Self-efficacy, Interest, and Belongingness – URM Students' Momentary Experiences in CS1

Alex Lishinski
alishins@utk.edu
University of Tennessee
Knoxville, TN, USA

Sarah Narvaiz
snarvaiz@vols.utk.edu
University of Tennessee
Knoxville, TN, USA

Joshua M. Rosenberg
jmrosenberg@utk.edu
University of Tennessee
Knoxville, TN, USA

ABSTRACT

Educational stakeholders want to understand and overcome the well-documented racial and gender disparities within computer science education. There are many factors that influence students' participation, performance, and persistence in CS courses, including motivational and affective factors. Prior research in CS education has documented the influence of these factors on students' CS outcomes generally, and on URM students in particular. What has been less investigated, is how students' motivational and affective experiences in CS develop and evolve from moment to moment, particularly for URM students. To better understand how these experiences develop, this paper presents the results of a study using intensive longitudinal methods which examined the differences in momentary experiences between racially underrepresented students and their represented peers in undergraduate introductory computer science courses. Using the Experience Sampling Method (ESM), we solicited responses on students' momentary self-efficacy, interest, and affective experiences 8-18 times in each of two semesters from a total of 110 CS students, of which 19 identified as racially underrepresented.

Analyzing the data using a Bayesian multivariate multilevel modeling approach, we found that students' pre-semester self-efficacy impacted their momentary frustration, interest, and self-efficacy throughout the duration of the semester, as expected. Likewise, students' baseline interest significantly impacted their momentary self-efficacy and interest. Baseline sense of belonging, by contrast, showed no significant impact on their momentary experiences. We also examined how students' affective experiences relate to course outcomes and found that baseline self-efficacy significantly predicts end-of-course grades for URM students. Overall, this study highlights that students' self-efficacy and interest are important for their momentary experiences, and course outcomes for URM students, while their sense of belonging did not make a significant impact. We expected that these influences might differ in magnitude for URM students, although a larger sample size and greater statistical power is needed to substantiate this possibility. Nevertheless, the findings from this study emphasize the importance of self-efficacy, and interest for URM students' momentary experiences, which are important for their other outcomes in CS classes.

KEYWORDS

CS1, Experience Sampling Method, Self-efficacy, Interest, Frustration

ACM Reference Format:
Alex Lishinski, Sarah Narvaiz, and Joshua M. Rosenberg. 2022. Self-efficacy, Interest, and Belongingness – URM Students' Momentary Experiences in CS1. In *Proceedings of the 2022 ACM Conference on International Computing Education Research V.1 (ICER 2022), August 7–11, 2022, Lugano and Virtual Event, Switzerland*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3501385.3543958

1 INTRODUCTION

Students from minoritized groups are not adequately represented in computer science (CS). Significantly fewer students of color take the Advanced Placement CS exams in US high schools (although the gap is shrinking) (Code.org, 2019). As of 2016, only 10.1% of the computer science bachelor's degrees in the US were awarded to Hispanic students and only 8.6% to African American or Black students (National Center for Science and Engineering Statistics, 2019). Enrollment of students of color has increased in recent years, but there are still concerns about attrition rates following introductory CS (CS1) courses (Computing Research Association, 2017). One factor that is well-researched with relation to students' success in CS is students' prior programming experience [11, 101, 103]. Broadening participation in CS will involve multiple approaches aimed at different parts of the problem, but particularly with regard to persistence beyond CS1 classes, it is essential that we better understand and support the factors which promote positive experiences in these classes. Improving the pedagogical practices in CS1 courses will help improve the outcomes for everyone, but some approaches may be particularly beneficial for addressing these participation gaps in CS. In order to advance our understanding of how to best improve persistence and broaden participation in CS education, much attention in computing education research has focused on understanding students' educational experiences framed as large chunks of time, like a whole course. This approach examines different student, instructor, school, and other broad contextual elements to understand how these various components predict or influence student outcomes in CS courses (e.g., [13, 23, 35, 61, 75, 77, 90]). While this broad approach is suitable for many contexts–particularly for providing insights into the general factors that impact what students might take away from an entire course–it can also be very useful to look at students' experiences on a more granular level, as has been done in undergraduate classes for other disciplines (e.g.,

ICER 2022, August 7–11, 2022, Lugano and Virtual Event, Switzerland

ACM ISBN 978-1-4503-9194-8/22/08…$15.00
https://doi.org/10.1145/3501385.3543958

[87] in science education), but such work is limited in computing education research.

For CS education, the limited insight into how students' classroom experiences influence their development of robust interest in the field, which will encourage them to seek sustained engagement. Prior research has shown us things like curricular elements that promote higher levels of interest for students (e.g. [34, 45], but in all such cases, there is an underlying mechanism connecting curricular inputs to students' outcomes that we often have little understanding of. Students' behaviors, thoughts, beliefs, emotions, and motivations, along with those of instructors, and other elements of the instructional context and modality all affect one another in a complex relationship that is certainly dynamic over time [54, 71]. Understanding every possible element of this picture is overwhelming, but zooming in on students' experiences to gain a better picture of what they are experiencing across time can shed light on why certain students in certain situations develop an interest in computing, while others do not. This method is particularly helpful to better understand how to support students from underrepresented groups, because not only might they have different experiences, the same experiences might influence their outcomes differently [94].

We know that students' beliefs and attitudes are strongly related to their outcomes in CS (e.g., [56, 90]). This is not surprising, and it is consistent with education research outside CS as well. We have less understanding of the details of how exactly students' experiences in class contribute to shaping and developing these beliefs and attitudes. We have a good appreciation for the importance of the granular learning experiences that students have in class. However, we must not neglect the importance of the granular motivational and affective experiences that students have. Experiences of frustration, accomplishment, motivation, discouragement are not even to be construed as parallel threads to the learning experiences that students have, but as an inherent part of the learning experience [9, 49], even though our attention in pedagogy tends to skew towards the more mechanical part of the learning experiences. These affective classroom experiences can be thought of as parallel outcomes to the learning ones that students achieve during a course of study, and they also have a significant impact on the more distal outcomes of interest and persistence in the field.

These motivational and affective experiences are important for all students but may be particularly important for supporting students from underrepresented groups and closing the participation gaps in CS. For example, self-efficacy, a student's belief about whether they can succeed in a given domain, is an important student motivational factor. For underrepresented students, research has found that self-efficacy mediates the impact of efforts that have been shown to increase their persistence in STEM, such as mentoring and research experiences [94]. This suggests that in addition to such efforts, we need to create experiences in the classroom that bolster self-efficacy. Self-efficacy beliefs impact many behaviors and attitudes that are related to achievement and persistence, so supporting students' self-efficacy will certainly be a part of addressing the participation gaps in CS.

In order to investigate the details of students' experiences and how they flow into each other and develop over time, it was necessary for us to collect data from students at frequent intervals. To that end, this study involves an intensive longitudinal data collection making use of the Experience Sampling Method (ESM) [37], to study the motivational and affective experiences of students in a more granular–rather than in a global–way. This type of rich repeated measures is well suited to measure constructs that change frequently over time, like emotions. In addition to this benefit of ESM, we were able to make the measures maximally contextual, targeting our data collection to moments of particular salience for CS students. This type of data collection aims for a greater level of depth than other forms of quantitative research, while still providing the breadth of insight achieved by modeling quantitative data. As we gain more understanding of the details of how students' experiences connect to each other and develop over time, and how they differ between students, future research can start to get a better understanding of why certain pedagogical approaches might work better than others, and how this might differ for different students.

1.1 Purpose of study

Prior research has not examined granular experiences in students' motivational and affective experiences, nor how these experiences differ for URM students. We want to better understand these because we think that these fine-grained details of student experiences will shed light on the ways that we can better support underrepresented students in computing, in order to work towards the goal of broadening participation in the field. By better understanding differences in the momentary experiences students have, and how these experiences are related to baseline motivational and emotional factors, we will have a more principled, empirically driven way of determining what sorts of pedagogical approaches are likely to improve the quality of experiences for these students.

To investigate the experiences of underrepresented students in CS classes, we conducted a longitudinal ESM data collection targeting motivational and affective components of students' experiences. The data analysis and discussion in this study are focused around the following research questions:

- RQ1: How do racially underrepresented students' momentary experiences differ compared to their represented peers enrolled in CS1?
- RQ2: How do baseline levels of sense of belonging, interest value, and self-efficacy impact the momentary experiences of underrepresented students' when compared to their represented peers enrolled in CS1?
- RQ3: What are the impact of these momentary experiences on course grade outcomes?

2 LITERATURE REVIEW

2.1 Self-efficacy

Self-efficacy beliefs are beliefs that an individual has about what they will be able to accomplish [4]. Greater self-efficacy means that an individual has more positive beliefs about their own ability to successfully perform the behaviors that will lead to success in a given endeavor. Self-efficacy beliefs are context-specific, which in education means that individuals have different sets of self-efficacy beliefs about their abilities to achieve success in different academic subjects [5, 99]. The self-regulated learning (SRL) model, which characterizes how students use behaviors like metacognition and strategies to guide themselves through the learning process, includes self-efficacy as a component, and its importance derives from the fact that those with greater self-efficacy are willing to expend greater effort and display greater persistence in overcoming difficulties [72, 88]. SRL is an iterative cycle of planning and forethought, performance, and reflection [105]. Self-efficacy beliefs influence students' behaviors at each stage of this cycle. They influence goals and planning in the forethought stage, and attention focusing and learning strategies in the performance stage, and are in turn revised by students' during the reflection stage in light of their outcomes and experiences during the learning process [71]. As such, self-efficacy beliefs are not a static construct for students that are stable over time; they are continuously developing and changing over time for each individual for each type of task or larger subject matter domain as they gain more experiences. And because they are both a precursor and a product of the self-regulated learning cycle, the development of self-efficacy is a reciprocal feedback loop process, which can be positive when students are in an environment that promotes experiences likely to bolster self-efficacy, or negative when students have experiences that erode self-efficacy. Self-efficacy has a robust empirical foundation that has shown it to be strongly related to student outcomes and persistence across student populations [15, 40, 67, 82, 98].

Self-efficacy is one of the most well-studied motivational constructs in education research; it has also often been the subject of research in CS education, particularly in recent years. Researchers in CS education have found that self-efficacy is strongly related to student outcomes in CS [52, 79, 101, 102], and they have also worked to develop better ways to measure the specifics of self-efficacy in CS contexts [7, 18, 92]. CS education researchers have also implicated self-efficacy as a contributing factor to the participation gaps in CS. Self-efficacy influences both short term academic outcomes, such as grades, but it also influences longer-term outcomes like persistence [41, 53]. This is particularly important in the context of CS1 because self-efficacy beliefs formed early in CS1 can have significant impacts (both positive and negative) on students' performance over the length of the course [55, 59]. For this reason, it is necessary that we create self-efficacy bolstering experiences for students in CS1, and to do so, it is necessary that we pay particular attention to the unique challenges experienced by women and students of color.

Furthermore, prior research has shown a need for greater support of self-efficacy for underrepresented groups in CS. For students of color, research has found that self-efficacy mediates the impact of efforts that have been shown to increase the persistence of these students in STEM, such as mentoring and research, which suggests that improved self-efficacy is the underlying reason for better outcomes resulting from such efforts [94].

