computing education so that one day we could truly achieve our dream of CS for all.

REFERENCES

- [1] John Airey and Cedric Linder. Language and the experience of learning university physics in sweden. *European journal of physics*, 27(3):553, 2006.
- [2] Maureen Snow Andrade. International students in english-speaking universities: Adjustment factors. *Journal of Research in International education*, 5(2):131–154, 2006.
- [3] Jeff Atwood. Non-english question policy. <https://stackoverflow.blog/2009/07/23/non-english-question-policy/>.
- [4] Jeff Atwood and Joel Spolsky. Stack overflow. <https://stackoverflow.com/>.
- [5] Albert Bandura. Self-efficacy. *The Corsini encyclopedia of psychology*, pages 1–3, 2010.
- [6] Sheila R Barnett, Raymond J Morin, Dan K Kiely, Margaret Gagnon, Gohar Azhar, Eric L Knight, Jerald C Nelson, and Lewis A Lipsitz. Effects of age and gender on autonomic control of blood pressure dynamics. *Hypertension*, 33(5):1195–1200, 1999.
- [7] Derrick A Bennett. How can i deal with missing data in my study? *Australian and New Zealand journal of public health*, 25(5):464–469, 2001.
- [8] Sylvia Beyer, Kristina Rynes, Julie Perrault, Kelly Hay, and Susan Haller. Gender differences in computer science students. In *Proceedings of the 34th SIGCSE technical symposium on Computer science education*, pages 49–53, 2003.
- [9] College Board. AP Computer Science A. <https://apcentral.collegeboard.org/courses/ap-computer-science-a/course>.
- [10] Peter E Childs and Finbar J O'FARRELL. Learning science through english: An investigation of the vocabulary skills of native and non-native english speakers in international schools. *Chemistry Education Research and Practice*, 4(3):233–247, 2003.
- [11] Oluwafolake Ogunbowo Dimandja. "We are not that different from you": A phenomenological study of undergraduate Muslim international student campus experiences. PhD thesis, University of Colorado Colorado Springs. Kraemer Family Library, 2017.
- [12] Daniel Ed Druckman and Robert A Bjork. *Learning, remembering, believing: Enhancing human performance*. National Academy Press, 1994.
- [13] Pedro Guillermo Feijóo-García, Keith McNamara Jr, and Jacob Stuart. The effects of native language on block-based programming introduction: A work in progress with hispanic population.
- [14] Chrystal George Mwangi, Alicia Peralta, Sharon Fries-Britt, and Nina Daoud. *Academic motivation and experiences of international students of color in STEM*, pages 197–211. 01 2016.
- [15] Philip J. Guo. Non-native english speakers learning computer programming: Barriers, desires, and design opportunities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–14, New York, NY, USA, 2018. Association for Computing Machinery.
- [16] Scott Hanselman. Do you have to know english to be a programmer. <https://hanselman.com/blog/DoYouHaveToKnowEnglishToBeAProgrammer.aspx>.
- [17] John D. Hunter. Matplotlib: A 2d graphics environment. 9:90, 2007.
- [18] Mrwan Ben Idris and Hany Ammar. The correlation between arabic student's english proficiency and their computer programming ability at the university level. *USA International Journal of Managing Public Sector Information and Communication Technologies (IJMPICIT)*, 2018.
- [19] Dave S Kerby. The simple difference formula: An approach to teaching nonparametric correlation. *Comprehensive Psychology*, 3:11–IT, 2014.
- [20] Päivi Kinnunen and Beth Simon. Cs majors' self-efficacy perceptions in cs1: results in light of social cognitive theory. In *Proceedings of the seventh international workshop on Computing education research*, pages 19–26, 2011.
- [21] David M Lane. *Online statistics education: An interactive multimedia course of study*. Rice University, 2007.
- [22] R. Lent, S. Brown, and Kevin C. Larkin. Relation of self-efficacy expectations to academic achievement and persistence. 1984.
- [23] Thomas Lumley, Paula Diehr, Scott Emerson, and Lu Chen. The importance of the normality assumption in large public health data sets. *Annual review of public health*, 23(1):151–169, 2002.
- [24] Lauri Malmi, Judy Sheard, Päivi Kinnunen, and Jane Sinclair. Theories and models of emotions, attitudes, and self-efficacy in the context of programming education. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, pages 36–47, 2020.
- [25] Gretchen McCulloch. Coding is for everyone as long as you speak english. <https://www.wired.com/story/coding-is-for-everyone-as-long-as-you-speak-english/>.
- [26] Wes McKinney. Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, page 51, 2010.
