Replication and Expansion Study on Factors Influencing Student Performance in CS2

Margaret Ellis
Virginia Tech
Blacksburg, VA, USA
maellis1@vt.edu

Sara Hooshangi
Virginia Tech
Blacksburg, VA, USA
shoosh@vt.edu

ABSTRACT

While many studies have focused on students' performance in CS1 courses, research related to the performance and persistence of students in CS2 classes is not as widely performed. In this work, we will extend our previous work to examine students' performance in CS2. We examined a data set that spanned over seven years on more than 5300 student records. In addition to typical factors studied by others (i.e. gender, race, CS1 performance), our work also took into account the relationship between various CS1 pathways to CS2, student major, and the number of previous college CS courses (including transfer credits) and student performance in CS2. CS1 grade is a good indicator of performance in CS2. Gender was not a significant factor in determining performance in CS2 and undeclared engineering majors stood out as high performers. CS majors passed the course at higher rates than other majors. Our large data set allowed for more granular analysis according to race and ethnicity and additional access to students' underserved status. Race and ethnicity had a significant correlation with performance, and so did the underserved status. Our large data set confirmed some of the findings of our previous work, while providing some new insight.

CCS CONCEPTS

• **Social and professional topics** → *Student assessment*.

KEYWORDS

CS2, Student Performance, Data Structures, Prior CS Knowledge, Diversity

ACM Reference Format:
Margaret Ellis and Sara Hooshangi. 2023. Replication and Expansion Study on Factors Influencing Student Performance in CS2. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2023), March 15–18, 2023, Toronto, ON, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3545945.3569867

1 INTRODUCTION

CS Educators are interested in understanding factors that contribute to students' success in computer science courses. The hope is that by further understanding student performance and persistence in computers science courses, the development of curriculum, instruction, and supports can be improved to increase overall performance and broaden participation in the field. There is a high demand for computing professionals and a computer science education can open doors for many students. Both CS1 and CS2 are gateway courses that are critical in attracting and preparing students for degrees in computer science.

Many previous studies focus on the factors that impact success in CS1 such as psychometrics, external factors, and previous experiences [10]. Fewer studies have focused on student performance in CS2. The most commonly studied predictor of success in CS2 is performance in CS1. Previous research has also examined prior math experience, demographics, self-efficacy, and other course related analysis [2, 5, 8, 14, 25, 26]. In our previous CS2 analysis, we analyzed persistence and performance of 610 students in a single semester. We introduced two new factors that had not been studied in the past. These factors are 1) pathways to CS2 (AP CS, institution CS1, transfer credits, non-standard pathways) and 2) the number of CS courses that a student has taken before arriving at CS2 [15].

With our larger data set of 5,300 CS2 students over the span of seven years, we revisit analysis of factors such as CS1 performance, gender, and major with this larger data set. We also investigate race and ethnicity with increased granularity. Additionally, our larger data set provides indicators for whether students are underserved, so we consider that as an additional factor related to CS2 performance.

2 PREVIOUS RESEARCH

A plethora of studies have examined the factors that influence success in introductory CS courses with the ultimate goal of determining interventions to improve student success early in the CS pipeline [14, 25]. In previous studies, various factors have been investigated and several metrics are commonly used to indicate student success, the most common being course grade. The most commonly studied predictors are engagement and performance in previous courses. Demographic and psychometric factors such as self-regulation and self-efficacy have also been studied [14, 25]. Beck et al. recently attempted to determine specific CS1 exam questions that can assess students' readiness for success in CS2 [2].

Bergin et al. determined that both mathematics preparation in high school and self-regulation correlated positively with success in CS1 [3, 4, 24]. Alavardo et al. have shown that students with AP credits have higher grades across many CS courses [1]. Catanese et al. have shown that transfer students perform well in higher level computer science courses as non-transfer students [7].

CS1 course performance has been repeatedly show to predict CS2 course performance [5, 8, 19]. Our previous study on a single semester of data from this course verified these findings and also indicated that gender and race were not significantly correlated with

SIGCSE 2023, March 15–18, 2023, Toronto, ON, Canada

ACM ISBN 978-1-4503-9431-4/23/03.
https://doi.org/10.1145/3545945.3569867

whether students passed, did not pass, or dropped the course. However, undeclared engineering majors stood out as high performers and students' CS pathway leading to CS2 was also significant. Students with CS1 transfer credits had significantly lower pass rates, while students with only one previous CS course credit were less likely to drop or not pass the course [15].

Our larger data set provides the opportunity to analyze success rates for more specific races. The National Center for Education Statistics data indicates that both Black and Hispanic students are less likely to finish a bachelor's degree within four years relative to their White and Asian peers [11]. George et al. paints a more complex picture of persistence specific to computing and based on student interest according to race. Two years after taking an introductory computing course, the Asian/Asian American group was the only race/ethnicity variable that was statistically significant in predicting interest in a computing career [13].

