Do Intentions to Persist Predict Short-Term Computing Course Enrollments? A Scale Development, Validation, and Reliability Analysis

Rachel Harred
North Carolina State University
Raleigh, North Carolina, USA
rlharred@ncsu.edu

Tiffany Barnes
North Carolina State University
Raleigh, North Carolina, USA
tmbarnes@ncsu.edu

Susan R Fisk
Kent State University
Kent, Ohio, USA
sfisk@kent.edu

Bita Akram
North Carolina State University
Raleigh, North Carolina, USA
bakram@ncsu.edu

Thomas W Price
North Carolina State University
Raleigh, North Carolina, USA
twprice@ncsu.edu

Spencer Yoder
North Carolina State University
Raleigh, North Carolina, USA
smyoder@ncsu.edu

ABSTRACT

A key goal of many computer science education efforts is to increase the number and diversity of students who persist in the field of computer science and into computing careers. Many interventions have been developed in computer science designed to increase students' persistence in computing. However, it is often difficult to measure the efficacy of such interventions, as measuring actual persistence by tracking student enrollments and career placements after an intervention is difficult and time-consuming, and sometimes even impossible. In the social sciences, attitudinal research is often used to solve this problem, as attitudes can be collected in survey form around the same time that interventions are introduced and are predictive of behavior. This can allow researchers to assess the potential efficacy of an intervention before devoting the time and energy to conduct a longitudinal analysis. In this paper, we develop and validate a scale to measure intentions to persist in computing, and demonstrate its use in predicting actual persistence as defined by enrolling in another computer science course within two semesters. We conduct two analyses to do this: First, we develop a computing persistence index and test whether our scale has high alpha reliability and whether our scale predicts actual persistence in computing using students' course enrollments. Second, we conduct analyses to reduce the number of items in the scale, to make the scale easy for others to include in their own research. This paper contributes to research on computing education by developing and validating a novel measure of intentions to persist in computing, which can be used by computer science educators to evaluate potential interventions. This paper also creates a short version of the index, to ease implementation.

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

ACM ISBN 978-1-4503-9431-4/23/03…$15.00
https://doi.org/10.1145/3545945.3569875

CCS CONCEPTS

• **Social and professional topics → Adult education**; • **Computing methodologies → Feature selection**.

KEYWORDS

Persistence, enrollment, validated scale, introductory computer science

ACM Reference Format:
Rachel Harred, Tiffany Barnes, Susan R Fisk, Bita Akram, Thomas W Price, and Spencer Yoder. 2023. Do Intentions to Persist Predict Short-Term Computing Course Enrollments? A Scale Development, Validation, and Reliability Analysis. 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/10.1145/3545945.3569875

1 INTRODUCTION

A key goal of computer science education efforts is to increase the number and diversity of students who persist in the field of computer science and in computing careers. This is because there are more computing jobs available than there are people to fill them, and there is a lack of diversity in computer science in academia and industry [5].

In response, many interventions have been designed to attract and retain students in the field of computing. These interventions range from separating novice programmers from experienced programmers in introductory courses [32], to sending lightweight email interventions [13] [1], to intensive belonging interventions ([16] [35], to longitudinal research studies on improving departmental support [6]. However, it is often difficult to assess the efficacy of such interventions because measuring actual persistence can be expensive, time consuming, and even impossible. This also limits the ability of computer science education research to produce novel and innovative interventions, as researchers are unlikely to commit to the long-term evaluation of an intervention that may or may not be successful. Lastly, the difficulty of longitudinal research on persistence also limits research interventions on marginalized populations in computing, as there are often such a small number of people in these populations that amassing enough for longitudinal analysis is exceedingly difficult.

Using surveys to assess behavior has been successfully utilized by social science researchers to address these issues; particularly, in regards to investigating the potential efficacy of interventions before conducting longitudinal analysis. Attitudinal research is often used to collect leading indicators for later behavior choices, as attitudes can be collected in survey form around the same time that interventions are introduced. And indeed, "intentions to persist are predictive of actual persistence in STEM fields" [12]. In fact, "...hundreds of research efforts occurring [since the late 1960s] support the contention that intention is the 'best' predictor of future behavior" [27]. However, using attitudinal measures of persistence is less common in computing education. In this paper, we develop and validate a scale to measure intentions to persist in computing. We conduct two analyses to do this: First, we develop a computing persistence index (modified from existing scales on intentions to persist) and test a) whether our scale has high Cronbach's alpha reliability, and b) whether our scale predicts actual, short term persistence in computing using students' course enrollments within the next two semesters. Second, we conduct analyses to reduce the number of items in the scale, to make the scale easier for researchers to include in their own research. We then test whether the reduced index a) has high alpha reliability and b) predicts actual persistence in computing using students' course enrollments.

