Journal for STEM Education Research (2020) 3:195–216
https://doi.org/10.1007/s41979-020-00031-1

RESEARCH ARTICLE

Undergraduate Student Concerns in Introductory STEM Courses: What They Are, How They Change, and What Influences Them

Clara L. Meaders[1] · A. Kelly Lane[2] · Anya I. Morozov[2] · J. Kenny Shuman[2] · Emma S. Toth[3] · Marilyne Stains[4] · MacKenzie R. Stetzer[5,6] · Erin Vinson[6] · Brian A. Couch[2] · Michelle K. Smith[1]

Published online: 29 April 2020

Abstract

Introductory STEM courses represent entry points into a major, and student experiences in these courses can affect both their persistence and success in STEM disciplines. Identifying course-based student concerns may help instructors detect negative perceptions, areas of struggle, and potential barriers to success. Using an open-response survey question, we identified 13 common concerns expressed by students in introductory STEM courses. We converted these student-generated concerns into closed-ended items that were administered at the beginning and middle of the semester to students in 22 introductory STEM course sections across three different institutions. Students were asked to reflect on each item on a scale from very concerned to not concerned. A subset of these concerns was used to create a summary score of course-based concern for each student. Overall levels of student concern decreased from the first week to the middle of the semester; however, this pattern varied across different demographic groups. In particular, when controlling for initial concern and course grades, female students held higher levels of concern than their peers. Since student perceptions can impact their experiences, addressing concerns through communication and instructional practices may improve students' overall experiences and facilitate their success.

Keywords Concerns · Introductory STEM courses · Gender · Undergraduates

Clara L. Meaders and A. Kelly Lane contributed equally to this work. Brian A. Couch and Michelle K. Smith contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s41979-020-00031-1) contains supplementary material, which is available to authorized users.

✉ Brian A. Couch
bcouch2@unl.edu

Extended author information available on the last page of the article

Introduction

Undergraduate students' experiences in introductory gateway science, technology, engineering, and mathematics (STEM) courses contribute to student retention and overall persistence in STEM majors and careers (Suresh 2006; Watkins and Mazur 2013). Half of college students who intend to graduate with STEM degrees fail to do so within six years of starting college (Eagan et al. 2014). The majority of students who leave STEM majors do not persist past the first year (Seymour and Hewitt 1997). Increasing student retention in STEM majors has been proposed as a key strategy to producing the overall number of graduates required to meet the growing need for a trained workforce (President's Council of Advisors on Science and Technology 2012). Improving student retention rates is also necessary to increase diversity in STEM because retention rates are lower for students from underrepresented groups (Alexander et al. 2009; Beasley and Fischer 2012; Cataldi et al. 2018; Hughes 2018; Riegle-Crumb et al. 2019).

Students cite course experiences including their perceptions of classroom climate, faculty behaviors, and interactions with faculty as reasons to leave STEM majors (Geisinger and Raman 2013; Seymour and Hewitt 1997; Suresh 2006; Watkins and Mazur 2013). One challenge for students may be that they can experience significant shifts in instructional methods between their high school and college STEM courses (Akiha et al. 2018), and students who are experiencing their first semester in college or are first generation have different predictions about classroom instruction when compared to their peers (Brown et al. 2017; Meaders et al. 2019). These factors may contribute to students having negative experiences in their introductory level courses. Consequently, identifying the course-related issues about which students perceive and express concern could lead to ways to improve introductory course experiences for all students.

To date, the undergraduate education literature has largely focused on describing students' adverse experiences in course environments through the lens of anxiety (e.g., Barrows et al. 2012; Chapell et al. 2005; Foley et al. 2017). Anxiety, which can be described as a negative, prospective emotion, is associated with lower college exam grades and retention in STEM courses (Barrows et al. 2012; Bellinger et al. 2015; Chapell et al. 2005; England et al. 2017; Foley et al. 2017; Pekrun et al. 2007). For example, England et al. (2017) found that 16% of students (total $n = 327$) in three introductory biology sections reported moderately high classroom anxiety, and students with greater anxiety were more likely to self-report lower grades and intent to leave the biology major.