Our larger data set provides information about whether or not a student is considered underserved. Students who are identified as underserved at our university include students who are first generation college students, receive Pell Grant funding, or identify as veterans. These three populations have unique characteristics but all experience an academic performance gap [6, 13, 16, 17, 30, 36]. A review of a CS2 course containing 86 students at Humbloldt State University reported a variation in course success rates for specific sub-populations. The first generation students had a 70.6% pass rate, while the not first generation students had a 82.1% success rate. Students in that course receiving financial aide only had a 69.8% success rate, compared to the 84.8% success rate for students who received no financial aide [21].

Many studies have demonstrated student success is affected by multiple factors including students' previous experience, self-efficacy and sense-of-belonging. Prior computing experience such as internships, co-curricular activities, or high school programming courses can significantly impact a student's experience in a computer science course [23]. In Rountree's study, expectancy for success was the strongest single indicator of success in a CS1 course, more so than demographics and background [27]. Krauss-Levy et al. found that a lower sense of belonging was negatively correlated with course performance and pass rates [18].

Students from minoritized groups tend to experience lower self efficacy and sense of belonging. For example, Krauss-Levy et al's found that women, first generations students and transfer students had a lower sense of belonging [18]. In a study of over 100 institutions, black and female students were more likely to have communal goal orientation and to have a weaker sense of belonging with computing [20]. Also, controlling for race and gender, Nguyen and Lewis found that competitive enrollment was a negative predictor of sense of belonging and self-efficacy for students without prior experience [22].

Other research has focused on examining the WDF rate and improving that ratio post-intervention [31]. WDF is the percentage of students receiving a grade of "D" or "F", or withdrawing from the course before the course ends. Further research into the factors that influence WDF can shed light on effective intervention methods that could impact students who are in the danger of exiting the CS pipeline.

3 RESEARCH QUESTIONS

This project aims to examine different factors that are correlated with student performance and retention in CS2. Our research questions are the following as a replication study with a larger data set:

- RQ1: Are demographic factors such as gender, race, student's major, and underserved status correlated with performance (passing or failing) CS2?
- RQ2: Is CS1 grade correlated with CS2 performance?
- RQ3: Is CS1 pathway to CS2 (standard CS1 at our institution or transfer credit for CS1) correlated with performance in CS2?

4 INSTITUTIONAL BACKGROUND

Our institution is a large research university located in United States with more than 30,000 undergraduate students and a rapidly growing CS department. Computer Science is one of 13 departments within the College of Engineering (COE). Students are admitted to the College of Engineering and are considered General Engineering majors (referred to in this paper as undeclared engineering) in their first year. They can declare a major within the college after their first year. All students in computer science major or minor are required to take our standard CS2 course which is taught in Java.

4.1 CS2 Pathways

The software and programming sequence of courses for our CS majors and minors is CS1 (Introduction to Software Design), CS2 (Introduction to Data Structures and Software Design), and CS3 (Data Structures and Algorithms) all taught in Java with no previous experience required in CS1. However, many students enter CS2 by satisfying the prerequisite with a course other than the standard CS1 course offered to CS majors at our institution.

Students may have earned credit for CS1 during high school by passing the APCS A exam, taking a dual enrollment computer science course, or a CS course at an International Baccalaureate (IB) High School. Some students may earn CS1 transfer credits directly from another university or a community college. Students in other majors at our institution can take an introductory course in Java or a sequence of Python courses before enrolling in CS2.

4.2 Dataset

This study focuses on the examination of students who took CS2 between Fall 2015 and Spring 2022. The course was heavily updated in the Fall of 2015 and underwent subtle improvements over the years, with a reorganization in Spring 2020 to address needs to deliver the course online and to incorporate an increased number of smaller coding practice opportunities. The objectives of the course, projects, and tests remained very similar over this seven year period.

Data was acquired from the University Registrar's office upon IRB approval. We only focused on students who attempted and received a grade for CS2 between Fall 2015 and Spring 2022. Table 1 shows the gender, race/ethnicity, and underserved status distribution of students who completed the course. A total of 5316 students were included in our data set. Gender and race distributions are similar to the composition of the CS undergraduate population in our department. For race, we have used similar categories found

in other Computing Research Association (CRA) publications, but our data does not separate non-resident students from domestic students. We used the National Science Foundation [12] definition of underrepresented minorities in the STEM field, where Hispanic of any race is categorized as Hispanic. Other/mixed category includes racial and ethnic groups combined (American Indians or Alaska Natives, Native Hawaiians or Other Pacific Islanders, and individuals who report more than one race and are not Hispanic). Students without a disclosed race are marked as Not Reported in our tables and figures. At our institution, non-international students are reported as underserved based on any one of three factors, 1) neither parent completed a bachelor's degree, commonly referred to as first-generation students, or 2) qualify for a Pell Grant based on the Free Application for Federal Student Aide (FAFSA), or 3) self-identify as veterans and receive GI benefit while enrolled [29].