This paper contributes to research on computing education in a number of ways. The primary contribution is that it develops and validates a measure of intentions to persist in computing, which can be used by computer science educators. Often computer science educators use other attitudinal measures (e.g., belonging, self-efficacy, etc.) to assess the efficacy of their interventions; however, these attitudes are only proxies for intentions to persist, as these measures do not directly predict persistence. An attitudinal measure of persistence is preferable because it measures persistence directly. The second contribution of this paper is that our developed measure can help computer science education researchers evaluate novel, promising interventions without the time and resource commitment of longitudinal data analysis. This is especially valuable for underrepresented populations with small sample sizes. The third and final contribution of this paper is that it creates a short version of the index, to ease implementation. Adding 6 questions to a survey imposes a far smaller burden on study participants while still providing a reliable indicator. This can allow researchers to assess the potential efficacy of interventions, before committing to a long-term analysis.

2 RELATED WORK

2.1 Attitudinal Measures in Computing Education Research

Increasing retention of college students in STEM has been an important objective for researchers for many years [15] and computing education researchers have taken a variety of approaches to increase students' persistence. Of particular relevance to our research, computing education researchers have also studied how attitudes (e.g., belonging, self-efficacy, etc.) predict persistence, as well as interventions to impact such attitudes. "The most widely known and studied model of student persistence" [17] is from Tinto [42]. Tinto's model is based on the idea that integration into social and academic environments predicts student persistence in finishing college. This model has been built upon by many researchers ([15] [28] [11]). Of particular relevance to this work is Hausmann, et al. (2007), who explicitly explored students' feelings of belonging as a predictor of intentions to persist in college, with an actual persistence measure of "I intend to complete my degree at <institution>" [17]. Other research echoes this finding, as a sense of belonging was found to increase students' intention to persist in the CS major, along with satisfaction of the course [33]. Other popular frameworks that explore how attitudes influence persistence include Graham et al.'s persistence framework to explain persistence in STEM majors, as defined by obtaining a STEM undergraduate degree [15]. They describe this concept of persistence in STEM as a manifestation of student motivation and confidence. Other research supports this framework; in particular, findings that higher self-efficacy predict higher persistence, long-term interest, and learning outcomes [23][24]. Research on the impacts of attitudes on persistence in online courses reproduce many of these findings across many subjects.

However, despite the plethora of research on attitudes in computing education, there have been limited attempts by computing education researchers to use attitudinal measures of persistence. Fisk et al. (2021) and Akram et al. (2022) have used attitudinal measures of computing persistence to evaluate a lightweight email intervention, but did not have the data to evaluate whether the changes in attitudes corresponded to changes in behavioral measures of persistence [13] [1]. Lin et al. (2016) used attitudinal measures of persistence to evaluate the role of persistence on students' self-efficacy beliefs and look for differences in gender. They found that persistence levels had significant effects on self-efficacy beliefs but that self-efficacy did not vary by gender. They did not use the persistence measures to predict actual persistence [23].

Thus, computer science education researchers are typically focused on attitudinal measures that predict persistence in computing but that do not directly measure students' attitudes about computing persistence.

2.2 Previous Research on Attitudinal Measures of Persistence

As stated in the introduction, there are many reasons that an attitudinal measure of computing persistence would be useful to computing education researchers. However, such a measure would only be useful if such attitudes actually predict behavior; i.e., intentions to persist would have to predict actual persistence. In the social sciences, research has shown that intentions to persist do, in fact, predict persistence [27]. In Merolla et al. (2012), intentions to persist were measured by one question on a survey: "How likely are you to pursue a science related research career?" Persistence was measured by enrollment in one of four science training programs offered at the university. They found that the more a student identified as a scientist, the more likely they were to stay enrolled in their training program and to indicate that they would pursue a science career. Stets et al. (2017) also find that intentions are predictive of actual science behavior [38]. There has also been some research done in computer science that uses surveys to identify intentions to persist and then predict persistence with those intentions. Barker et al. investigated which factors most related to intentions to persist in the computer science major [2]. This study found that student-student interaction was the highest predictor of students' intentions to persist in the major. Another study by Katz et al. (2006) found that many factors that predicted achievement, like confidence and interest in computer science, also predicted persistence in the computer science major, measured by enrollment in upper-level computer science courses [20].