Previous work has shown that self-efficacy is important for student outcomes in CS. What has been less explored is self-efficacy as an outcome in and of itself, one that will affect longer-term outcomes such as persistence. We also have little understanding of how students develop their self-efficacy on a more granular moment-to-moment basis and how this is related to other baseline characteristics of students. A better understanding of these would shed light on how to support different groups of students to develop their self-efficacy in CS and enable them to participate in CS as much as they want to. This is particularly important in light of recent research that shows that CS students already make judgments of their own ability that skew negative and are often detrimental [32, 33]. In sum, though research on students' self-efficacy in CS education contexts has been carried out, there is the need for further work that examines the experiences of underrepresented students–and so in a more granular, rather than global (across the entire course or semester) manner.

2.2 Affective Experiences - Frustration

Similar to motivational characteristics like self-efficacy, students' affective experiences are an important influence on the quality of their academic experiences, given the ways in which research has shown that the emotional character of students' experiences influences behavior [74]. The notion of affective load refers to the way that students must manage affective states while learning, analogously to how cognitive load refers to the ways that students must manage the cognitive demands of a learning situation [68]. Affective load is very important to consider in CS education because affective states and experiences significantly influence attention and other cognitive processes and this is especially important in a domain focused on problem solving.

Frustration is one type of affective experience that has been found in prior research to be central to the ways that affective experiences influence behaviors in the SRL cycle [44, 66], as well as in problem solving based activities [19, 36]. Experiences of frustration are of particular importance to computer science students. Frustration has been noted as one of the most common affective experiences experienced by CS students, and the process of working through programming tasks has been typically modeled as cycling back and forth between experiences of success and frustration [8, 9, 48, 49]. Even professional software engineers experience strong frustration, with a survey of 45 software engineers finding that 67% report experiencing severe recent frustration [26]. Furthermore, research has shown that early experiences of frustration can impact students' outcomes much later in CS courses [58]. Other research by Eckerdal et al. [22], has found strong negative experiences of frustration to be common in connection with learning threshold concepts, those that are central to the discipline. Thus, examining the affective experience of frustration–particularly through a method sensitive enough to detect variation in frustration over time and across assignments–is important for studies intended to explore the experiences students have in CS classes.

2.3 Interest

Educational psychology scholars have provided multiple theoretical frameworks for the deceptively simple concept of interest. These frameworks, focusing on a different component or understanding interest through a different lens may all be useful in different situations. There are models of interest which describe it as developmental stages [38], an emotional state [2], and beliefs about a potential object of interest [85]. Like self-efficacy, having an interest in a subject increases the SRL behaviors that an individual will engage in. On all of the different models of interest, however, interest is not a static trait of an individual. The model of interest proposed by Hidi and Renninger [38], their four-phase model describes interest as a developmental process, where individuals grow from an initial spark of interest until their interest is well-developed and self-sustaining. These four stages are, in order, triggered situational interest, maintained situational interest, emerging individual interest, and finally well-developed individual interest. Individuals start where their interest in a subject is triggered by characteristics of the task and situation, which causes them to enter a temporary state of increased attention and engagement. This state of interest can then, in subsequent situations, be induced to last a longer period of time, until repeated exposures lead to the individual internalizing the temporary psychological state into an enduring predisposition, in which individuals have internalized knowledge and conception of value towards the subject, greater positive feelings and a self-sustaining desire to continue to engage.

Interest has been often invoked in prior CS education research, but less often through the lens of one of these formal frameworks from educational psychology. Oftentimes interest is treated as a simple outcome that allows a pedagogical approach or outreach effort to be evaluated. For example, some previous studies have examined different novel computing curricular approaches and found that they were able to successfully increase students' interest [34, 45]. The concept of interest has also often been invoked in studies seeking to understand underlying factors in the participation gaps in CS. For example, a study by Cheryan et al. [13] showed the way that stereotype cues lead to lower interest in CS for women at the undergraduate level. Other studies have looked at other influences on interest for women and other minoritized groups [12, 20, 46].

What these prior studies have not examined is the development of interest on a finer-grained level, how it develops as a result of students' experiences on a moment-to-moment basis. This can be very useful for understanding students' development of interest in CS, if we uncover more about how we support students from the initial point of contact with students where interest is triggered, all the way to a well-developed individual interest in CS if that's what they desire. It remains unclear how we can best do this, but insofar as efforts to broaden participation want to go beyond initial experiences of interest, but to longer-lasting individual interest that propels students into deeper, longer-term engagement with CS. Students going from phase 1 to phase 3 over the course of a CS1 class is not out of the question, and while some students will make it there even with minimal attention paid to supporting the development of their interest, if we want to help as many students as possible reach their potential in CS, particularly those from underrepresented groups that already face many obstacles, this is an important issue to research and better understand in a CS context.

2.4 Sense of Belonging

Sense of belonging is a well-researched motivational construct which, within educational contexts, typically refers to a students' feeling of being accepted, supported, and included by peers, faculty, and institutions [31]. Specifically, using Tinto's retention model [96], studies have investigated how sense of belonging impacts student retention [3, 39, 65, 70, 95, 97]. Sense of belonging has also been studied for students participating in STEM studies. An overwhelming majority of studies have found that sense of belonging has larger impacts on outcomes such as retention and achievement for underrepresented students in the STEM field [93]. Specifically, women and students of color in STEM report lower levels of sense of belonging compared to their male and White counterparts [30, 43, 91]. Sense of belonging is a useful construct for education researchers to use to understand how institutions of higher education can better support students throughout the duration of their studies, as it is known to be one of the most significant factors in student success and persistence [1].

Some scholars have focused explicitly on underrepresented students' sense of belonging in STEM education contexts. Specifically, to better understand the factors that promote belongingness for underrepresented students, Rainey [78] interviewed around 200 undergraduate STEM majors who primarily identified as female or a person of color. Students reported that their frustration in understanding concepts within their studies was one of the primary factors which either increased or decreased their sense of belonging. Specifically, underrepresented students who left their STEM studies cited a lack of interpersonal relationships and low competence as the reason for their attrition. Similar findings were reported in Strayhorn's study of more than 750 undergraduate students in STEM areas of study across six large public universities in the United States. Strayhorn found that racially underrepresented students reported significantly lower levels of sense of belonging compared to their represented counterparts after controlling for academic backgrounds [93].

While few, more recent studies within computer science have examined the sense of belonging for underrepresented students. Sax et al. [84] examined sense of belonging for underrepresented students in introductory computing classes and found results that conflicted with prior research on URM computing students and sense of belonging. Results showed that women had lower levels of sense of belonging at the start of the semester whereas racially underrepresented students' sense of belonging did not differ from womens' male counterparts. However, Nguyen and Lewis [69] analyzed a large sample of freshman computer science majors' sense of belonging and found that women reported lower levels of sense of belonging compared to men and racially underrepresented students reported lower levels of sense of belonging than their represented counterparts. Similarly, Walton and Cohen found that women reported lower levels of belongingness compared to men, and Asian and Black students reported lower levels of belongingness compared to White students [100]. Lastly, Krause-Levy et al. [50] found no differences in the sense of belonging for students who identified as Black, Latinx, Native American, and Pacific Islander compared to students who did not identify with these groups. Given the mixed findings on racially underrepresented students' sense of belonging within computing courses, additional research in the area is needed.

2.5 URM Experiences in CS

While the inclusion of women and people of color in STEM fields is improving, these groups have historically been underrepresented (and even systematically excluded) from most STEM fields [81]. Given the systemic exclusion of underrepresented groups, researchers have sought out to study factors, such as environmental backgrounds and psychosocial factors, that allow these groups to enter and persist within STEM. Many researchers have found that students' psychological factors, their "academic mindset," is an important factor for students to persist in the STEM field [80, 100], including feeling accepted in their academic environment [42, 78]. As outlined in Lytle et al.'s review [60], several psychosocial factors are important to consider when addressing retention and engagement issues for STEM undergraduate students, including the aforementioned factors of self-efficacy and sense of belonging among others.

Students from underrepresented groups have been inadequately represented in undergraduate computing programs [64, 107]. Not only are these groups underrepresented they also tend to have lower retention rates compared to their White and Asian male counterparts [106]. To better understand and begin to quell the barriers of entry and persistence, more recent research has focused on studying these groups' affective experiences. Salguero et al. [83] analyzed undergraduate survey data enrolled in introductory CS courses to understand the multidimensional challenges students face. The analysis revealed that low sense of belonging and low confidence (or self-efficacy) were among the factors that students attributed to challenges faced during their semester enrolled in CS. Other studies examining underrepresented students, women and racially diverse students in particular, have noted that these groups are at risk for starting with and lowering their self-efficacy given the stereotype of the identity that computer scientists are White or Asian male [24, 25]. Given that psychosocial factors including sense of belonging are linked to retention [17, 50, 78, 93], additional research is needed to understand these groups' affective experiences while enrolled in an introductory computing course.

3 METHOD

3.1 Context and Participants

This paper describes a study taking place in two semesters across the 2020-2021 academic year at a large public university in the Southeastern United States. There were a total of 110 students who participated in this study. These students were recruited from three different introductory programming courses. The first class (CS1A), from which 77% of the students were drawn, is an introductory undergraduate course, taught in C++, which serves primarily computer science and other engineering majors, and is part of a larger sequence of CS courses intended for that audience. The second class (CS1B), from which 12% of the students were drawn, is an introductory undergraduate course taught in python, which serves an audience of students outside of computer science and engineering who need a broader standalone introduction to CS. The third class (CS1C), from which 11% of the students were drawn, is an introductory course serving an audience of graduate students in science and engineering disciplines who need a broad introduction to CS for similar reasons as the students in CS1B. Students' status as being from an underrepresented racial/ethnic group was determined by the student survey data in which students reported if they identified as underrepresented and 17% of the students were defined as such. Each of the three CS1 courses were taught by a single instructor, and the differences between courses were accounted for in our analysis.

3.2 The Experience Sampling Method

The Experience Sampling Method (ESM) is an intensive longitudinal method that allows researchers to collect detailed data about the experiences of individuals' daily lives at a level of granularity not normally possible with typical forms of survey research [37]. ESM offers a number of benefits over other forms of survey research, first and foremost being increased ecological validity. ESM data collections involve surveying students who are actively in the process of carrying out daily life activities of interest, allowing researchers to survey participants in the moment and get a more accurate accounting of feelings, behaviors, and activities than is typically available in surveys that rely on participants retrospective recall and/or summarizing over many instances. ESM also offers the advantages of being minimally intrusive in participants' natural behaviors in the same way as a diary, while at the same time taking advantage of the ability to use standard survey items with documented measurement properties. This research approach then lets researchers get at a level of detail on participants' experiences through repeated measures that is reminiscent of qualitative research, but at a scale typically infeasible in most qualitative research studies.

For all its benefits, ESM research does have certain drawbacks, like all research methods and these must be weighed against the advantages. Major issues for ESM research can include participant self-selection bias and respondent fatigue [89]. As in many ESM studies, our data collection was inherently limited to the extent that our participants had to be recruited to participate in an extracurricular activity that would require some time and effort from them, despite our efforts to minimize these. So our study and the data we collected is limited, like all observational studies, by the fact that our student pool is a subset of the larger population of interest that was willing to participate in something that didn't directly benefit their course outcomes. Respondent fatigue is also an issue that we encountered in our data collection as well, as most students didn't complete all of the surveys that we sent to them. We think, however, that this is less of an issue in our data collection than in most ESM studies, as we sent only 1-2 surveys in any given week over the course of a whole semester, whereas many ESM data collections involve several surveys a day for several consecutive days without suffering in terms of validity and reliability [16].