One type of anxiety extensively studied in undergraduate education is test anxiety, or fear of failure during exams, and this research has often included an investigation of gender differences. Studies frequently report that female students have overall higher levels of test anxiety than male students (e.g., Chapell et al., 2005; Harris et al. 2019; Núñez-Peña et al. 2016). With respect to how anxiety relates to performance, there is often an inverse relationship between test anxiety and exam performance (e.g., Harris et al. 2019) and between test anxiety and undergraduate GPA (e.g., Chapell et al., 2005). However, studies show mixed results with respect to how the relationship between anxiety and performance plays out by gender. While female students have overall higher levels of test anxiety, some studies find that high test anxiety similarly affects performance for both male and female students (Chapell et al. 2005; Harris et al.

2019; Seipp 1991) or negatively influences female student performance alone (Salehi et al. 2019). In some cases, test anxiety does not impact performance for either gender (England et al. 2019). These studies highlight the complex interplay between emotional states, student background characteristics, and course performance.

Instructors' actions and course design can potentially contribute to anxiety, even in active learning environments (Eddy et al. 2015; England et al. 2017). An investigation of specific active learning strategies revealed that practices such as call and response, worksheets, clicker questions, and peer discussions caused anxiety for some students (Cooper et al. 2018; England et al. 2017). In addition, a student's demographic background can contribute to how they experience different types of instructional strategies. For example, female students report higher levels of anxiety related to whole-class discussions, and international students report higher levels of anxiety regarding peer discussions (Eddy et al. 2015).

While there are many investigations into undergraduate anxiety, other emotions such as concern are less frequently explored. In this article, we aim to identify the salient course-based concerns that undergraduate students have about their introductory STEM courses. We define "concerns" as sources of apprehension, uncertainty, or difficulty that students perceive to affect their interaction with or success in a course. Concerns can arise early in the semester and can relate to a variety of in-class or out-of-class activities. Furthermore, concerns as we have defined them encompass sources of stress as well as other challenges that do not meet the threshold of anxiety. We focused on concerns rather than anxiety because we wanted to identify a broad list of issues that instructors could act on to improve students' course experiences.

Given the importance of course experiences to student persistence, we sought to investigate concerns based on their potential to highlight course components and challenges that could be addressed by the instructor and potentially improve student outcomes. We focused on introductory courses because these courses have reputations as "weed-out courses" (Mervis 2011) and include experiences that influence students' decisions to remain in STEM (Alting and Walser 2007; Chang et al. 2008; Seymour and Hewitt 1997). Additionally, we explored if any concerns varied between students with different demographics because certain groups leave STEM majors at higher rates than others (Alexander et al. 2009; Beasley and Fischer 2012; Cataldi et al. 2018; Hughes 2018; Riegle-Crumb et al. 2019). Differences in concerns between demographic groups could reveal specific areas that instructors can address to provide students with positive experiences. To describe student concerns in introductory STEM courses, we investigated the following research questions: (1) what concerns do students hold about their introductory courses, (2) how do students' concerns in introductory courses change within a semester, and (3) do concerns differ based on student demographic characteristics?

Materials and Methods

Identifying the Range of Concerns

To explore course-based concerns, we distributed surveys to undergraduate students in introductory STEM courses (Table 1). In Fall 2017, we conducted pilot surveys during both the first-week and mid-semester in 13 introductory STEM course sections (disciplines surveyed included biology, chemistry, and physics) at two public, research-intensive universities and received 2181 student responses from the first-week and 1920 responses from the mid-semester data collection (Supplemental Table 1). The courses included in the pilot survey were taught by nine different instructors who typically used lecture or interactive lecture. For the interactive lecture, the instructors utilized clicker questions, peer discussion, or small group activities (personal communication and past observations).