3 METHODOLOGICAL APPROACH - ASSESSING SCALES

In the present research, we use adapted survey questions from Correll [7] and [redacted] as the basis of our scale. These questions are intended to measure students' persistence intentions and include items like, "How likely are you to take another course in computer science?" and "How likely are you to apply for jobs requiring high levels of computer science ability?" See table 3 for the full set of items. Since these items have high face validity and have been used in previous research, but their predictive validity has not been assessed, they are ideal candidates to use as the basis of our scale investigation.

We next review the literature on how to assess the reliability and predictive validity of a scale and explain how this literature informs our approach to assessing our persistence scale.

3.1 Assessing the Reliability of a Scale

When creating a scale, the first thing you want to do is to make sure all of your items (in this case, attitudinal survey items) are all measuring the same construct. We use standard statistical procedures described by Cronbach [8] [30], to determine if all of the items load on the same factor. Cronbach's alpha measures the internal consistency and reliability of how closely related a set of items are as a group. We calculated Cronbach's alpha for each item in the scale and it is presented in table 2. Our internal reliability coefficient is a = 0.95 which is high and over the threshold of acceptable reliability, 0.70. In fact, with an alpha higher than 0.90, experts suggest considering shortening the scale [9].

3.2 Assessing the Predictive Validity of a Scale

A scale has predictive validity if "the test accurately predicts what it is supposed to predict" [39]. Our review of the literature found that there are numerous approaches to assessing the predictive validity of a scale: for instance, correlation of factors on the scale with the outcome measure/predicted value (Cronbach's alpha and Pearson PMC) [31] [4] [14] [34] [22] [10] [43], regression to show the scale was predictive of the target variable [31] [44] [4] [18] [21], ROCAUC [22], SEM [10], chi-square [10] [43], RMSEA [10], CFI [10]. More specifically, Morisky et. al. determined predictive validity "through association with [the measured variables] and the [target variable] including confirmatory factor analysis", then used a logistic regression analysis to see how well the measured factors predicted the target outcome [31]. In another study, predictive validity was measured by correlations of final course grade with the subscales, using factor analysis, coefficient alphas and zero-order correlations [34]. Willoughby et. al. used confirmatory factor analysis to evaluate the criterion validity of the scale and then estimated a logistic regression model to see if the scale predicted certain outcomes [44]. But what all of these approaches have in common is establishing predictive validity by showing an association between the scale and the outcome variable and showing that the scale accurately predicts what it is supposed to predict. Thus, we follow others from the literature [31] [44] [4], we conduct a logistic regression analysis to determine if our scale predicts students' actual persistence in computing (measured as enrollment in another semester of CS). We use a logistic regression because our dependent variable - enrollment - is binary. To determine what constitutes an adequate value for predictive validity, we again reviewed the literature. In general, we found that a p-value of less than .001 was the generally accepted level to meet "good" prediction values. Boateng et. al. used regression analysis to evaluate the predictive validity of their scales, with acceptable p-values of <0.001 and <0.01, with 95% confidence interval [4]. Based on these p-values, they concluded that predictive validity of the scale was supported. Willoughby et. al. established predictive validity with a logistic regression model to see if their scale predicted their target variable and with a p-value < 0.0001, it did. They concluded that their scale was predictive of their target variable [44]. In another study, Herche et al measured predictive validity using multiple regression models and reported a p-value of <0.001. With this, the authors concluded that their scale predicted their target behavior with convincing accuracy [18].