3.3 Data Collection

A central consideration for the design and implementation of our ESM data collection was minimizing the burden on students to participate. By making the data collection as seamlessly integrated into students' normal routines as possible, we thought we could best minimize the risk of self-selection bias and respondent fatigue compromising the quality of our data. Over the now decades-long history of ESM research, increasing technology has improved the ability of researchers to collect ESM data in an increasingly unobtrusive way. One goal of our ESM research program was to further optimize the data collection procedures to further advance in this direction. To that end, we used a system based on conducting ESM surveys via text messages sent directly to students' phones. Whereas most modern ESM studies tend to conduct their data collections using purpose-built apps, we thought this approach, while much more unobtrusive than paper and pencil data collections with researcher-provided signaling devices, could nevertheless be further optimized.

We developed a web application that would allow us to send our surveys via text messaging rather than relying on a 3rd party app. One advantage of the app that we developed is that it allowed us complete control over all aspects of the data collection. This advantage did not by itself motivate the time spent developing the app, as the available purpose-built mobile apps for this type of data collection are relatively full-featured and customizable. However, the major advantage of the full customizability of the app is that it allowed us to change the nature of our data collection to suit our needs.

Our original use of the app for sending ESM surveys was to send timed survey prompts, based on the times that students were finishing their various class sessions. However, this was disrupted by the onset of the COVID-19 pandemic which switched instruction to online, not just in the Spring 2020 semester, but in the whole following academic year in which this data collection took place. The switch to online instruction changed the modality of the introductory programming courses that we studied to asynchronous, which meant that there were no particular times that could be assumed to be salient for all students. So to address this change, we had to change the functionality of the app to allow us to use a more dynamic, event contingent survey collection plan. To that end, we connected the app with the course LMS to determine salient moments when students were working on class activities, triggering surveys to be sent out when students submitted certain assignments to the LMS. This way, although students may have answered the surveys at different times across the semester relative to one another, these instances were analogous across students because they were keyed to specific assignments that students worked on.

In the case of this study, the data collection plan was as follows. For each of the three CS1 courses participating in the study, we selected course assignments that met two criteria. First, these assignments needed to be substantive programming assignments where students wrote full programs. Second, the assignments were associated with submissions in the course LMS. Given the criteria, we identified 14, 11, and 3 such assignments for CS1A, CS1B, and CS1C respectively during the spring 2021 semester, and 13 and 7 such assignments for CS1A and CS1B in the fall 2020 semester. When the course LMS received a submission on one of these assignments from the students participating in our study, our app sent a survey prompt for our ESM survey approximately 5 minutes later. Subsequent submissions of the same assignment did not trigger additional surveys. Submissions were open to students ahead of the due dates, so there was some variability in when students actually completed the assignments, which we accounted for in our models by including a fixed effect for the day the survey was actually completed, rather than modeling based on the nominal due date. In addition to these surveys that were associated with particular assignments and triggered by student submissions, we also sent out 5 surveys at fixed times. In other words, these 5 surveys were sent to all students at the same time, at a time chosen to be unlikely to coincide with CS1 activities, and were equally spaced across the semester. The response rate for all ESM surveys was 77.1% over the two semesters. This means that 77.1% of our outgoing survey prompts were answered. The survey data also showed that the connection of surveys to course assignments was a more effective modality relative to fixed time surveys, as the response rate for the fixed time surveys was 64.4% versus 81.9% for the surveys linked to assignments.

Beyond the additional flexibility afforded by using the app for our data collection, it offered some important advantages in minimizing the burden of the data collection on students. Rather than requiring them to install and use a third party ESM app, the use of our app, leveraging the text message interface on their phones that students are already familiar with, made it easier for students to respond to our survey questions. Furthermore, beyond the use of a familiar interface, we think that the use of simple text messages in telling us about personal feelings, which can be a sensitive topic for some students, made the whole process less clinical and intimidating than responding in an app, as students are very likely used to using text messaging in all facets of their personal life. Furthermore, we thought that it would be easier for students to simply tune out the notifications from an app, whereas we were more likely to get their attention with a text message. Of course students were repeatedly told that they were free to opt-out of the study at any time for any reason, and our text message surveys respected the standard SMS opt-out keywords, so it was extremely easy for students to opt-out if they no longer wanted to participate.

3.4 Measures

Our data collection consisted of two major components. The ESM surveys took place during the semester and prior to the onset of ESM data collection students completed a pre-survey. The ESM surveys (included in Appendix A) were conducted via text message using our app, and these consisted of three Likert scale items corresponding to a momentary experience of interest that we were trying to capture. The pre-survey was administered on an online survey platform and consisted of multiple scales corresponding to constructs of interest, in addition to other questions that allowed us to connect students' responses across the semester, and questions soliciting feedback on the data collection itself from students.

The ESM questions asked students to tell us about their momentary experiences in their introductory programming courses on three dimensions. We queried their frustration and interest by asking them to rate on a Likert scale their current level agreement with the statements “I feel frustrated” and "I feel interested in computer science" following the same simple template for asking about momentary affect and attitudes as many other ESM studies [6, 47]. We asked about their momentary self-efficacy using an item that we adapted from the Motivated Strategies for Learning Questionnaire (MSLQ) self-efficacy scale [76], which read “I feel confident about being able to do the work going forward.” All of the items were rated by students on a 1-5 scale where 1 indicates strong disagreement and 5 indicates strong agreement. Each time students received the survey, they were also given this prompt prior to rating the items: “Please indicate your agreement at this moment with the following statements about your experiences in CS1." See Appendix A for exact wording of survey prompt and items. The constructs were measured at each occasion using a single item rather than longer scales or subscales in order to make responding to the surveys as efficient for students as possible. This is consistent with the prevailing practice in ESM research and is supported by methodological work showing that correlations of single items with constructs of interest are very similar to correlations of full scales, and as such the answers to substantive research questions are the same.

The pre-survey consisted of full scales of the constructs discussed in section 2 (self-efficacy, interest, sense of belonging), which we used as baseline measures in building our models of students' ESM experiences. Each of these scales were set up to use the same 5 point Likert scale as the ESM items where 1 indicates strongly disagree and 5 indicates strongly agree. For self-efficacy, for which we used the 8 item scale from the MSLQ. This scale asks students to rate their self-efficacy beliefs relative to a specific course without mentioning specific CS content (e.g. "I expect to do well in this class"), which contrasts this scale with some of the more recently developed self-efficacy instruments in CS education research which key students' self-efficacy beliefs to their confidence about specific CS concepts [92]. While the latter type of self-efficacy instrument would be an excellent end of course outcome for students' self-efficacy, the more generic forward-looking scale was more appropriate for use as a beginning of semester baseline measure. We measured students' baseline level of interest in CS using a scale measuring the interest value that students perceive towards a subject, using a scale drawn from Conley [14], which contains a set of task value subscales (interest value, utility value, attainment value, and cost) that come from an expectancy-value framework, and which were originally developed by Eccles and Wigfield [21]. These scales measure different reasons that students might value CS as a field, with the interest value subscale used in this analysis focusing on the degree to which students valued CS due to an attraction to its inherent qualities, as opposed to valuing it as a means to an end or because of how it would reflect on their identity. We also used a set of 8 items measuring students' sense of belongingness, which we adapted from items used by Mendoza-Denton et al. [63]. These items ask students to rate their feelings of belongingness from a few different angles (e.g. how happy they feel to be there, how welcome they feel, how comfortable they are around peers). We modified these items, not for their content, but rather for their format, so that responses were consistent with the other Likert type scales used here. Whereas the original items as used by Mendoza-Denton et al. [63] encoded the unique information for each item into the scale itself (e.g., How do you feel towards the university? On a scale from thrilled to be here to miserable; on a scale from definitely fit in to do not fit in, etc.), we converted these to simple declarative sentences (e.g. I am happy to be here; I fit in, etc.) that we could use with the same 5 point Likert scale mentioned previously.

3.5 Data Analysis

The models that we fit to answer our research questions are Bayesian multivariate multilevel models. The multilevel component consists of the fact that we have repeated measures nested within students. The multivariate component consists of the fact that we have multiple outcome measures - the three ESM variables, as well as an end-of-course grade outcome. The Bayesian component of our analysis is that the models were fit using Bayesian estimation methods, and our interpretation of the model parameters thus estimated was also done within a Bayesian statistical inference framework.

Our research questions deal with understanding how the experiences of URM students, as they relate to the experiences of momentary frustration, self-efficacy, and interest differ from those of their non-URM peers. We also investigated whether the baseline factors from the pre-survey were substantively different in their influence on the experiences of URM students. To answer these questions we also present results about the broader character of experiences in CS for the broader group of students.

We paid particular attention in setting up the models to the structural elements that would be included, as there were some complex features of the data that these models would let us account for. To account for the within-subjects dependency, whereby an individuals' own repeated responses are more similar to each other than those of others, we included in the model a random effect for each person, as is commonplace in analyses of ESM data [37]. We also accounted for groupings in the data based on commonalities within specific assignments with a random effect for assignment. Though less common in analyses of ESM data, a number of scholars have argued for the importance of modeling such systematic variation (e.g., [86]), and we heed those calls in our analysis.

Furthermore, we accounted for differences in intervals between measures and different schedules for different people by adding a fixed effect for the day of the semester on which each survey took place. Finally, because prior analyses of ESM data suggests that students' responses nearer in time were more likely to be similar than those that were further spaced out [73], a feature of the data known as serial autocorrelation, we added autocorrelation terms within-person to the model. Though these models are clearly complex, we think that this complexity is warranted given the nature of ESM data, including the data we collected in CS1 classes. We also think the complexity of the models does not obscure the interpretability of the analyses which we present next. After estimating the model parameters expressing the connection between our variables, we tested whether these values were significant using a Bayesian inference procedure based on the Region of Practical Equivalence (ROPE). By contrast with the frequentist parameter null hypothesis-testing procedures, this Bayesian inference procedure selects a region around a value that you want to test against that would be practically equivalent to that value called a ROPE. Then the hypothesis is tested by determining whether or not there is overlap (to a predetermined threshold) between the Bayesian Highest Density Interval (HDI) and the ROPE [51]. To estimate these Bayesian multivariate, multilevel models, we used the brms R package [10], and ROPE testing and associated plots made use of the bayestestR and bayesplot R packages [27, 62].

4 RESULTS

In order to understand students' momentary experiences and the differences in these experiences for URM students and their peers, all models discussed below included student and assignment as random effects, autocorrelation was accounted for. To address our research questions, covariates for students' baseline levels of self-efficacy, sense of belonging were included as fixed effects. For ease of model interpretation, all continuous variables (pre-survey baseline measures as well as the repeated measures ESM variables) were standardized for all of the models to have the properties M = 0, SD = 1.

We first fit a "null" model, which included only the random effects terms, in order to estimate the variability (in SD units) in each of the three outcomes at the assignment, student, and response level (the residual). These can be transformed into Intra-Class Correlation (ICC) statistics that help to explain how much variation is present at the assignment and student level and how much is then left over after accounting for those factors. We found that there was substantial variation at the student level, with ICCs ranging from a low of .48 for frustration (indicating that 48% of the variation in reports of frustration was between students) to .50 and .57 for self-efficacy and interest, respectively. There was less but still noteworthy variation at the assignment level, from a low of .12 for interest to .13 and .17 for confidence and frustration, respectively. Overall, this suggests that there was substantial variation between individual students and assignments. Generally, frustration varied more between assignments (and less between individuals) than self-efficacy and especially interest, suggesting that frustration is somewhat more malleable than interest, which tends to be more stable for individual students. This also suggests that accounting for both of these sources of variation is important, and not doing so can bias the estimates from the model [28].