Table 1: Gender, race/ethnicity and underserved status

			Total
Gender	Male	4250 (80.0%)	5316 (100%)
	Female	1058 (19.9%)	
	Unknown	8 (0.1%)	
Race/ Ethnicity	White	2447 (46.0%)	5316 (100%)
	Asian	1924 (36.2%)	
	Hispanic	314 (5.9%)	
	Other/Mixed	273 (5.1%)	
	Black	206 (3.9%)	
	Not Reported	152 (2.9%)	
Underserved	No	4298 (80.9 %)	5316 (100%)
	Yes	1012 (19.0%)	
	Unknown	6 (0.1%)	

Figure 1 shows the trend in our CS2 enrollment over the seven years of analyzed data. We see a steady increase in the number of students who enroll in CS2 each year, consistent with the growth that has been observed in the number of undergraduate CS students in our department. Each vertical bar represents total enrollment for fall, spring, and summer in an academic year. For 21-22 academic year, the data for summer enrollment was not available at the time of this analysis. This contributes to the lower number of students shown for that academic year. The percentage of female students averages around 20% over the seven years, with a slight increase over the past three years, from 17% in 2019-2022 academic year to 23% in 2021-2022 academic year.

For this data set we also had indicators for whether a student was underserved as defined earlier. Overall, 19% of students were underserved across the data set. The percentage of students who have an underserved status has fluctuated over the years, with its peak at 23% in 2016-2017 academic year, dropping to 16% in 2020-2021, and then slowly increasing to 19% at the end of 2021-2022 academic year. Figure 2 shows the trend in our CS2 enrollment by underserved status over the seven years of data.

5 RESULTS

For our 7-year CS2 data set, we analyzed the following factors: gender, race, student major, CS1 pathway, number of CS courses taken before CS2, underserved status, and CS1 grade. We investigated the relationship between these factors and CS2 grades. We also analyzed the correlation of these factors with whether students passed the course or received a WDF. For this course, students must earn a C or better to proceed to the next course, so C- students were categorized in the WDF group.

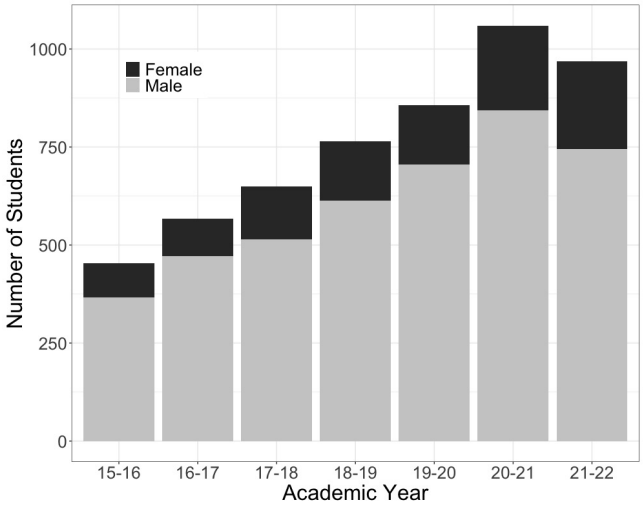

Figure 1: CS2 enrollment over the years by gender

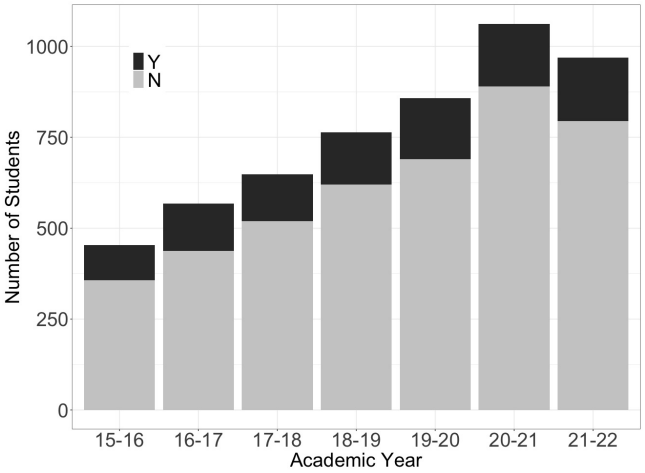

Figure 2: CS2 enrollment by underserved status

5.1 RQ1: Are demographic factors such as gender, race, student's major, and underserved status correlated with performance (passing or failing) CS2?

To examine demographic factors (gender, race, major, and underserved status) that may be correlated with student performance in CS2, we looked at each of these factors separately and examined the performance of students belonging to each group within a given factor. A few students had received atypical grades such as incomplete, repeat or no grade. To simplify our analysis, we had removed these few cases from our data set before any analysis.

Table 2 shows passing rate based on binary gender. The number of female students in our CS2 class is consistent with the overall rate of female students in our department. The result of this table suggests that male and female students are performing similarly in the course, and passing at the same rate.