3.3 Reducing the Number of Items in a Scale

Our method for scale reduction was informed by the reduction of the PANAS scale by Thompson[40]. The PANAS scale is a measure of affect used in psychology research. In order to adapt this scale for use with international-English speakers, a short form PANAS scale was created by using exploratory factor analysis and principal component analysis techniques. The survey was effectively cut in half but still was able to adequately inform researchers of participants' affects. Here, we replicate their methods to produce a reduced scale for 1-year persistence into future computer science courses. Using R for analysis, we implemented feature selection by exploratory factor analysis and principal component analysis with varimax rotation, grouping our scale into a single factor and looking at the loadings for each question. We then used Cronbach's alpha to verify our scale and ensure the questions still held together. We tested our reduced scale by using it to predict persistence of students in computer science, measured by whether the students enrolled in subsequent courses of computer science in the two semesters following the survey. For this we implemented a logistic regression because we are predicting a binary outcome and measured the McFadden's pseudo-R^2 value to identify the amount of variance that our scale accounted for.

4 METHODS

4.1 Data collection

Over the fall 2020 and spring 2021 semesters we surveyed 892 students in 5 different introductory computer science courses: CSC 110, 111, 113 and 116 (two sections) at a southern, R1 university. Our computer science department is housed within the school of engineering and is well-established. The computer science courses

Table 1: The courses in our dataset, with their descriptions and type of students typically enrolled

Course	Content	Student Type
CSC 110	Block-based programming	General or CS beginners
CSC 111	Introduction to Python	Civil, Construction, and Environmental Engineering majors
CSC 113	Introduction to MATLAB	Aerospace, Mechanical, and Biomedical Engineering majors
CSC 116	Introduction to Java	Computer Science majors
CSC 216	Software Development	Computer Science majors

are described in Table 1. The first survey was at the beginning of the semester (T1) and collected demographic information as well as information about assessment of ability, belonging, intention to persist and feelings about the class and the class professor. The second two surveys were taken after each major grade in the course (a test or project grade, T2 and T3) and collected the same information as the first survey except the demographic information. Enrollment data was also collected on each student showing whether or not they enrolled in the next two courses, CSC 116 or 216, in the spring 2021 or fall 2021 semesters. There were no data duplicates; if a student took more than one introductory CS course, it was noted under "enrolled courses", but only data from the first survey was used.

4.2 Sample of students

Our student sample was 23% women (N = 208), 73% men (N = 647), with 4% identifying as neither or no answer (N = 37). Our participants identified as 74% White (N = 664), 12% Asian (N = 105), 4% Black or African American (N = 40), 6% Hispanic or Latinx (N = 52), 1% Native American (N = 11), 0.1% Pacific Islander (N = 1), 4% South Asian (N = 38) and 1% identified as Other (N = 12). For context, the enrolled student population at our university, both undergraduate and graduate, is 63.7% White, 7.2% Asian, 6.53% Black or African American, 5.88% Hispanic or Latino, 3.74% Two or More Races, 0.375% American Indian or Alaska Native, and 0.0777% Native Hawaiian or Other Pacific Islanders, as of 2021.

4.3 Computing persistence intentions index

As previously stated, we modeled our index off of Correll [7], and [13]. The index was created using 10 survey questions about intentions to persist in computer science, see table 2. Students were asked to indicate how likely they were to engage in future computer science activities using a 1 to 7 Likert scale. The answers were then averaged together to create a persistence index. The data is self-reported, but, as cited before, previous research has shown that "intention is the 'best' predictor of future behavior" [27].

For our persistence index measure, we ran a Cronbach's alpha test to measure the internal reliability of the different items contained therein. We received an alpha score of 0.95 which tells us that our factors are highly consistent with each other. The reliability if an item is dropped remains constant except for the factor that measures if a student decided to minor in computer science. This tells us that we may be able to remove this item from our scale and not have much change in the prediction. These scores are shown in Table 2.

Table 2: Questions in our survey measuring intentions to persist, with their alpha score if item is dropped from the index and correlation coefficient as associated with enrollment in another course

Item	Raw Alpha	Correlation with Enrollment
Take another course in computer science	0.95	0.41
Get involved with undergraduate CS research	0.95	0.44
Get involved with CS clubs	0.94	0.49
Compete in a hackathon	0.95	0.42
Apply for a CS internship	0.94	0.55
Minor in computer science	0.96	0.22
Major in computer science	0.95	0.65
Apply to graduate school in CS	0.95	0.47
Apply to graduate programs requiring high levels of CS ability	0.95	0.29
Apply for jobs requiring high levels of CS ability	0.95	0.41

4.4 Actual computing persistence

Actual computing persistence was measured by whether the students enrolled in one of two following computer science courses in the two semesters following the data collection. This measure became the enrollment measure. Out of 892 student data points, 244 students enrolled in either of the two later computer science courses. For the fall 2020 semester, 105 out of 409 students enrolled in a later course and for the spring 2021 semester 139 out of 483 students enrolled in a later course.