The following models outlined in the remainder of this section correspond to each of the three research questions: model 1 addresses RQ1, model 2 addresses RQ2, and model 3 addresses RQ3. For model 1, students' momentary experiences were investigated by examining the random effects, residuals, and fixed effects, which included categorical variables accounting for differences between the three CS1 courses, differences between the two study semesters, and differences between non-URM and URM students. For model 2, we added the covariates for students' baseline measures of interest value for CS, self-efficacy, and belongingness to examine how these factors moderated students' momentary experiences. For the final model, model 3, we further built upon model 2 to examine the relationships of baseline covariates and ESM experiences with student outcomes.

4.1 RQ1: Differences in Momentary Experiences

To answer RQ1 (*How do racially underrepresented students' momentary experiences differ compared to their represented peers enrolled in CS1?*), we examined the differences in momentary experiences between CS1 courses, semesters, and URM status. Table 1 shows the estimates from model 1 for the autoregressive effect (AR[1]), the standard deviation (SD) for the assignment and person random effects, and the fixed effects (b) for day of semester and URM status, for each of the three momentary experience ESM variables. When examining the variation in momentary experiences, we found comparatively little variation in students' reported momentary experiences across assignments relative to variation across students. The *SD*s for these experiences across assignments varied, on average, from .03 (situational interest) to .07 (frustration) with self-efficacy at .05. However, on average, there were substantial amounts of variation for these experiences at the student level with an *SD* of 0.75 for confidence (momentary self-efficacy), 0.76 for frustration, and 0.82 for situational interest.

The fixed effects estimates for the differences in the three momentary experiences between URM students and their peers showed no significant differences. Although these differences were not statistically significant by our ROPE criterion, across both semesters, URM students were more frustrated (.05 *SD* units higher) and expressed lower self-efficacy (0.11 *SD* units lower), yet were more interested (0.12 *SD* units higher) than their racially represented peers. Additionally, the correlations between the three ESM variables are shown in Table 2, for the two random effects variables. Across assignments, frustration and self-efficacy were not related (r = -0.05), but frustration and situation interest were moderately negatively correlated (r = -0.33), and self-efficacy and interest had a small, weak association (r = 0.17). However, correlations of these experiences at the student level showed moderate to strong correlations in anticipated directions (frustration and self-efficacy, r = -0.63; frustration and situational interest, r = -0.31; self-efficacy and situational interest, r =0.56). To better understand the students' experiences, in particular, the differences between these two groups of students, pre-semester covariates were introduced in model 2.

4.2 Model 2 pre-semester interest and self-efficacy covariates on momentary experiences.

To answer RQ2 (*How do baseline levels of sense of belonging, interest value, and self-efficacy impact the momentary experiences of underrepresented students' when compared to their represented peers enrolled in CS1?*), we used three covariates from the beginning of semester survey measuring students' baseline levels of interest in CS, self-efficacy, and sense of belonging. Model 2 includes those as covariates, predicting students' momentary experiences, but it also includes an interaction to examine whether the relationship between students' baseline covariates and their momentary experiences differed between URM students and their peers. These results are shown in Table 3. Model 2 shows a significant relationship between students' baseline self-efficacy and their momentary frustration and self-efficacy, and a significant relationship between baseline interest in CS and their momentary interest. Results in Table 3 show that for every 1 unit increase in students' baseline self-efficacy, their reported levels of frustration decreased by 0.32 *SD* units, and their reported levels of momentary self-efficacy increased by 0.34 *SD* units. Similarly, for every 1-unit increase in students' baseline interest value, their reported levels of momentary interest increased by 0.69 *SD* units. The ROPE inferences are shown in Figure 1. These results confirm the significance of these three effects relative to the 89% HDI.

Model 2 also included interaction effects between URM students and their peers for their baseline levels of self-efficacy, interest, and sense of belonging on their momentary affective experiences (Figures 2 - 4), although these estimates were not significant with respect to the ROPE procedure, as seen in Figure 1. Specifically,

Table 1: Model 1: Autoregressive, Random, and Fixed Effects

	Confident		Frustrated		Interested	
	Value	CrI (95%)	Value	CrI (95%)	Value	CrI (95%)
AR[1]	0.29	0.22 – 0.36	0.28	0.21 – 0.35	0.4	0.33 – 0.47
sd – Assignment Intercept	0.05	0 – 0.1	0.07	0.01 – 0.15	0.03	0 – 0.07
sd – Person Intercept	0.75	0.65 – 0.88	0.76	0.65 – 0.9	0.84	0.71 – 0.98
b – Intercept	-0.22	-0.59 – 0.15	0.09	-0.3 – 0.45	0.22	-0.17 – 0.62
b – Day	-0.01	-0.01 – -0.01	0.01	0.01 – 0.01	-0.01	-0.01 – -0.01
b – URM	-0.11	-0.52 – 0.31	0.05	-0.38 – 0.49	0.12	-0.31 – 0.56

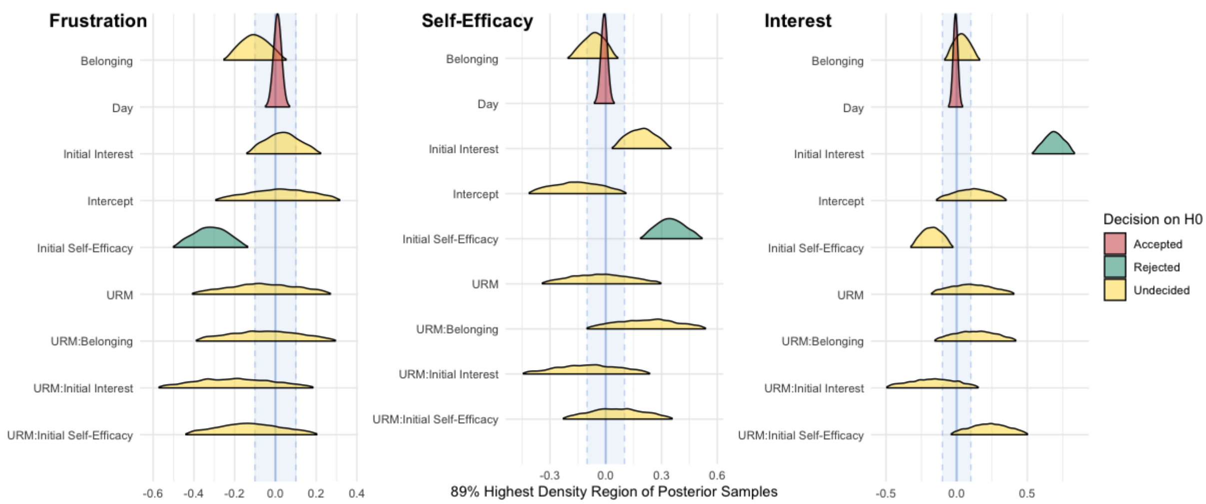

Figure 1: ROPE inferences: Covariates and Interaction Effect Estimates for 3 ESM outcomes (Green = estimate is not practically equivalent to 0, Red = estimate is practically equivalent to 0, Yellow = undecided given the uncertainty in the estimate)

Table 2: Model 1: Correlations Across ESM Experiences by Random Effect Grouping Factors

Parameter	Value	CI (95%)
cor: Assignment – frustrated ~ confident	-0.05	-0.85 – 0.83
cor: Assignment – frustrated ~ interested	-0.33	-0.94 – 0.72
cor: Assignment – confident ~ interested	0.17	-0.79 – 0.91
cor: Person – frustrated ~ confident	-0.63	-0.76 – -0.48
cor: Person – confident ~ interested	0.56	0.39 – 0.7
cor: Person – frustrated ~ interested	-0.31	-0.51 – -0.1

for every 1 unit increase in baseline interest, URMs students had lower change in frustration (-0.20 *SD* units), lower change in momentary interest (-0.16 *SD* units), and lower change in self-efficacy (-0.09 *SD* units). Similarly, for every 1 unit increase in baseline self-efficacy, URM students had lower change in frustration (-0.13 *SD* units), higher change in self-efficacy (0.06 *SD*), and higher change in momentary interest (0.23 *SD*), compared to their represented peers. Students' sense of belonging similarly did not differ significantly between URM students and their peers, with for every 1 unit increase in baseline levels of sense of belonging moderated URM students' had lower change in frustration levels (-0.05 *SD* units), higher change in self-efficacy (0.23 *SD* units) and higher change interest (0.13 *SD* units) when compared to their peers. Again, however, none of these interactions reached significance on the ROPE procedure.

4.3 Affective Experiences and Course Outcomes

Having examined students' momentary experiences and initial self-efficacy, interest in CS, and belongingness as predictors, we sought to understand how these experiences and covariates were related to students' final grades in their CS class. Model 3 built upon model 2, keeping the same models for the 3 ESM variables, but adding another model for final grade, to examine how these momentary experiences were related to end-of-course outcomes. Table 4 presents the results for model 3 and Figure 5 shows HDI and ROPE results. Model 3, like Model 2, also included interaction terms to determine whether the relationships between variables differed between URM students and their peers, and these are shown in Figure 6. These results show two significant and one non-significant interaction between URM students and baseline affective experiences on end-of-course grades. Both baseline interest and self-efficacy significantly

Table 3: Model 2 Summary: Fixed effects estimates for all covariates and interactions for the 3 ESM outcomes

	Confident		Frustrated		Interested	
Predictors	Estimates	CI (95%)	Estimates	CI (95%)	Estimates	CI (95%)
Intercept	-0.16	-0.45 – 0.15	0.03	-0.32 – 0.39	0.11	-0.18 – 0.39
Course CS1B	-0.16	-0.64 – 0.33	0.01	-0.53 – 0.55	-0.44	-0.89 – 0.02
Course CS1C	0.30	-0.21 – 0.81	-0.15	-0.70 – 0.43	0.09	-0.37 – 0.58
Day	-0.01	-0.01 – -0.01	0.01	0.01 – 0.01	-0.01	-0.01 – -0.01
Fall Semester	0.66	0.34 – 0.98	-0.57	-0.94 – -0.19	0.28	-0.03 – 0.58
URM	-0.03	-0.41 – 0.35	-0.06	-0.48 – 0.33	0.09	-0.26 – 0.44
Initial Self-efficacy	0.34	0.16 – 0.52	-0.32	-0.51 – -0.11	-0.18	-0.34 – -0.02
Initial Interest	0.19	0.02 – 0.36	0.03	-0.16 – 0.23	0.69	0.53 – 0.85
Initial Belong	-0.06	-0.20 – 0.07	-0.10	-0.26 – 0.06	0.04	-0.09 – 0.16
URM:Self-eff	0.06	-0.29 – 0.41	-0.13	-0.51 – 0.26	0.23	-0.09 – 0.54
URM:Interest	-0.09	-0.51 – 0.31	-0.20	-0.64 – 0.27	-0.16	-0.55 – 0.23
URM:Belong	0.23	-0.15 – 0.61	-0.05	-0.46 – 0.37	0.13	-0.23 – 0.48