Table 2: Pass/fail rate based on gender

Gender	Pass	WDF	Total
Male	3968 (93.4%)	282 (6.6%)	4250 (100.0%)
Female	999 (94.4%)	59 (5.6%)	1058 (100.0%)
Unknown	8(100%)	0 (0.0%)	8 (100.0%)
Total	4975 (93.4%)	341 (6.4%)	5316 (100.0%)

We further analyzed the relationship between gender and performance by running a Chi-squared analysis based on two possible CS2 outcomes; passing or not passing (WDF). Our analysis reveals that for gender ($\chi 2 = 2.13, p = .34$), there is no significant difference between CS2 outcome. This is consistent with the result of our previous study, where we saw no correlation between gender and performance [15].

We also looked at race and ethnicity as a factor that may be correlated with performance in the class. Figure 3 shows the grade distribution for various demographic groups as box-plots. Table 3 shows passing rate based on race/ethnicity. The initial analysis of these results indicates that Hispanic and Black students are passing at a slightly lower rate than Asian and White students. To further examine this observation, we ran a Chi-squared analysis that indicated a significant difference based on these race/ethnicity categories ($\chi 2 = 17.65, p = .003$).

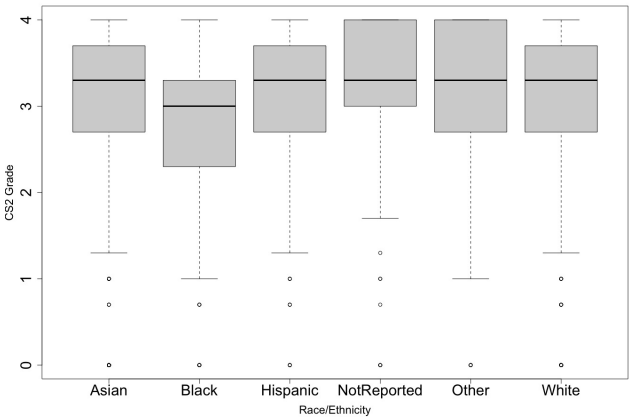

Figure 3: CS2 performance based on race/ethnicity

We further analyzed our result by performing pairwise Chi-square analysis between Black/White, Black/Asian, Hispanic/White and Hispanic/Asian students. The pairwise comparison showed a significant difference between Black students' rate of passing when compared to White students ($\chi 2 = 6.97, p = .008$) and Asian students ($\chi 2 = 6.65, p = .009$). Similarly Hispanic students' passing rate when compared to White students ($\chi 2 = 8.35, p = .003$) and Asian students ($\chi 2 = 7.85, p = .005$) were significant. No significant difference was observed between Asian and White students, as expected from the passing rates shown Table 3. These results, which were possible to obtain with the larger data set, demonstrate that further research and interventions are needed to support Black and Hispanic students.

Table 3: Pass/fail rate based on race/ethnicity

Race	Pass	WDF	Total
White	2305 (94.2%)	142 (5.8%)	2447 (100%)
Asian	1812 (94.2%)	112 (5.8%)	1924 (100.0%)
Hispanic	282 (89.8%)	32 (10.2%)	314 (100.0%)
Other/Mixed	252 (92.3%)	21 (7.7%)	273 (100.0%)
Black	184 (89.3%)	22 (10.7%)	206 (100.0%)
Not Reported	140 (92.1%)	12 (7.9%)	152 (100.0%)
Total	4975 (93.4%)	341 (6.4%)	5316 (100.0%)

We also looked at how students' major was correlated with performance. Figure 4 shows CS2 grade distribution for different majors. In this analysis, we separated the majors into four categories: "CS" for students majoring in computer science; "Other STEM", for students majoring in a STEM field other than CS' "Undeclared Engineering", for freshmen students in the College of Engineering who are yet to declare a major, and finally "Other" to account for students who are not in any STEM fields. The graph indicates that undeclared engineering students have the highest average grade followed by CS majors, which is the same order determined in the Spring 2021 data set [15].

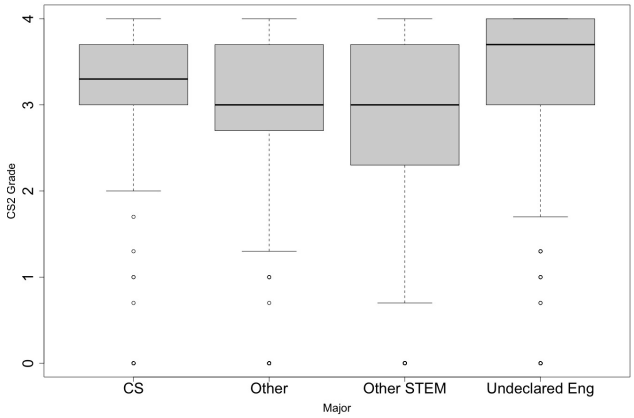

Figure 4: CS2 performance based on major

Table 4 shows passing rate based on major. The CS students are passing at higher rates (98.3%) than other majors, followed by undeclared engineering students. The students coming from non-STEM fields have the highest rate of failure. To further examine this result, we ran a Chi-squared analysis which indicated that a significant difference is observed based on major ($\chi 2 = 150.7, p < .001$). We further analyzed our data by performing pairwise Chi-square analysis between all major pairs and in all cases the p-value was less than 0.001. While the performance of undeclared engineering majors is higher than CS students, overall CS students

are passing at a higher rate than others. The students from non-STEM majors had the highest rate of failure relative to any other demographic group analyzed (14%) and more interventions should be considered to increase their success in CS2.