5 RESULTS

5.1 Do intentions to persist predict actual persistence in computing?

R^2 measures the proportion of the variation in the dependent variable that is predictable from the independent variable. For our purposes, we ran a logistic regression using our persistence index measure to predict student enrollment within 1 year (actual persistence). The p-value<0.0001 which indicates that our persistence index is statistically significant with the response variable of enrollments in CS courses within one year in the model. We can interpret that our persistence index is doing a good job of predicting enrollment in CS courses within one year. The McFadden's pseudo-R^2 is 0.34 which tells us that the index is accounting for 34% of the variation in students' actual persistence in computing. We used McFadden's pseudo-R^2 because we want to fit our logistic regression model using the method of maximum likelihood [26] [29]. When predicting human behavior, the R^2 value is usually below 50% as there are other mitigating factors that contribute to said behavior. When using McFadden's pseudo-R^2, the values tend to be lower, with values between 0.2 and 0.4 representing an excellent fit [25]. In a 2013 study on Native American undergraduate students' persistence intentions, a linear regression was used to test the extent to which each variable predicted academic persistence [41]. The adjusted R^2 for that test was 0.25 and the authors concluded that this indicated that the variables contributed significantly to the prediction of persistence intentions. In a study on identity as a predictor for undergraduate persistence intentions, the authors did block hierarchical multiple regression analyses to test their hypotheses to see how identity fit into other measures like academic and social integration in the prediction of persistence, defined here as graduating the undergraduate program [36]. They showed that

Table 3: Questions in our survey measuring intentions to persist, with the exploratory factor analysis loadings based on one factor and the principal component analysis loadings based on one factor. The two lowest loadings for each analysis are bolded.

Item	EFA Loadings	PCA Loadings
Take another course in computer science	0.79	**0.81**
Get involved with undergraduate CS research	0.89	0.91
Get involved with CS clubs	0.92	0.93
Compete in a hackathon	0.82	0.85
Apply for a CS internship	0.93	0.94
Minor in computer science	**0.60**	-
Major in computer science	0.83	0.87
Apply to graduate school in CS	0.84	0.86
Apply to graduate programs requiring high levels of CS ability	**0.76**	-
Apply for jobs requiring high levels of CS ability	0.81	**0.81**

identity had an R^2 value of 0.10 and found that identity was a significant predictor of persistence. Another study on community college student persistence found multiple factors affecting persistence, such as academic and social integration and educational objective [3]. Here, persistence was defined as enrolling in the college in the subsequent semester (from fall to spring). The cumulative R^2 for educational objective was 0.11 and the authors explain that while this is a modest value it indicates that there are other factors not captured in the model that have an effect on persistence.

5.2 Can we reduce the number of items in our persistence index and/ or improve the persistence index?

We have shown that our initial persistence index is fairly predictive; however, the large number of items contained therein might be an impediment to adopting this scale in research. Therefore, we next attempt to reduce the number of items in the scale while still maintaining a valid index. We show the correlation between each item in the scale and enrollment in Table 2. As in Thompson, et al.'s research [40], we follow the procedure to reduce our index by first conducting an exploratory factor analysis on all the items in our scale. We proceeded by attempting to combine the items into a one-factor solution. This analysis suggests that we remove the variables 'Minor in computer science' and 'Apply to graduate programs requiring computer science ability', shown in Table 3. After removing these variables, we ran a principal component analysis with varimax rotation, separating the factors into a one-factor solution, shown in Table 3. We removed the two variables least related to the single factor factors: 'Another course in computer science' and 'Apply to jobs requiring computer science ability'. The Cronbach's alpha score of our reduced scale is 0.95, showing that our scale is still reliable [19].

There is some research showing that PCA is not the best choice for doing feature selection in data analytics, suggesting that the best choice is using random forest permutation feature importance [37]. Due to this, we cross-validated our index reduction technique using permutation feature importance as well, and received the same result, with the same questions being selected to be removed. This provides more evidence that the reduced index will be similarly effective for predicting persistence as measured by enrollment in another computer science course in the next two semesters.