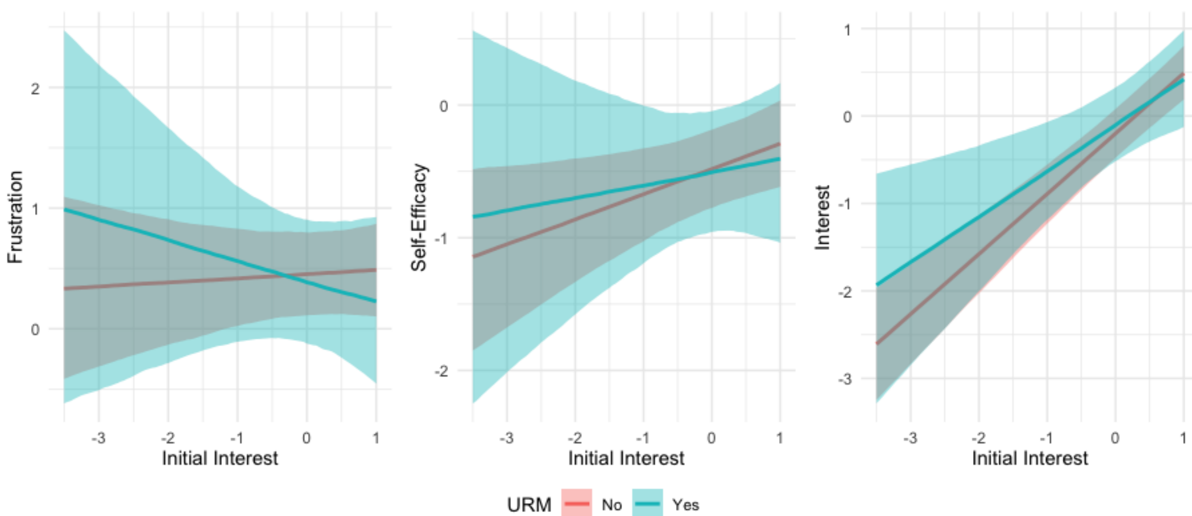

Figure 2: Interaction effects: Initial Interest on 3 ESM Outcomes, URM vs non-URM

differed from that of non-URM students in their relationship with URM students' end-of-course grade. URM students who reported higher levels of baseline self-efficacy had a higher end-of-course grade (by 0.56 *SD*) compared to their represented peers. However, in an unexpected direction, URM students' who reported higher levels of baseline interest in CS and sense of belonging had lower end-of-course grades on average (by 0.66 *SD*) compared to their represented peers. In other words, higher levels of baseline self-efficacy were significantly associated with higher final grades, and higher levels of baseline interest value for CS were significantly associated with lower final grades for URM students.

We also inspected the residual correlation values for the 3 ESM momentary measures and the final grade outcome. These correlations are shown in table 5. These values show the correlation between these variables at the level of individual survey responses, accounting for the autocorrelation, the random effects grouping factors, and the other covariates that we included in the model. These estimates all exclude 0 from their 95% credible intervals, indicating that there is still significant covariance between these measures even after accounting for everything we included in our models. The three ESM variables are correlated in the expected ways: frustration negatively related to confidence and interest, while confidence and interest positively related to one another. The correlations with final grade again follow intuitive lines, with momentary frustration negatively correlated with final grade, whereas momentary confidence and interest positively correlated. These results show that these momentary responses are significantly related to outcomes.

4.4 Conclusions

To summarize the key takeaways of our analyses by research question:

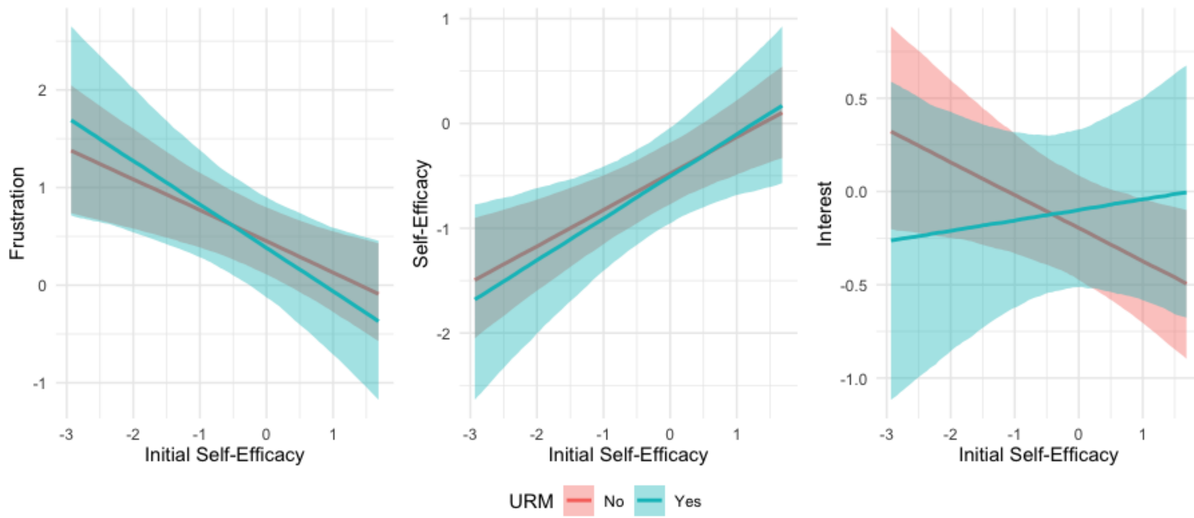

Figure 3: Interaction effects: Self-efficacy on 3 ESM Outcomes, URM vs non-URM

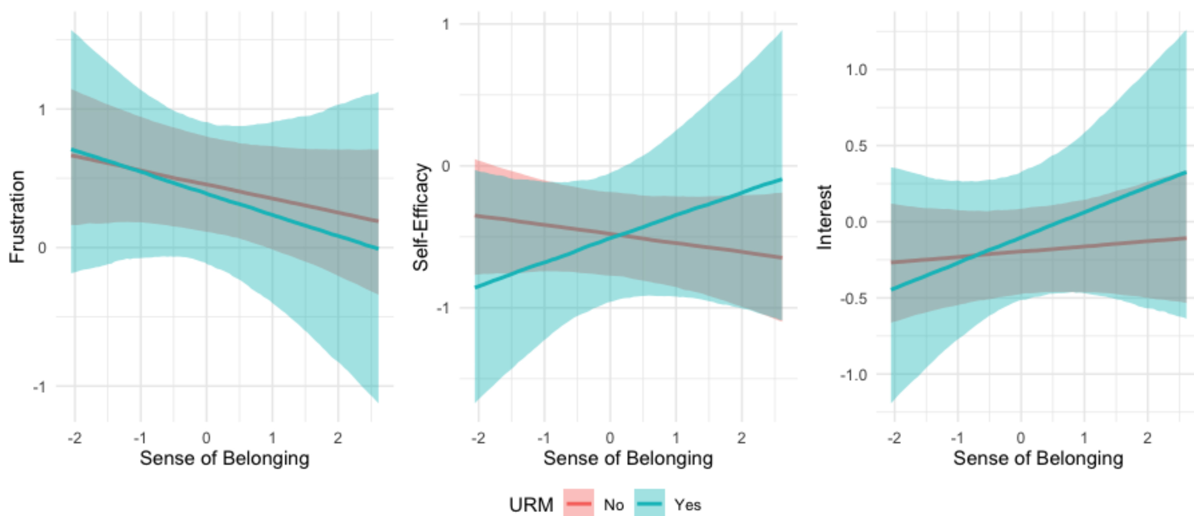

Figure 4: Interaction effects: Sense of Belonging on 3 ESM Outcomes, URM vs non-URM

RQ1: How do racially underrepresented students' momentary experiences differ compared to their represented peers enrolled in CS1?

- On average, there was substantially more variation in momentary experiences across students rather than across their CS1 assignments.
- On average, students with higher levels of baseline self-efficacy and interest had lower levels of frustration across the semester.
- On average, students with higher levels of baseline self-efficacy had higher levels of momentary interest across the semester.
- While not statistically significant differences, on average, URM students were more frustrated, more interested, and reported lower levels of self-efficacy compared to their racially represented peers.

RQ2: How do baseline levels of sense of belonging, interest value, and self-efficacy impact the momentary experiences of underrepresented students' when compared to their represented peers enrolled in CS1?

- On average, students with higher levels of baseline self-efficacy reported significantly lower levels of frustration and higher levels of self-efficacy across the semester.

Table 4: Model 3 Summary: Fixed effects estimates for all covariates and interactions for the 4 outcomes (3 ESM variables + final course grade)

	confident		**frustrated**		**interested**		**grade**	
Predictors	Estimates	CI (95%)	Estimates	CI (95%)	Estimates	CI (95%)	Estimates	CI (95%)
Intercept	-0.16	-0.53 – 0.19	-0.05	-0.44 – 0.35	0.25	-0.11 – 0.60	-0.11	-0.25 – 0.04
Day	-0.01	-0.01 – -0.01	0.01	0.01 – 0.01	-0.01	-0.01 – -0.01	NA	NA
Course CS1B	-0.28	-0.93 – 0.39	0.14	-0.55 – 0.83	-0.40	-1.04 – 0.25	0.13	-0.17 – 0.43
Fall Semester	0.66	0.28 – 1.03	-0.49	-0.91 – -0.06	0.13	-0.24 – 0.52	0.12	-0.03 – 0.28
URM	0.08	-0.37 – 0.53	-0.33	-0.81 – 0.15	0.15	-0.30 – 0.59	0.12	-0.07 – 0.33
Initial Self-efficacy	0.35	0.17 – 0.54	-0.33	-0.51 – -0.13	-0.12	-0.31 – 0.07	0.19	0.12 – 0.27
Initial Interest	0.21	0.02 – 0.41	0.02	-0.20 – 0.24	0.60	0.40 – 0.80	-0.03	-0.11 – 0.05
Initial Belong	-0.04	-0.18 – 0.11	-0.09	-0.25 – 0.06	-0.01	-0.15 – 0.14	0.00	-0.06 – 0.07
URM:Self-eff	0.02	-0.34 – 0.37	-0.23	-0.61 – 0.15	0.24	-0.12 – 0.60	0.56	0.38 – 0.73
URM:Interest	0.02	-0.47 – 0.50	-0.17	-0.68 – 0.35	-0.07	-0.55 – 0.43	-0.66	-0.90 – -0.43
URM:Belong	0.27	-0.14 – 0.65	-0.18	-0.60 – 0.23	0.17	-0.19 – 0.54	-0.18	-0.39 – 0.02