Table 4: Pass/fail rate based on student's major

Major	Pass	WDF	Total
CS	1270 (98.3%)	22 (1.7%)	1292 (100.0%)
Undeclared Eng	1706 (95.8%)	75 (4.2%)	1781 (100.0%)
Other STEM	1458 (90.3%)	156 (9.7%)	1614 (100.0%)
Other	541 (86.0%)	88 (14.0%)	629 (100.0%)

Finally, we looked at the relationship between underserved status and performance in the CS2 class. As shown in Table 1, 19% of students are considered underserved. The average grade for these students is similar to other groups, and we see no difference in the distribution of their CS2 grades. However Table 5 shows that the passing rate is slightly less for those in the underserved category. We further analyzed the relationship between underserved status and performance by running a Chi-squared analysis. Our analysis reveals that for the underserved status, there is a significant difference between CS2 passing outcome ($chi2 = 11.23, p < .001$). These results, which were possible to obtain with the underserved indicator in this data set, demonstrate that further research and interventions are needed to support students who are considered underserved.

Table 5: Pass/fail rate based based on underserved status

Underserved	Pass	WDF	Total
N	4046 (94.1%)	252 (5.9%)	4298 (100.0%)
Y	923 (91.2%)	89 (8.8%)	1012 (100.0%)
Unknown	6 (100.0%)	0(0.0%)	6 (100.0%)

5.2 RQ2: Is CS1 grade correlated with CS2 performance?

For our data set, CS1 grade reasonably predicts CS2 grade. Figure 5 shows the relationship between CS1 grade and CS2 grade. Students need a grade of C or better in CS1 before they can complete CS2, hence the grade distribution for CS1 only includes the range of C to A grades. As shown in Figure 5, the grades of students who took CS1 at our institution are correlated with their CS2 grades. This correlation is statistically significant with a $p< .001$ and a correlation coefficient of 0.42. The mean grade for CS1 is 3.06 and 3.14 for CS2. As shown in the box-plot, there is more variation in performance for students who received a lower CS1 grade, but overall CS1 grade is a significant factor in CS2 grade. These findings align well with many previous studies that have correlated CS1 and CS2 grade as discussed in Section 2.

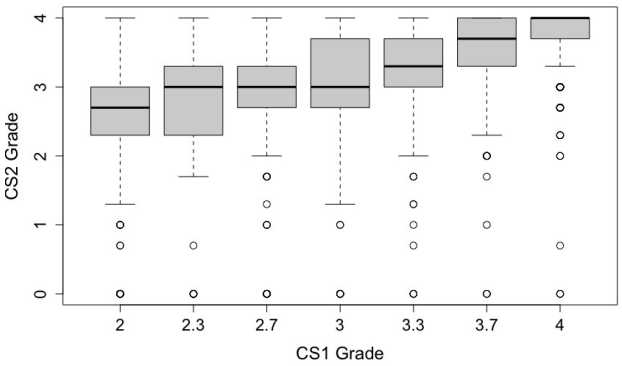

Figure 5: CS2 performance based on CS1 grades

5.3 RQ3: Is CS1 pathway to CS2 (standard CS1 at our institution or transfer credit for CS1) correlated with performance in CS2?

Students reach our CS2 classes from a variety of different academic pathways. Based on information provided in our data set, we created three categories to indicate student pathway to CS2 as follows: "CS1" for our traditional CS1 path; "CC Transfer" for community college transfer credit; "Other Transfer" for transfer credit from all other sources. "Other Transfer" encompasses AP CS, IB, four year institution transfers, and other pathways at our institution, but this granularity was not available at the time of analysis. Table 6 summarizes students' path to CS2 for these groups. About two-thirds of our students take our CS1 course, and the rest transfer the credit into our program.

We also looked into the number of prior CS college-level courses that a student had credit for before enrolling in CS2. In the 7-year data set, 90% of students came to CS2 from a single college course, which in all likelihood is CS1 or an equivalent transfer credit. Meanwhile, 7% came with 2 previous courses, and only 3% had taken 3 or more courses before enrolling in CS2. There was less variation in the number of previous CS courses the students had taken in the larger historical data set than in the recent single semester data set where 31% of the students had credit for 2 or 3 previous college-level CS courses.

Table 6: Pass/fail rate based on CS1 pathway

CS1 Path	Pass	WDF	Total
CS1	2863 (93.6%)	195 (6.4%)	3058 (100.0%)
CC* Transfer	192 (94.1%)	12 (5.9%)	204 (100.0%)
Other Transfer	1920 (93.5%)	134 (6.5%)	2054 (100.0%)
Total	4975 (93.6%)	341 (6.4%)	5316 (100.0%)

*CC: Community College

We further analyzed the relationship between groups in each factor by running a Chi-squared analysis on each factor. Our Chi-squared analysis reveals that for neither CS1 pathway ($\chi 2 = 0.44, p = .93$), nor number of prior CS courses ($\chi 2 = 0.02, p = 1.0$), was significantly correlated with CS2 grade. This is in contrast to what we

observed in our previous study on a single recent semester [15], where we saw a significance difference. We partially attribute this to the fact that there is now more variation in students' CS1 course and number of prior CS courses which is thus not reflected in the larger, more historic data set. Recognizing this changing landscape is a reminder to continue researching how to support students with various CS college course pathways to CS2.