When we re-ran the logistic regression model using our reduced scale, we got a pseudo-R^2 of 0.36 ($p<0.0001$), showing that the reduced scale is even more predictive than the full index model [41]. Our explanation of the improvement in variance from 0.34 to 0.36 is that our new model better encapsulates the variance in our survey responses. This new reduced scale is useful because it allows us to measure the persistence of students with fewer questions, eliminating those that were not as predictive of persistence as the core questions that remain. To show that this result is generalizable, we performed our index reduction technique on one semester's worth of data, the fall semester, and applied the model to the spring semester. In the fall semester, through feature selection and PCA, we found that the same four questions were selected for removal from the index. Our logistic regression prediction improved from pseudo-R^2=0.38 using the full index model, to pseudo-R^2=0.40 (higher is better) using the reduced index model. When applied to the spring semester's data, the prediction improved from pseudo-R^2=0.31 with the full index model to pseudo-R^2=0.33 with the reduced index model. This shows that our index reduction technique is generalizable from one set of data to another. This is important because it means that others may use this technique on their own datasets and have confidence in the results.

To find the best timing of these questions, we repeated our reduction technique on the survey items at times T1, T2, and T3. We received identical results regarding which items could be removed, with improved pseudo-R^2 calculations as follows: for T1, our pseudo-R^2 improved from 0.27 to 0.30, for T2, our pseudo-R^2 improved from 0.43 to 0.44, and for T3, our pseudo-R^2 improved from 0.35 to 0.36. This also indicates that the best individual time for measuring persistence is T2. Our pseudo-R^2 for the reduced scale was greatest (0.44) at T2, so we suggest for the best predictive result future research should do a survey at the end of the first milestone project or test grade.

We next conducted a missing values analysis. As the survey times progressed, we received less and less responses. At T1, we had 764 responses to the survey. At T2, we had 660 responses and at 3 we had only 609 responses. To account for this missing data, we replaced the value for each question with the average value we calculated for each student. This means that at T2 we replaced the missing value for each question with its value at T1, and at T3 we replaced missing values with the average of any values from T1 and T2. This led to us having 764 responses at each time. We ran our index reduction technique again for T2 and T3, to get a sense of when our pseudo-R^2 value is strongest. All analysis led to the same questions being dropped from the index as before. At T2, we had a lower pseudo-R^2 value of 0.34 (from 0.44) and at T3 we had a slightly lower pseudo-R^2 value of 0.34 (from 0.36). The pseudo-R^2 value at T1 was 0.30 which shows us that the best predictor at specific times was at T2, which is when the survey was taken after the first major grade milestone in each class (large project or test grade). This also shows us that we can get a good idea about whether students would persist in taking another computer science course at that time, which allows instructors to use this information to better intervene with their students. They might choose to encourage those students who have high grades, especially those who are not predicted to persist after this second survey.

One issue with our index reduction is that all indicators pointed to removing the question "How likely are you to take another course in computer science" from our index, when our independent variable that we are using to measure persistence is, in fact, whether the student enrolls in another course of computer science. This is counterintuitive. To investigate this issue, we removed the computer science majors and any students who would most likely be seniors in class standing from our data set. We then looked at the number of students who answered this question with a 5 or higher on the Likert scale, who were indicating that they would likely take another course. There were 426 such students. However, the number of students who actually enrolled in a subsequent course of computer science was only 146. This means that there were many students who answered that they would take another course of computer science but did not enroll in one in the next two semesters. Currently, we only have data to measure persistence in the following two semesters. It could be that those students do actually enroll in another course of computer science later. Therefore, we recommend that if instructors want to include this question in their surveys, they may wish to make it more precise by asking "How likely are you to take another course in computer science in the next year?"

6 CONCLUSION

This work shows that survey questions regarding intentions to persist do, in fact, predict persistence of students in CS0 and CS1 courses into another computer science course within the next 2 semesters. We were also able to use feature selection techniques to reduce our number of survey questions and improve its prediction. This was cross-validated in a number of ways including random forest feature selection and repeating the methods on separate semesters of data to make sure the technique was generalizable. We recommend that instructors and researchers can implement this new scale to predict whether students in introductory courses will enroll in subsequent computer science courses. This is important because a shorter survey for intentions to persist allows instructors and/or researchers to more easily predict persistence and allow for intervention if needed. This is especially critical for students from underrepresented groups in computer science, as encouraging persistence may help boost those participation numbers.