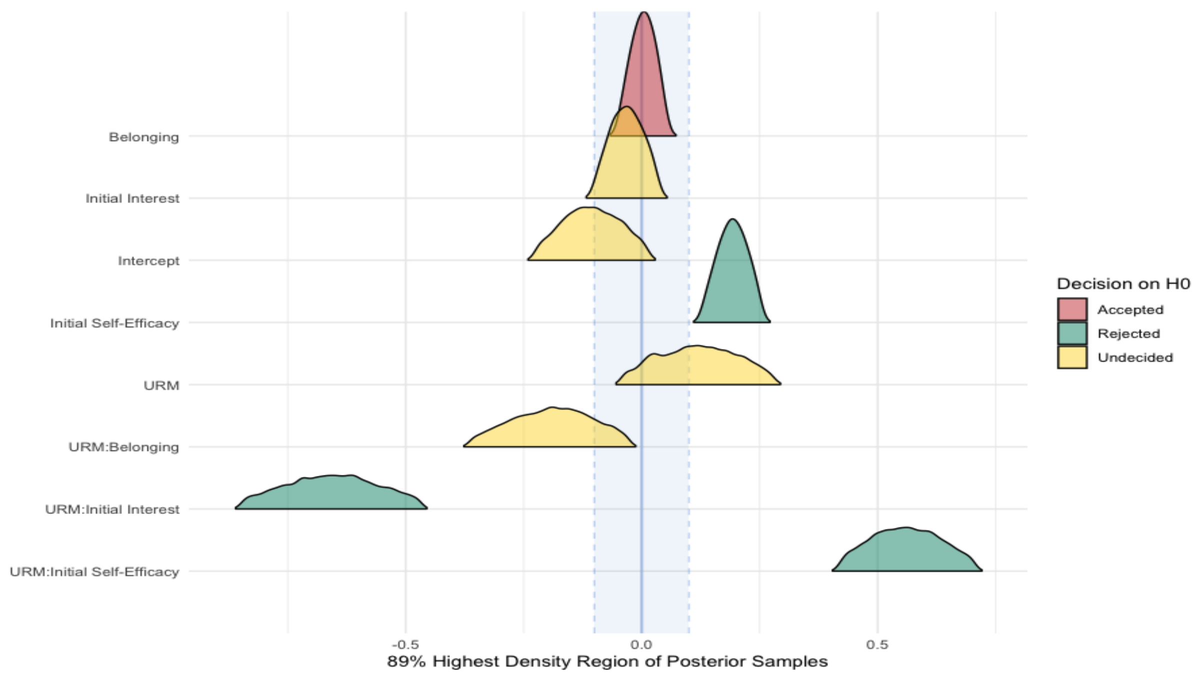

Figure 5: ROPE inferences: Covariates and Interaction Effect Estimates for Final Grade (Green = estimate is not practically equivalent to 0, Red = estimate is practically equivalent to 0, Yellow = undecided given the uncertainty in the estimate)

- On average, students with higher levels of baseline interest reported significantly higher levels of momentary interest in CS during the semester.
- The relationship between baseline covariates (self-efficacy, interest in CS, and sense of belonging) and the ESM experiences were not significantly different for URM students.
- URM students' levels of baseline self-efficacy were significantly more associated with higher final grades than their peers
- For URM students higher levels of baseline interest in CS were significantly associated with lower final grades, which differed from their peers

RQ3: What are the impact of these momentary experiences on course grade outcomes?

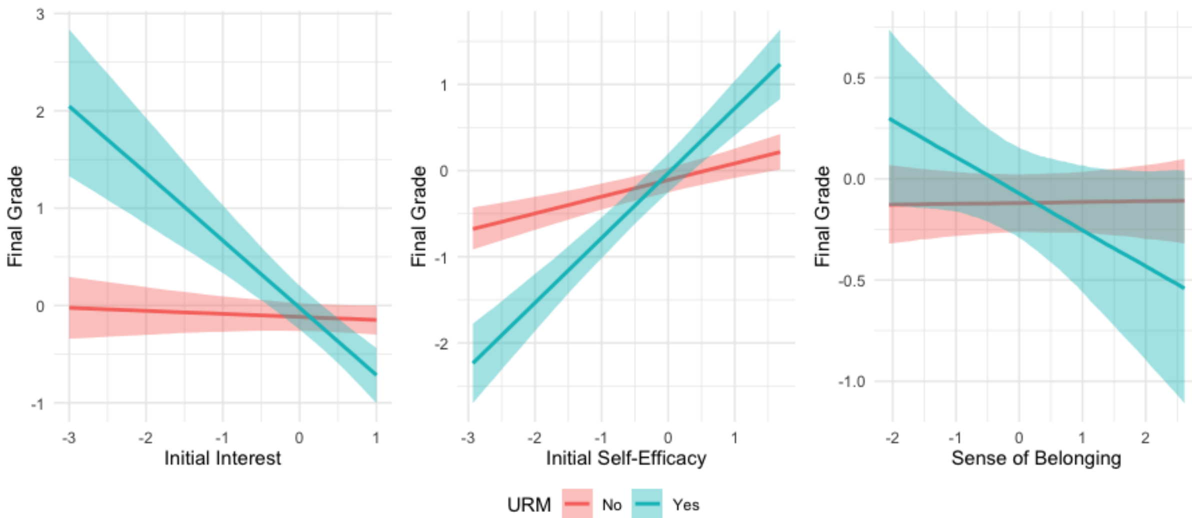

Figure 6: Interaction effects: 3 Baseline Covariates on Final Course Grade, URM vs non-URM

Table 5: Model 3 Residual Correlations

	Estimate	Est.Error	l-95% CI	u-95% CI
cor: Frustration - Confident	-0.33	0.04	-0.4	-0.25
cor: Frustration - Interest	-0.14	0.04	-0.22	-0.06
cor: Confident - Interest	0.39	0.03	0.33	0.45
cor: Frustration - Final Grade	-0.28	0.08	-0.42	-0.12
cor: Confident - Final Grade	0.22	0.07	0.07	0.35
cor: Interest - Final Grade	0.17	0.07	0.03	0.31

5 DISCUSSION

This study contributes to our understanding of how students' motivational and affective experiences in CS1 courses evolve moment to moment, with a particular focus on URM students and whether their experiences differ. The findings for RQ1 are well-informed by prior research and further support and contribute to existing literature, such as the high degree of variation of affective experiences (i.e., self-efficacy and frustration) across CS students. These experiences vary across students due to a variety of internal and external factors, such as CS identity and support systems. Additionally, within the STEM literature, students with lower levels of negative affect, such as frustration, experienced higher levels of positive affect, such as confidence in their ability. Lastly, while not significant, findings of low levels of self-efficacy and high levels of frustration align with research for URM students in STEM and CS fields. Future research should further explore the relationship between self-efficacy and frustration to better understand the significance of the relationship between these constructs.

The findings for RQ2 also further support the previous literature on URM students' experiences in the STEM and CS field with regards to the strong and important relationship between sense of belonging, self-efficacy, and frustration. However, future research should further investigate the relationship between baseline interest and momentary frustration and self-efficacy amongst URM CS students as our results suggest that initial interest does not moderate URM students' momentary self-efficacy. Given these findings, practitioners or educators of CS students can help underrepresented students build and sustain confidence in their abilities and promote an environment of belongingness that may decrease the negative emotions or psychosocial challenges faced in the STEM field.

With regard to course outcomes, URM students' baseline self-efficacy was significantly, positively associated with final grades. As detailed in the literature review, this positive relationship between self-efficacy and grades is well established and we expected to see it, although we did not necessarily anticipate that self-efficacy would have a significantly stronger relationship with grades for URM students. If this finding is robust, this has the implication that future pedagogy and research efforts relating to the experiences of URM students in CS1 should pay particular attention to self-efficacy. The other significant finding regarding student grades, was the unexpected negative association between initial interest in CS and grades for URM students. We did not expect to see this negative relationship or the significant difference between URM students and their peers with respect to this relationship. If this finding is robust, it has potentially large implications for how future research and outreach efforts should consider the values of URM students. One possibility is that URM students use different sources of value to motivate their academic success, which would connect to different components of the task value framework [21]. There is prior research supporting this notion that underrepresented students access different sources of value to persist in their educational pursuits [104]. This finding calls for additional research to better understand how URM students value CS and how their values are related to outcomes.

Overall the results discussed in this paper speak to the importance of positive affective experiences in CS1. We observed the strong (negative) relationship between, for example, experiences of confidence (self-efficacy) and experiences of frustration, as well as the connection between experiences of self-efficacy with final grades, particularly for URM students. A pedagogical and research implication of these findings is that we need to better bolster students self-efficacy in CS1. Computing education research has made some suggestions for how we might accomplish this [49, 57], which CS instructors might well implement currently, but further research is needed to bear out a set of best practices regarding self-efficacy, particularly for efforts to broaden participation in computing. Additionally, while not significant, higher levels of sense of belonging was positively associated with increased momentary self-efficacy and decreased frustration for URM students. Despite the mixed findings in the sense of belonging literature for URM CS students, CS instructors should strive to foster and sustain a physical and social classroom environment where URM students feel accepted, supported, and included. However, it is clear that additional research is needed to study the impact sense of belonging has for URM students within the CS field.

In addition to the suggestions for future research noted above, we feel it is important to address the question of the utility of ESM to examine momentary experiences of CS students given the majority of the variability was across students as opposed to assignments/timepoints. While we anticipated finding more variation across assignments/timepoints, with the premise that patterns in variability based on the course context would be informative for future research and practitioners, we still believe that there is potential for further research on CS students' momentary experiences using ESM. On a pragmatic level, there are many different ways that ESM could be used to investigate students' experiences, and this study is a rare example of this method being used in computing education research. Overall, there are many different constructs that could be investigated that might be more context sensitive, and which can be measured in many ways; there are many different ways to conduct an ESM data collection, and there are many different ways to analyze ESM data. On a more philosophical level, we also conceptualize students' experiences as an outcome in themselves that is worthy of study. If we saw the variability in experiences track closer to differences in the course context from moment to moment, that could certainly lend itself to teaching applications and further research regarding the way that course content is structured. Nevertheless, understanding the variability in students course experiences and how these connect to beginning of course covariates for different students has a different but related application in how to be responsive to the different needs of different students. This is of particular importance in the realm of closing participation gaps for traditionally underrepresented groups of students.

6 LIMITATIONS

Recent discussions within quantitative critical (QuantCrit) research have called researchers to recognize that quantitative studies may be unintentionally biased by researchers' positionality within the context of a study and to exercise caution when comparing majority and minority groups [29]. Given this recent call, the first limitation in this study is not being able to study URM students' affective experiences without a reference point. URM students are grossly underrepresented and undeniably face challenges in persisting in the CS field. In order to bring further awareness and contribute to the understanding of the factors that can promote their success in the CS field, we first need to understand the factors that promote persistence relative to their peers in order to ensure practitioners and institutions can support these needs. The second limitation of the study is the small sample size of students who identified as URM. We encourage future studies to examine the momentary experiences of URM students with a larger sample size to ensure adequate statistical power. Another specific target for future research is to explore different ways of modeling ESM data, looking at more complex patterns in change over time, and further investigating patterns of non-response in ESM data and how they may be related to covariates and outcomes of interest.