6 DISCUSSION

In this follow-up paper, we revisited several factors that may be correlated with student performance in a CS2 class on a large data set that spanned over seven years and contained more than 5300 student records. Our analysis corroborates previous research on some of the predictors of CS2 performance and also provides some new and specific insights that can be further investigated. CS1 grades are significant factors in our students' CS2 performance. Our results confirm findings in previous studies [5, 8, 10] and also our own previous one-semester study [15] in this regard. Furthermore, we saw no evidence of gender as being a significant factor in student performance or passing rate for this CS2 course. Prior works have also shown no correlation between gender and performance in introductory CS courses [33, 34].

On the other hand, we found race and ethnicity to be correlated with student performance in this CS2 course over the 7 year time frame. The lower passing rate of these students is also indicative that Black and Hispanic students are at a disadvantage when compared to their White and Asian peers, and risk dropping out of the CS pipeline at a higher rates. Further work needs to be done to support Black and Hispanic students in our CS2 course. It is possible these students' performance is affected by their level of previous experience, self-efficacy, and sense-of-belonging, all of which can be interrelated and affect academic performance [18, 23, 27, 28]. More research and interventions need to be undertaken. For example, if these students have communal goal orientation and a weaker sense of belonging similar to students in [20] then possibly volunteering opportunities, mentoring, and group projects may improve sense of belonging and thus performance. A psychological intervention that is designed to instill that hardship and doubt are common to all CS2 students regardless of race may also be able to improve Black students' sense of belonging and self-efficacy and thus performance [32].

In addition, student's major was also a significant factor in predicting CS2 grade. While undeclared engineering students, who are College of Engineering students who have not yet declared a specific major, had a higher average grade in CS2, overall CS majors had the higher rate of passing the course. The difference in passing rate between various major categories in this study (CS, undeclared engineering, other STEM, and other) was significant. This indicates that CS students are passing CS2 at a higher rate. In our previous study [15], we saw that the undeclared engineering group outperformed others and passed at a higher rate. There is now a larger and higher performing percentage of undeclared engineering students who are taking CS2. These students are 34% of the 7-year population and 52% of the Spring 2021 population. Placing CS2 students in cohorts may help improve the performance of the other STEM majors and other majors similarly to how Decker saw a decreased WDF rate in CS1 when students were placed in cohorts based on previous experience [9].

Underserved status was also significantly correlated with the passing rate of a student. Students who were considered underserved passed at a lower rate, and this indicates that intervention or support mechanisms should be put in place to provide support to these groups. Transparent teaching strategies as outlined by Winkelmes et al. could improve confidence, sense of belonging, and mastery of skills for low-income and underrepresented CS2 students [35]. For example, ensuring that problem-based assignments clearly state knowledge gained and skills practiced that are relevant to students 5 years in the future.

We saw no significant difference between students who transferred CS1 credit from a community college. In our previous study students who transferred CS1 credit had a lower passing rate. Our analysis on the larger historic data set is more aligned with work by Catanese et al. that showed transfer students perform as well as non-transfer students in computing courses [7]. It is possible that historically transfer students transfer CS1 credit, but more recently an increased number of non-transfer students also obtain credit outside our institution. Unfortunately, our data set did not indicate transfer student status and did not have enough granularity to examine the specific source of all possible CS1 pathways.

6.1 Threats to Validity

Because there are no defined experimental and control groups in this study, a direct causality between process measurements and our outcomes cannot be inferred. Differential experience could be a potential threat to this study, as students come from diverse academic backgrounds. Our findings are based on data for students at a larger research institution. As such our results might not be generalizable to all undergraduate students. However, the characteristics of our students are representative of typical undergraduate students who enroll in CS2 offered at large research institutions in the United States.

7 CONCLUSION

This is a replication study using a large data set over many years of student records. We verified that CS1 grade is a good indicator of CS2 performance, and that gender is not correlated with performance. We found that Black and Hispanic students pass CS2 at a lower rate and more attention must given to their success in CS2 classes. Students who have already declared a CS major pass the course at higher rates than other majors. Further study is needed to understand current students' prior pathways and computing experience before enrolling in CS2. Further study is also needed to investigate these factors, not just in isolation, but in conjunction with one another and previous computing experience. Student self-efficacy and sense of belonging should also be considered in order to better understand how to support students for success in CS2.