6.1 Limitations

The primary limitation of this work is that we only have data for enrollment for the subsequent two semesters after the survey was taken, meaning that the scale can only be used to predict short-term enrollments. Thus, we do not have sufficient data to determine whether our scale can predict long term retention or persistence in computer science.

Another limitation is that this work was done in a very specific university context that might not hold in other universities or community colleges, or other contexts where we might want to measure persistence in computing such as K-12.

6.2 Future directions

Future work will include gathering data on enrollment in computer science courses for our participants for more semesters and re-evaluating the predictive ability of the persistence index for longer term enrollments.

REFERENCES

[1] Bita Akram et al. "Increasing Students' Persistence in Computer Science through a Lightweight Scalable Intervention". In: *Proceedings of the 27th ACM Conference on on Innovation and Technology in Computer Science Education Vol. 1*. 2022, pp. 526–532.

[2] Lecia J Barker, Charlie McDowell, and Kimberly Kalahar. "Exploring factors that influence computer science introductory course students to persist in the major". In: *ACM Sigcse Bulletin* 41.1 (2009), pp. 153–157.

[3] Trudy H Bers and Kerry E Smith. "Persistence of community college students: The influence of student intent and academic and social integration". In: *Research in higher Education* 32.5 (1991), pp. 539–556.

[4] Godfred O Boateng et al. "Best practices for developing and validating scales for health, social, and behavioral research: a primer". In: *Frontiers in public health* 6 (2018), p. 149.

[5] *Broadening participation in computing (BPC)*. URL: https://www.nsf.gov/pubs/2021/nsf21571/nsf21571.htm.

[6] J McGrath Cohoon. "Recruiting and retaining women in undergraduate computing majors". In: *ACM SIGCSE Bulletin* 34.2 (2002), pp. 48–52.

[7] Shelley J Correll. "Constraints into preferences: Gender, status, and emerging career aspirations". In: *American sociological review* 69.1 (2004), pp. 93–113.

[8] Lee J Cronbach. "Coefficient alpha and the internal structure of tests". In: *psychometrika* 16.3 (1951), pp. 297–334.

[9] Robert F DeVellis and Carolyn T Thorpe. *Scale development: Theory and applications*. Sage publications, 2021.

[10] Per Eisele. "The predictive validity of the team diagnostic survey: Testing a model with performance and satisfaction as output variables". In: *Team Performance Management* (2015).

[11] Susan A Elkins, John M Braxton, and Glenn W James. "Tinto's separation stage and its influence on first-semester college student persistence". In: *Research in higher education* 41.2 (2000), pp. 251–268.

[12] Martin Fishbein and Icek Ajzen. *Predicting and changing behavior: The reasoned action approach*. Psychology press, 2011.

[13] Susan R Fisk et al. "Increasing Women's Persistence in Computer Science by Decreasing Gendered Self-Assessments of Computing Ability". In: *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1*. 2021, pp. 464–470.

[14] Floyd J Fowler Jr and Floyd J Fowler. *Improving survey questions: Design and evaluation*. Sage, 1995.

[15] Mark J Graham et al. "Increasing persistence of college students in STEM". In: *Science* 341.6153 (2013), pp. 1455–1456.

[16] Carolyn Hart. "Factors associated with student persistence in an online program of study: A review of the literature." In: *Journal of Interactive Online Learning* 11.1 (2012).

[17] Leslie Hausmann, Janet Schofield, and Rochelle Woods. "Sense of Belonging as a Predictor of Intentions to Persist Among African American and White First-Year College Students." In: *Research in Higher Education* 48.7 (2007).

[18] Joel Herche. "A note on the predictive validity of the CETSCALE". In: *Journal of the Academy of Marketing Science* 20.3 (1992), pp. 261–264.

[19] Andrew Hoegh and Barbara M Moskal. "Examining science and engineering students' attitudes toward computer science". In: *2009 39th IEEE Frontiers in Education Conference.* IEEE. 2009, pp. 1–6.

[20] Sandra Katz et al. "Gender, achievement, and persistence in an undergraduate computer science program". In: *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 37.4 (2006), pp. 42–57.

[21] Mary C Kernan and Geoffry S Howard. "Computer anxiety and computer attitudes: An investigation of construct and predictive validity issues". In: *Educational and psychological measurement* 50.3 (1990), pp. 681–690.