- [27] Mahmoud M Mhashi and ALIM Alakeel. Difficulties facing students in learning computer programming skills at tabuk university. *Recent Advances in Modern Educational Technologies*, pages 15–24, 2013.
- [28] Martin Moakler and Mikyong Kim. College major choice in stem: Revisiting confidence and demographic factors. *The Career Development Quarterly*, 62, 06 2014.
- [29] Oracle. Jdk 14 documentation. <https://docs.oracle.com/en/java/javase/14/>, 2009.
- [30] Yogendra Pal. A framework for scaffolding to teach programming to vernacular medium learners. *Dissertation. Indian Institute of Technology Bombay*, 2016.
- [31] Yogendra Pal and Sridhar Iyer. Effect of medium of instruction on programming ability acquired through screencast. In *2015 International Conference on Learning and Teaching in Computing and Engineering*, pages 17–21. IEEE, 2015.
- [32] Donald B Rubin. Multiple imputation after 18+ years. *Journal of the American statistical Association*, 91(434):473–489, 1996.
- [33] Pooja Sankar. Piazza. <https://piazza.com/>, 2009.
- [34] Pooja Sankar, Jessica Gilmartin, and Melissa Sobel. An examination of belongingness and confidence among female computer science students. *SIGCAS Comput. Soc.*, 45(2):7–10, July 2015.
- [35] Congressional Research Service. Foreign stem students in the united states. <https://crsreports.congress.gov/product/pdf/IF/IF11347>.
- [36] Adalbert Gerald Soosai Raj, Kasama Ketsuriyonk, Jignesh M Patel, and Richard Halverson. What do students feel about learning programming using both english and their native language? In *2017 International Conference on Learning and Teaching in Computing and Engineering (LaTICE)*, pages 1–8. IEEE, 2017.
- [37] Adalbert Gerald Soosai Raj, Kasama Ketsuriyonk, Jignesh M Patel, and Richard Halverson. Does native language play a role in learning a programming language? In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, pages 417–422, 2018.
- [38] Adalbert Gerald Soosai Raj, Eda Zhang, Saswati Mukherjee, Jim Williams, Richard Halverson, and Jignesh M Patel. Effect of native language on student learning and classroom interaction in an operating systems course. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, pages 499–505, 2019.
- [39] Adalbert Gerald Soosai Raj, Hanqi Zhang, Viren Abhyankar, Saswati Mukerjee, Eda Zhang, Jim Williams, Richard Halverson, and Jignesh M Patel. Impact of bilingual cs education on student learning and engagement in a data structures course. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, pages 1–10, 2019.
- [40] Phil Steinhilber, Andrew Petersen, and Jan Vahrenhold. Revisiting self-efficacy in introductory programming. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, pages 158–169, 2020.
- [41] Seda Sümer, Senel Poyrazlı, and Kamini Grahame. Predictors of depression and anxiety among international students. *Journal of Counseling and Development*, 86(4):429–437, January 2008.
- [42] Kathleen Swigger, Ferda Alpaslan, Robert Brazile, and Michael Monticino. Effects of culture on computer-supported international collaborations. *International Journal of Human-Computer Studies*, 60(3):365–380, 2004.
- [43] Gabriela Valdez. US higher education classroom experiences of undergraduate Chinese international students. *Journal of International Students*, 5(2):188–200, 2015.
- [44] Raphael Vallat. Pingouin: statistics in python. *The Journal of Open Source Software*, 3(31):1026, November 2018.
- [45] S. van der Walt, S. C. Colbert, and G. Varoquaux. The numpy array: A structure for efficient numerical computation. *Computing in Science Engineering*, 13(2):22–30, 2011.
- [46] Ashok Kumar Veerasamy and Anna Shillabeer. Teaching english based programming courses to english language learners/non-native speakers of english. *International Proceedings of Economics Development and Research*, 70:17, 2014.
- [47] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C.J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [48] Michael Waskom, Olga Botvinnik, Paul Hobson, John B. Cole, Yaroslav Halchenko, Stephan Hoyer, Alistair Miles, Tom Augspurger, Tal Yarkoni, Tobias Megies, Luis Pedro Coelho, Daniel Wehner, cynddl, Erik Ziegler, diego0020, Yury V. Zaytsev, Travis Hoppe, Skipper Seabold, Phillip Cloud, Miikka Koskinen, Kyle Meyer, Adel Qalieh, and Dan Allan. seaborn: v0.5.0 (november 2014), November 2014.
- [49] Kun Yan and David C Berliner. Chinese international students' academic stressors in the united states. *College Student Journal*, 43(4), 2009.