REFERENCES

[1] Christine Alvarado, Gustavo Umbelino, and Mia Minnes. 2018. The Persistent Effect of Pre-College Computing Experience on College CS Course Grades. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (Baltimore, Maryland, USA) *(SIGCSE '18)*. Association for Computing Machinery, New York, NY, USA, 876–881. https://doi.org/10.1145/3159450.3159508

[2] Leland Beck, Patty Kraft, and Alexander W. Chizhik. 2022. Predicting Student Success in CS2: A Study of CS1 Exam Questions. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1* (Providence, RI, USA) *(SIGCSE 2022)*. Association for Computing Machinery, New York, NY, USA, 140–146. https://doi.org/10.1145/3478431.3499276

[3] Susan Bergin and Ronan Reilly. 2005. Programming: Factors That Influence Success. *SIGCSE Bull.* 37, 1 (Feb. 2005), 411–415. https://doi.org/10.1145/1047124.1047480

[4] Susan Bergin, Ronan Reilly, and Desmond Traynor. 2005. Examining the Role of Self-Regulated Learning on Introductory Programming Performance. In *Proceedings of the First International Workshop on Computing Education Research* (Seattle, WA, USA) *(ICER '05)*. Association for Computing Machinery, New York, NY, USA, 81–86. https://doi.org/10.1145/1089786.1089794

[5] Halil Bisgin, Murali Mani, and Suleyman Uludag. 2018. Delineating Factors that Influence Student Performance in a Data Structures Course. In *2018 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9. https://doi.org/10.1109/FIE.2018.8659300

[6] Emily Forrest Cataldi. 2018. First-generation students: College access, persistence, and postbachelor's outcomes. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018421

[7] Helen Catanese, Carl Hauser, and Assefaw H. Gebremedhin. 2018. Evaluation of Native and Transfer Students' Success in a Computer Science Course. *ACM Inroads* 9, 2 (April 2018), 53–57. https://doi.org/10.1145/3204471

[8] Holger Danielsiek and Jan Vahrenhold. 2016. Stay on These Roads: Potential Factors Indicating Students' Performance in a CS2 Course. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (Memphis, Tennessee, USA) *(SIGCSE '16)*. Association for Computing Machinery, New York, NY, USA, 12–17. https://doi.org/10.1145/2839509.2844591

[9] Adrienne Decker, Christopher Egert, and Erin Cascioli. 2020. Cohorting Incoming Students in a CS1 Course: Experiences and Reflections from the First Year of Implementation. *J. Comput. Sci. Coll.* 35, 8 (aug 2020), 186–197.

[10] Richard J. Enbody, William F. Punch, and Mark McCullen. 2009. Python CS1 as Preparation for C++ CS2. *SIGCSE Bull.* 41, 1 (mar 2009), 116–120. https://doi.org/10.1145/1539024.1508907

[11] National Center for Education Statistics (NCES). 2021. Digest of Education Statistics. *Retrieved August 19, 2022 from https://nces.ed.gov/programs/digest/d21/tables/dt21_326.10.asp*, (2021).

[12] National Science Foundation. 2022. Women, Minorities, and Persons with Disabilities in Science and Engineering. *Retrieved August 19, 2022 from https://ncses.nsf.gov/pubs/nsf21321/* (2022).

[13] Kari L. George, Linda J. Sax, Annie M. Wofford, and Sarayu Sundar. 2022. The tech trajectory: Examining the role of college environments in shaping students' interest in Computing Careers. *Research in Higher Education* 63, 5 (2022), 871–898. https://doi.org/10.1007/s11162-021-09671-7

[14] Arto Hellas, Petri Ihantola, Andrew Petersen, Vangel V. Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, Chris Messom, and Soohyun Nam Liao. 2018. Predicting Academic Performance: A Systematic Literature Review. Association for Computing Machinery, New York, NY, USA, 175–199. https://doi.org/10.1145/3293881.3295783

[15] Sara Hooshangi, Margaret Ellis, and Stephen H. Edwards. 2022. Factors Influencing Student Performance and Persistence in CS2. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1* (Providence, RI, USA) *(SIGCSE 2022)*. Association for Computing Machinery, New York, NY, USA, 286–292. https://doi.org/10.1145/3478431.3499272

[16] Terry T. Ishitani. 2006. Studying attrition and degree completion behavior among first-generation college students in the United States. *The Journal of Higher Education* 77, 5 (2006), 861–885. https://doi.org/10.1353/jhe.2006.0042

[17] Brandy M. Jenner. 2017. Student veterans and the transition to Higher Education: integrating existing literatures. *Journal of Veterans Studies* 2, 2 (2017), 26. https://doi.org/10.21061/jvs.14

[18] Sophia Krause-Levy, William G. Griswold, Leo Porter, and Christine Alvarado. 2021. The Relationship Between Sense of Belonging and Student Outcomes in CS1 and Beyond. In *Proceedings of the 17th ACM Conference on International Computing Education Research* (Virtual Event, USA) *(ICER 2021)*. Association for Computing Machinery, New York, NY, USA, 29–41. https://doi.org/10.1145/3446871.3469748