[22] Daniel Kotz, Jamie Brown, and Robert West. "Predictive validity of the Motivation To Stop Scale (MTSS): a single-item measure of motivation to stop smoking". In: *Drug and alcohol dependence* 128.1-2 (2013), pp. 15–19.

[23] Guan-Yu Lin. "Self-efficacy beliefs and their sources in undergraduate computing disciplines: An examination of gender and persistence". In: *Journal of Educational Computing Research* 53.4 (2016), pp. 540–561.

[24] Alex Lishinski and Joshua Rosenberg. "All the Pieces Matter: The Relationship of Momentary Self-efficacy and Affective Experiences with CS1 Achievement and Interest in Computing". In: *Proceedings of the 17th ACM Conference on International Computing Education Research.* 2021, pp. 252–265.

[25] D McFadden. "Quantitative methods for analyzing travel behavior of individuals: some recent 857 developments: Institute of Transportation Studies". In: *University of California* 858 (1977).

[26] Daniel McFadden. "The measurement of urban travel demand". In: *Journal of public economics* 3.4 (1974), pp. 303–328.

[27] David M Merolla et al. "Structural precursors to identity processes: The role of proximate social structures". In: *Social Psychology Quarterly* 75.2 (2012), pp. 149–172.

[28] Barbara S Metzner and John P Bean. "The estimation of a conceptual model of nontraditional undergraduate student attrition". In: *Research in higher education* 27.1 (1987), pp. 15–38.

[29] Martina Mittlböck and Michael Schemper. "Explained variation for logistic regression". In: *Statistics in medicine* 15.19 (1996), pp. 1987–1997.

[30] Donald E Morisky, Lawrence W Green, and David M Levine. "Concurrent and predictive validity of a self-reported measure of medication adherence". In: *Medical care* (1986), pp. 67–74.

[31] Donald E Morisky et al. "Predictive validity of a medication adherence measure in an outpatient setting". In: *The journal of clinical hypertension* 10.5 (2008), pp. 348–354.

[32] Linda Ott, Briana Bettin, and Leo Ureel. "The impact of placement in introductory computer science courses on student persistence in a computing major". In: *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education.* 2018, pp. 296–301.

[33] Ilias O Pappas et al. "Assessing student behavior in computer science education with an fsQCA approach: The role of gains and barriers". In: *ACM Transactions on Computing Education (TOCE)* 17.2 (2017), pp. 1–23.

[34] Paul R Pintrich et al. "Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ)". In: *Educational and psychological measurement* 53.3 (1993), pp. 801–813.

[35] Rita Manco Powell. "Improving the persistence of first-year undergraduate women in computer science". In: *ACM SIGCSE Bulletin* 40.1 (2008), pp. 518–522.

[36] Thomas N Robinson III. "Identity as a mediator of institutional integration variables in the prediction of undergraduate persistence intentions". In: *Journal of Adolescent Research* 18.1 (2003), pp. 3–24.

[37] Hossein Shafizadeh-Moghadam. "Fully component selection: An efficient combination of feature selection and principal component analysis to increase model performance". In: *Expert Systems with Applications* 186 (2021), p. 115678.

[38] Jan E Stets et al. "The science identity and entering a science occupation". In: *Social science research* 64 (2017), pp. 1–14.

[39] Hamed Taherdoost. "Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research". In: *How to test the validation of a questionnaire/survey in a research (August 10, 2016)* (2016).

[40] Edmund R Thompson. "Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS)". In: *Journal of cross-cultural psychology* 38.2 (2007), pp. 227–242.

[41] Mindi N Thompson, Michelle Johnson-Jennings, and Rachel S Nitzarim. "Native American undergraduate students' persistence intentions: A psychosociocultural perspective." In: *Cultural Diversity and Ethnic Minority Psychology* 19.2 (2013), p. 218.

[42] Vincent Tinto. *Leaving college: Rethinking the causes and cures of student attrition.* ERIC, 1987.

[43] Roger Tourangeau and Kenneth A Rasinski. "Cognitive processes underlying context effects in attitude measurement." In: *Psychological bulletin* 103.3 (1988), p. 299.

[44] Michael T Willoughby et al. "The measurement of executive function at age 3 years: psychometric properties and criterion validity of a new battery of tasks." In: *Psychological assessment* 22.2 (2010), p. 306.