REFERENCES

[1] Mi Young Ahn and Howard H. Davis. 2020. Students' sense of belonging and their socio-economic status in higher education: a quantitative approach. *Teaching in Higher Education* 0, 0 (2020), 1–14. https://doi.org/10.1080/13562517.2020.1778664
[2] Mary Ainley. 2007. Being and feeling interested: Transient state, mood, and disposition. In *Emotion in Education*, Paul A Schultz and Reinhard Pekrun (Eds.). 141–157.
[3] Alexander W Astin. 1999. Student involvement: A developmental theory for higher education. (1999).
[4] Albert Bandura. 1977. Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological review* 84, 2 (mar 1977), 191–215. http://www.ncbi.nlm.nih.gov/pubmed/847061
[5] Albert Bandura. 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Prentice-Hall, Englewood Cliffs, NJ.
[6] Patrick N. Beymer, Joshua M. Rosenberg, Jennifer A. Schmidt, and Neil J. Naftzger. 2018. Examining Relationships among Choice, Affect, and Engagement in Summer STEM Programs. *Journal of Youth and Adolescence* 47, 6 (2018), 1178–1191. https://doi.org/10.1007/s10964-018-0814-9
[7] Jyoti Bhardwaj. 2017. In search of self-efficacy: development of a new instrument for first year Computer Science students. *Computer Science Education* (2017), 1–21. https://doi.org/10.1080/08993408.2017.1355522
[8] Nigel Bosch and Sidney D'Mello. 2013. Sequential patterns of affective states of novice programmers. *CEUR Workshop Proceedings* 1009, Aiedcs (2013), 1–10.
[9] Nigel Bosch and Sidney D'Mello. 2017. The Affective Experience of Novice Computer Programmers. *International Journal of Artificial Intelligence in Education* 27, 1 (2017), 181–206. https://doi.org/10.1007/s40593-015-0069-5
[10] Paul-Christian Bürkner. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software* 80, 1 (2017), 1–28.
[11] Jennifer Campbell, Diane Horton, and Michelle Craig. 2016. Factor for Success in Online CS1. *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education* (2016), 320–325. https://doi.org/10.1145/2899415.2899457
[12] Sapna Cheryan, Allison Master, and Andrew N. Meltzoff. 2015. Cultural stereotypes as gatekeepers: Increasing girls' interest in computer science and engineering by diversifying stereotypes. *Frontiers in Psychology* 6, FEB (2015), 1–8. https://doi.org/10.3389/fpsyg.2015.00049
[13] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M Steele. 2009. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology* 97, 6 (2009), 1045–1060. https://doi.org/10.1037/a0016239
[14] Anne Marie M. Conley. 2012. Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology* 104, 1 (2012), 32–47. https://doi.org/10.1037/a0026042

[15] Marcus Credé and L Alison Phillips. 2011. A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences* 21 (2011), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002
[16] Mihaly Csikszentmihalyi and Reed Larson. 2014. Validity and Reliability of the Experience-Sampling Method. In *Flow and the Foundations of Positive Psychology*. 1–298. https://doi.org/10.1007/978-94-017-9088-8
[17] Teresa Dahlberg, Tiffany Barnes, Kim Buch, and Audrey Rorrer. 2011. The STARS alliance: Viable strategies for broadening participation in computing. *ACM Transactions on Computing Education (TOCE)* 11, 3 (2011), 1–25.
[18] Holger Danielsiek, Laura Toma, and Jan Vahrenhold. 2017. An Instrument to Assess Self-Efficacy in Introductory Algorithms Courses. *Proceedings of the 2017 ACM Conference on International Computing Education Research* (2017), 217–225.
[19] Ivana Di Leo, Krista R. Muis, Cara A. Singh, and Cynthia Psaradellis. 2019. Curiosity… Confusion? Frustration! The role and sequencing of emotions during mathematics problem solving. *Contemporary Educational Psychology* 58, March (2019), 121–137. https://doi.org/10.1016/j.cedpsych.2019.03.001
[20] Remy Dou, Karina Bhutta, Monique Ross, Laird Kramer, and Vishodana Thamotharan. 2020. The Effects of Computer Science Stereotypes and Interest on Middle School Boys' Career Intentions. *ACM Transactions on Computing Education* 20, 3 (2020). https://doi.org/10.1145/3394964
[21] Jacquelynne S Eccles and Allan Wigfield. 1995. In the mind of the achiever: The structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin* 21, 3 (1995), 215–225.
[22] Anna Eckerdal, Robert Mccartney, Jan Erik Moström, Kate Sanders, Lynda Thomas, and Carol Zander. 2007. From Limen to Lumen: Computing students in liminal spaces. *Proceedings of the third international workshop on Computing education research* (2007), 123–132.
[23] Albrecht Ehlert and Carsten Schulte. 2009. Empirical comparison of objects-first and objects-later. *ICER'09 - Proceedings of the 2009 ACM Workshop on International Computing Education Research* (2009), 15–26. https://doi.org/10.1145/1584322.1584326
[24] Allan Fisher and Jane Margolis. 2002. Unlocking the clubhouse: the Carnegie Mellon experience. *ACM SIGCSE Bulletin* 34, 2 (2002), 79–83.
[25] Lawrence M Fisher. 2016. Booming enrollments. *Commun. ACM* 59, 7 (2016), 17–18.
[26] Denae Ford and Chris Parnin. 2015. Exploring causes of frustration for software developers. In *2015 IEEE/ACM 8th International Workshop on Cooperative and Human Aspects of Software Engineering*. IEEE, 115–116.
[27] Jonah Gabry and Tristan Mahr. 2022. bayesplot: Plotting for Bayesian Models. https://mc-stan.org/bayesplot/ R package version 1.9.0.
[28] Andrew Gelman. 2006. Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. *Technometrics* 48, 3 (2006), 432–435. https://doi.org/10.1198/004017005000000661
[29] David Gillborn, Paul Warmington, and Sean Demack. 2018. QuantCrit: education, policy,'Big Data'and principles for a critical race theory of statistics. *Race Ethnicity and Education* 21, 2 (2018), 158–179.
[30] Catherine Good, Aneeta Rattan, and Carol S Dweck. 2012. Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of personality and social psychology* 102, 4 (2012), 700.
[31] Carol Goodenow. 1993. Classroom belonging among early adolescent students: Relationships to motivation and achievement. *The Journal of early adolescence* 13, 1 (1993), 21–43.
[32] Jamie Gorson and Eleanor O'Rourke. 2020. Why do CS1 Students Think They're Bad at Programming?: Investigating Self-efficacy and Self-assessments at Three Universities. *ICER 2020 - Proceedings of the 2020 ACM Conference on International Computing Education Research* (2020), 170–181. https://doi.org/10.1145/3372782.3406273
[33] Jamie Gorson and Eleanor O'Rourke. 2021. CS1 Student Assessments of Themselves Relative to Others: The Role of Self-Critical Bias, Gender, and Sense of Belonging. In *International Society of the Learning Sciences*.
[34] Mark Guzdial. 2013. Exploring hypotheses about media comptation. In *ICER 2013 - Proceedings of the 2013 ACM Conference on International Computing Education Research*. 19–26. https://doi.org/10.1145/2493394.2493397
[35] Brian Hanks, Sue Fitzgerald, Renée McCauley, Laurie Murphy, and Carol Zander. 2011. Pair programming in education: a literature review. *Computer Science Education* 21, 2 (2011), 135–173. https://doi.org/10.1080/08993408.2011.579808
[36] Markku S Hannula. 2015. Emotions in Problem Solving. In *Selected Regular Lectures from the 12th International Congress on Mathematical Education*, Sun Joo Cho (Ed.). Springer International Publishing, 269–288. https://doi.org/10.1007/978-3-319-17187-6
[37] Joel M. Hektner, Jennifer A. Schmidt, and Mihaly Csikszentmihalyi. 2007. *Experience Sampling Method: Measuring the quality of everyday life.* Sage.
[38] Suzanne Hidi and Ann Renninger. 2006. The Four-Phase Model of Interest Development. *Educational Psychologist* 41, 2 (2006), 111–127. https://doi.org/10.1207/s15326985ep4102 arXiv:0710.4428v1
[39] Marybeth Hoffman, Jayne Richmond, Jennifer Morrow, and Kandice Salomone. 2002. Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention: Research, Theory & Practice* 4, 3 (2002), 227–256.
[40] Toni Honicke and Jaclyn Broadbent. 2016. The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review* 17 (2016), 63–84. https://doi.org/10.1016/j.edurev.2015.11.002
[41] Mica Hutchison, Deborah K. Follman, Melissa Sumpter, and George M. Bodner. 2006. Factors Influencing the Self-Efficacy Beliefs of First-Year Engineering Students. *Journal of Engineering Education* 95, 1 (2006), 39. https://doi.org/10.1002/j.2168-9830.2006.tb00876.x
[42] Tiffany A Ito and Erin McPherson. 2018. Factors influencing high school students' interest in pSTEM. *Frontiers in psychology* (2018), 1535.
[43] Dawn R Johnson. 2012. Campus racial climate perceptions and overall sense of belonging among racially diverse women in STEM majors. *Journal of College Student Development* 53, 2 (2012), 336–346.
[44] Martin H. Jones, Shannon R. Audley-Piotrowski, and Sarah M. Kiefer. 2012. Relationships among adolescents' perceptions of friends' behaviors, academic self-concept, and math performance. *Journal of Educational Psychology* 104, 1 (2012), 19–31. https://doi.org/10.1037/a0025596
[45] Yasmin B Kafai, Kristin Searle, Cristóbal Martinez, and Bryan Brayboy. 2014. Ethnocomputing with Electronic Textiles : Culturally Responsive Open Design to Broaden Participation in Computing in American Indian Youth and Communities. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE '14)*.
[46] Sandra Katz, David Allbritton, John Aronis, Christine Wilson, and Mary Lou Soffa. 2006. Gender, Achievement, and Persistence in an Undergraduate Computer Science Program. *Data Base for Advances in Information Systems* 37, 4 (2006), 42–57. https://doi.org/10.1145/1185335.1185344
[47] Jen Katz-Buonincontro and Rajashi Ghosh. 2014. Using workplace experiences for learning about affect and creative problem solving: Piloting a four-stage model for management education. *International Journal of Management Education* 12, 2 (2014), 127–141. https://doi.org/10.1016/j.ijme.2014.03.003
[48] Päivi Kinnunen and Beth Simon. 2010. Experiencing Programming Assignments in CS1: The Emotional Toll. *Proceedings of the Sixth international workshop on Computing education research (ICER '10)* (2010), 77–85. https://doi.org/10.1145/1839594.1839609
[49] Päivi Kinnunen and Beth Simon. 2011. CS Majors' Self-Efficacy Perceptions in CS1: Results in Light of Social Cognitive Theory. *Proceedings of the seventh international workshop on Computing education research - ICER '11* (2011), 19–26. https://doi.org/10.1145/2016911.2016917
[50] 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*. 29–41.
[51] Stephan Krusche, Barbara Reichart, Paul Tolstoi, and Bernd Bruegge. 2016. Experiences from an Experiential Learning Course on Games Development. *Proceedings of the 47th ACM Technical Symposium on Computer Science Education (SIGCSE '16)* (2016), 582–587. https://doi.org/10.1145/2839509.2844599
[52] Feng Yang Kuo, Wen Hsiung Wu, and Cathy Lin. 2013. An investigation of self-regulatory mechanisms in learning to program visual basic. *Journal of Educational Computing Research* 49, 2 (2013), 225–247. https://doi.org/10.2190/EC.49.2.f
[53] Li Lin, Taehun Lee, and Lori Anderson Snyder. 2018. Math self-efficacy and STEM intentions: A person-centered approach. *Frontiers in psychology* (2018), 2033.
[54] Alex Lishinski. 2017. *Cognitive, Affective, and Dispositional Components of Learning Programming*. Ph. D. Dissertation. https://www.alexlishinski.com/media/Lishinski-dissertation.pdf
[55] Alex Lishinski and Joshua Rosenberg. 2021. *All the Pieces Matter: The Relationship of Momentary Self-efficacy and Affective Experiences with CS1 Achievement and Interest in Computing*. Vol. 1. Association for Computing Machinery. 252–265 pages. https://doi.org/10.1145/3446871.3469740
[56] Alex Lishinski and Aman Yadav. 2018. Motivation, Attitudes, and Dispositions. In *Handbook of Computing Education Research*, Sally Fincher and Anthony Robins (Eds.). Cambridge University Press.
[57] Alex Lishinski and Aman Yadav. 2021. Self-evaluation Interventions: Impact on Self-efficacy and Performance in Introductory Programming. *ACM Transactions on Computing Education* (2021).
[58] Alex Lishinski, Aman Yadav, and Richard Enbody. 2017. Students' Emotional Reactions to Programming Projects in Introduction to Programming: Measurement Approach and Influence on Learning Outcomes. *Proceedings of the 2017 ACM Conference on International Computing Education Research* (2017), 30–38.
[59] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. 2016. Learning to Program: Gender Differences and Interactive Effects of Students' Motivation, Goals and Self-Efficacy on Performance. *Proceedings of the 12th Annual International ACM Conference on International Computing Education Research (ICER '16)* (2016).
[60] Ashley Lytle, Alexander John De Rosa, and Frank T Fisher. 2021. A Review of Psychosocial Factors Associated with Undergraduate Engagement and Retention in STEM. In *2021 Fall ASEE Middle Atlantic Section Meeting*.