[19] Lucas Layman, Yang Song, and Curry Guinn. 2020. Toward Predicting Success and Failure in CS2: A Mixed-Method Analysis. In *Proceedings of the 2020 ACM Southeast Conference* (Tampa, FL, USA) *(ACM SE '20)*. Association for Computing Machinery, New York, NY, USA, 218–225. https://doi.org/10.1145/3374135.3385277

[20] Colleen Lewis, Paul Bruno, Jonathan Raygoza, and Julia Wang. 2019. Alignment of Goals and Perceptions of Computing Predicts Students' Sense of Belonging in Computing. In *Proceedings of the 2019 ACM Conference on International Computing Education Research* (Toronto ON, Canada) *(ICER '19)*. Association for Computing Machinery, New York, NY, USA, 11–19. https://doi.org/10.1145/3291279.3339426

[21] Adamou Fode Made and Abeer Hasan. 2020. Creating a More Equitable CS Course through Peer-Tutoring. *J. Comput. Sci. Coll.* 35, 10 (apr 2020), 33–38.

[22] An Nguyen and Colleen M. Lewis. 2020. Competitive Enrollment Policies in Computing Departments Negatively Predict First-Year Students' Sense of Belonging, Self-Efficacy, and Perception of Department. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (Portland, OR, USA) *(SIGCSE '20)*. Association for Computing Machinery, New York, NY, USA, 685–691. https://doi.org/10.1145/3328778.3366805

[23] Anne-Kathrin Peters, Anders Berglund, Anna Eckerdal, and Arnold Pears. 2014. First Year Computer Science and IT Students' Experience of Participation in the Discipline. In *2014 International Conference on Teaching and Learning in Computing and Engineering*. 1–8. https://doi.org/10.1109/LaTiCE.2014.9

[24] Keith Quille and Susan Bergin. 2018. Programming: Predicting Student Success Early in CS1. a Re-Validation and Replication Study. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education* (Larnaca, Cyprus) *(ITiCSE 2018)*. Association for Computing Machinery, New York, NY, USA, 15–20. https://doi.org/10.1145/3197091.3197101

[25] Keith Quille and Susan Bergin. 2019. CS1: how will they do? How can we help? A decade of research and practice. *Computer Science Education* 29, 2-3 (2019), 254–282. https://doi.org/10.1080/08993408.2019.1612679 arXiv:https://doi.org/10.1080/08993408.2019.1612679

[26] Farzana Rahman and Jaya Tyagi. 2021. Investigating the Role of Different Prep Pathways on CS2 Performance Across Three Different Majors. In *Proceedings of the 22st Annual Conference on Information Technology Education* (SnowBird, UT, USA) *(SIGITE '21)*. Association for Computing Machinery, New York, NY, USA, 141–146. https://doi.org/10.1145/3450329.3476851

[27] Nathan Rountree, Janet Rountree, and Anthony Robins. 2002. Predictors of Success and Failure in a CS1 Course. *SIGCSE Bull.* 34, 4 (dec 2002), 121–124. https://doi.org/10.1145/820127.820182

[28] Anya Tafliovich, Jennifer Campbell, and Andrew Petersen. 2013. A Student Perspective on Prior Experience in CS1. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education* (Denver, Colorado, USA) *(SIGCSE '13)*. Association for Computing Machinery, New York, NY, USA, 239–244. https://doi.org/10.1145/2445196.2445270

[29] Virginia Tech. 2022. University DataCommons. *Retrieved August 19, 2022 from https://udc.vt.edu/irdata/dictionary* (2022).

[30] Paul Tough. 2014. "Who Gets to Graduate?". *The New York Times Magazine, May 15. New York: New York Times Company.* (2014).

[31] Arto Vihavainen, Jonne Airaksinen, and Christopher Watson. 2014. A Systematic Review of Approaches for Teaching Introductory Programming and Their Influence on Success. In *Proceedings of the Tenth Annual Conference on International Computing Education Research* (Glasgow, Scotland, United Kingdom) *(ICER '14)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/2632320.2632349

[32] Gregory M. Walton and Geoffrey L. Cohen. 2007. A Question of Belonging: Race, Social Fit, and Achievement. *Journal of Personality and Social Psychology* 92, 1 (2007), 82–96.

[33] Brenda Cantwell Wilson. 2002. A Study of Factors Promoting Success in Computer Science Including Gender Differences. *Computer Science Education* 12, 1-2 (2002), 141–164. https://doi.org/10.1076/csed.12.1.141.8211

[34] Brenda Cantwell Wilson and Sharon Shrock. 2001. Contributing to success in an introductory computer science course: a study of twelve factors. *Acm sigcse bulletin* 33, 1 (2001), 184–188. https://doi.org/10.1145/364447.364581

[35] Mary-Ann Winkelmes, Matthew L. Bernacki, Jeffrey Butler, Michelle Zochowski, Jennifer Dawn Golanics, and Kathryn Harriss Weavil. 2016. A Teaching Intervention That Increases Underserved College Students' Success.

[36] Sherry A. Woosley and Dustin K. Shepler. 2011. Understanding the Early Integration Experiences of First-Generation College Students. *College student journal* 45 (2011), 700.