[61] Mary Lou Maher, Celine Latulipe, Heather Lipford, and Audrey Rorrer. 2015. Flipped classroom strategies for CS education. *SIGCSE 2015 - Proceedings of the 46th ACM Technical Symposium on Computer Science Education* (2015), 218–223. https://doi.org/10.1145/2676723.2677252

[62] Dominique Makowski, Mattan S. Ben-Shachar, and Daniel Lüdecke. 2019. bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software* 4, 40 (2019), 1541. https://doi.org/10.21105/joss.01541

[63] Rodolfo Mendoza-Denton, Geraldine Downey, Valerie J. Purdie, Angelina Davis, and Janina Pietrzak. 2002. Sensitivity to status-based rejection: Implications for African American students' college experience. *Journal of Personality and Social Psychology* 83, 4 (2002), 896–918. https://doi.org/10.1037/0022-3514.83.4.896

[64] Catherine Mooney, Anna Antoniadi, Ioannis Karvelas, Lána Salmon, and Brett A Becker. 2020. Exploring Sense of Belonging in Computer Science Students. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*. 563–563.

[65] Jennifer Morrow and Margot Ackermann. 2012. Intention to persist and retention of first-year students: The importance of motivation and sense of belonging. *College student journal* 46, 3 (2012), 483–491.

[66] Krista R. Muis, Marianne Chevrier, and Cara A. Singh. 2018. The Role of Epistemic Emotions in Personal Epistemology and Self-Regulated Learning. *Educational Psychologist* 53, 3 (2018), 165–184. https://doi.org/10.1080/00461520.2017.1421465

[67] Karen D Multon, Steven D Brown, and Robert W Lent. 1991. Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology* 38, 1 (1991), 30–38. https://doi.org/10.1037/0022-0167.38.1.30

[68] Diane Nahl. 2005. Affective Load. In *Theories of information behavior.*, Karen E. Fisher, Sanda Erdelez, and Lynne McKechnie (Eds.). Information Today, Medford, N.J., 39–43.

[69] 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. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE* (2020), 685–691. https://doi.org/10.1145/3328778.3366805

[70] Patrick O'Keeffe. 2013. A sense of belonging: Improving student retention. *College Student Journal* 47, 4 (2013), 605–613.

[71] Ernesto Panadero. 2017. A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology* 8, APR (2017), 1–28. https://doi.org/10.3389/fpsyg.2017.00422

[72] Ernesto Panadero, Anders Jonsson, and Juan Botella. 2017. Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review* 22 (2017), 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

[73] Erika A Patall, Rebecca R Steingut, Ariana C Vasquez, Scott S Trimble, Keenan A Pituch, and Jen L Freeman. 2018. Daily autonomy supporting or thwarting and students' motivation and engagement in the high school science classroom. *Journal of Educational Psychology* 110, 2 (2018), 269.

[74] Reinhard Pekrun, Stephanie Lichtenfeld, Herbert W. Marsh, Kou Murayama, and Thomas Goetz. 2017. Achievement Emotions and Academic Performance: Longitudinal Models of Reciprocal Effects. *Child Development* 88, 5 (2017), 1653–1670. https://doi.org/10.1111/cdev.12704

[75] Markeya S. Peteranetz, Shiyuan Wang, Duane F. Shell, Abraham E. Flanigan, and Leen Kiat Soh. 2018. Examining the impact of computational creativity exercises on college computer science students' learning, achievement, self-efficacy, and creativity. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education* 2018-Janua (2018), 155–160. https://doi.org/10.1145/3159450.3159459

[76] Paul R. Pintrich, David A.F. Smith, Teresa Garcia, and Wilbert J. McKeachie. 1993. Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement* 53, 3 (sep 1993), 801–813. https://doi.org/10.1177/0013164493053003024

[77] Johanna Pirker, Maria Riffnaller-Schiefer, and Christian Gütl. 2014. Motivational active learning - Engaging university students in computer science education. *ITICSE 2014 - Proceedings of the 2014 Innovation and Technology in Computer Science Education Conference* (2014), 297–302. https://doi.org/10.1145/2591708.2591750

[78] Katherine Rainey, Melissa Dancy, Roslyn Mickelson, Elizabeth Stearns, and Stephanie Moller. 2018. Race and gender differences in how sense of belonging influences decisions to major in STEM. *International journal of STEM education* 5, 1 (2018), 1–14.

[79] Vennila Ramalingam, Deborah LaBelle, and Susan Wiedenbeck. 2004. Self-efficacy and mental models in learning to program. *Proceedings of the 9th annual SIGCSE conference on innovation and technology in computer science education - ITiCSE '04* (2004), 171–175. https://doi.org/10.1145/1007996.1008042

[80] Aneeta Rattan, Krishna Savani, Dolly Chugh, and Carol S Dweck. 2015. Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science* 10, 6 (2015), 721–726.

[81] E Rivers. 2017. Women, minorities, and persons with disabilities in science and engineering. *National Science Foundation* (2017).

[82] Steven B. Robbins, Kristy Lauver, Huy Le, Daniel Davis, Ronelle Langley, and Aaron Carlstrom. 2004. Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological bulletin* 130, 2 (2004), 261–288. https://doi.org/10.1037/0033-2909.130.2.261

[83] Adrian Salguero, William G Griswold, Christine Alvarado, and Leo Porter. 2021. Understanding Sources of Student Struggle in Early Computer Science Courses. In *Proceedings of the 17th ACM Conference on International Computing Education Research*. 319–333.

[84] Linda J Sax, Jennifer M Blaney, Kathleen J Lehman, Sarah L Rodriguez, Kari L George, and Christina Zavala. 2018. Sense of belonging in computing: The role of introductory courses for women and underrepresented minority students. *Social Sciences* 7, 8 (2018), 122.

[85] Ulrich Schiefele. 2001. The Role of Interest in Motivation and Learning. In *Intelligence and personality: Bridging the gap in theory and measurement*, J.M. Collis and S. Messick (Eds.). Erlbaum, Mahwah, NJ, 163–194.

[86] Jennifer A Schmidt, Patrick N Beymer, Joshua M Rosenberg, Neil N Naftzger, and Lee Shumow. 2020. Experiences, activities, and personal characteristics as predictors of engagement in STEM-focused summer programs. *Journal of Research in Science Teaching* 57, 8 (2020), 1281–1309.

[87] Barbara Schneider, Joseph Krajcik, Jari Lavonen, Katariina Salmela-Aro, Michael Broda, Justina Spicer, Justin Bruner, Julia Moeller, Janna Linnansaari, Kalle Juuti, et al. 2016. Investigating optimal learning moments in US and Finnish science classes. *Journal of Research in Science Teaching* 53, 3 (2016), 400–421.

[88] Dale H. Schunk and Peggy A. Ertmer. 2000. Self-regulation and academic learning: Self-efficacy enhancing interventions. In *Handbook of self-regulation*, Monique Boekaerts, Paul R. Pintrich, and M. Zeidner (Eds.). 631–649.

[89] Christie Scollon, Chu Kim-Prieto, and Ed Diener. 2003. Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies* 4, 1 (2003), 5–34. http://ideas.repec.org/a/spr/jhappi/v4y2003i1p5-34.html

[90] Duane F Shell, Abraham E Flanigan, and Markeya S Peteranetz. 2016. Students ' Initial Course Motivation and Their Achievement and Retention in College CS1 Courses. *Proceedings of the 47th ACM Technical Symposium on Computer Science Education (SIGCSE '16)* (2016), 639–644.

[91] Jessi L Smith, Karyn L Lewis, Lauren Hawthorne, and Sara D Hodges. 2013. When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Personality and Social Psychology Bulletin* 39, 2 (2013), 131–143.

[92] Phil Steinhorst, Andrew Petersen, and Jan Vahrenhold. 2020. Revisiting Self-Efficacy in Introductory Programming. *ICER 2020 - Proceedings of the 2020 ACM Conference on International Computing Education Research* (2020), 158–169. https://doi.org/10.1145/3372782.3406281

[93] Terrell L Strayhorn. 2011. Sense of belonging and African-American student success in STEM: comparative insights between men and women. In *Beyond stock stories and folktales: African Americans' paths to STEM fields.* Emerald Group Publishing Limited.

[94] Moin Syed, Eileen L. Zurbriggen, Martin M. Chemers, Barbara K. Goza, Steve Bearman, Faye J. Crosby, Jerome M. Shaw, Lisa Hunter, and Elizabeth M. Morgan. 2019. The Role of Self-Efficacy and Identity in Mediating the Effects of STEM Support Experiences. *Analyses of Social Issues and Public Policy* 19, 1 (2019), 7–49. https://doi.org/10.1111/asap.12170

[95] Vincent Tinto. 1975. Dropout from higher education: A theoretical synthesis of recent research. *Review of educational research* 45, 1 (1975), 89–125.

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

[97] Vincent Tinto. 1996. Reconstructing the First Year of College. *Planning for higher education* 25, 1 (1996), 1–6.

[98] Jeffrey C. Valentine, David L. Dubois, and Harris Cooper. 2004. The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review. *Educational Psychologist* 39, July 2015 (2004), 111–133. https://doi.org/10.1207/s15326985ep3902

[99] Thomas I Vaughan-Johnston and Jill A Jacobson. 2020. Theory Issues in Measurement Self-efficacy Theory. In *The Wiley Encyclopedia of Personality and Individual Differences: Models and Theories*. Vol. I. 375–379.

[100] 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.

[101] Christopher Watson, Frederick W. B. Li, and Jamie L. Godwin. 2014. No Tests Required: Comparing Traditional and Dynamic Predictors of Programming Success. *Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE '14)* (2014), 469–474.

[102] Susan Wiedenbeck. 2005. Factors affecting the success of non-majors in learning to program. *Proceedings of the 2005 international workshop on Computing education research - ICER '05* (2005), 13–24. https://doi.org/10.1145/1089786.1089788

[103] Chris Wilcox and Albert Lionelle. 2018. Quantifying the benefits of prior programming experience in an introductory computer science course. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science*

Education 2018-Janua (2018), 80–85. https://doi.org/10.1145/3159450.3159480
[104] Tara J Yosso*. 2005. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race ethnicity and education* 8, 1 (2005), 69–91.
[105] Barry J Zimmerman and Anastasia Kitsantas. 1999. Acquiring Writing Revision Skill: Shifting From Process to Outcome Self-Regulatory Goals. *Journal of Educational Psychology* 91, 2 (1999), 241–250.
[106] Stuart Zweben. 2019. Enrollment and retention in US computer science bachelor's programs in 2016-17. *ACM Inroads* 10, 4 (2019), 47–59.
[107] Stuart Zweben and Betsy Bizot. 2019. Taulbee Survey Total Undergrad CS, Enrollment Rises Again, but with Fewer New Majors; Doctoral, Degree Production Recovers From Last Year's Dip. *Computing Research Association* (2019).

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1937700.

A ESM SURVEY

Please indicate your agreement at this moment with the following statements about your experiences in COSC 102/111/505 on a 1-5 scale, with 1 indicating strong disagreement, 3 indicating that you neither agree nor disagree, and 5 indicating strong agreement.

- I feel frustrated.
- I feel confident about being able to do the work going forward.
- I feel interested in computer science.