Two open-ended questions were included in the pilot survey: (1) "How do you expect the use of class time in [course title] to be different from the [subject] course(s) you took in high school?" and (2) "What concerns, if any, do you have regarding these differences in how class time is used?" We included the first question to focus student responses on differences between their high school and college courses. We performed inductive content analysis on all responses to the second question (Thomas 2006). Two co-authors (coders) AKL and AIM separately read through 100 responses collected at one institution during the first week of class. Together, they developed a list of ideas appearing in those responses, which served as an initial codebook. The coders separately tested this codebook on sets of 50–100 responses cycling through both institutions and the first-week and mid-semester datasets. After each iteration, the coders met to compare results and modify the codebook, and this process continued until the coders constructed a finalized codebook.

Next, the coders calculated a consensus estimate of interrater reliability to determine if they held similar understandings of the codes or if any code definitions required further refinement (Stemler 2004). Each coder independently coded 50 random responses, which included all combinations of institutions and first-week and mid-semester surveys. If 90% agreement per code was not reached, the previous steps were repeated.

A single student response could be coded for more than one concern. Ten percent of student responses were coded as "other" because their concerns varied widely from the codes, the responses were challenging to interpret, or the responses were not on topic (e.g., criticizing the instructor instead of describing a concern). The coders reviewed all responses coded as "other" together to make sure no potential codes were overlooked, which resulted in two additional codes (concerns due to large class sizes and a lack of efficiency in how class time was used), which were less common ideas but were shared among students. These responses were coded to consensus by both coders.

Table 1 Overview of data collection

Semester	Administration	Description
Fall 2017	First-Week and Mid-Semester	Pilot survey: Open-ended questions identifying course-based concerns
Fall 2018	First-Week and Mid-Semester	Final survey: • Closed-ended questions about 13 identified concerns • Open-ended question identifying top concern

Final Survey Content

The final codebook included 17 codes, 13 of which we modified to become closed-ended items for the final survey implemented in Fall 2018 (Supplemental Table 2). The codes were adapted from their original wording to use language that would be accessible to students and align with the question format. We excluded codes that captured student responses that explicitly described their lack of concerns or their preferences for college courses compared to their high school courses. Consequently, the final list of 13 items only included student-generated concerns and represented both common and uncommon ideas from the pilot survey (Supplemental Appendix A).

In the final survey (Table 1), we asked students to indicate their level of concern for each of the 13 items through a 3-point Likert scale question, with 0 being "Not Concerned," 1 being "Somewhat Concerned," and 2 being "Very Concerned." The question layout included three boxes for students to drag and drop each item into the bin that best captured their level of concern (Supplemental Appendix B). The surveys were built and distributed using Qualtrics (Provo, UT). Each survey included two additional short answer questions asking students to identify their top course-based concern out of the 13 items and explain why it was their top concern. At the end of each survey, we included seven optional demographic questions (Supplemental Appendix B).

Final Survey Administration

During Fall 2018, we distributed final surveys online, out-of-class during the first week and mid-semester to students in 22 introductory STEM course sections at three research-intensive universities. These courses were taught by faculty participating in a professional development group that met on a monthly basis to explore ways to help students with the transition from high school to college STEM courses. Student responses were voluntary, and faculty varied in their distribution of participation credit points as incentives for participation. The course subjects in our study broadly covered STEM disciplines and included biology, chemistry, computer science, earth science, ecology and environmental science, economics, engineering, forestry, mathematics, physics, and statistics (National Science Foundation, National Center for Science and Engineering Statistics 2015). The total course enrollment was 3916 students.

Data Analysis

We received 2436 student responses for the first-week survey and 1671 responses from the mid-semester survey. We removed student responses from the final dataset according to the steps detailed in Supplemental Figure 1. Broadly, for each survey we removed responses from students that (1) were not complete survey responses; (2) were responses from students who filled out the survey for a single course multiple times (after keeping their first response); or (3) did not include responses to all of the demographic questions. We then matched students' responses for those who responded to both the first-week and mid-semester survey by name and student ID. For matched students who answered both surveys, we removed those who changed their answers for demographic questions from the first-week to the mid-semester survey from the dataset.

If a matched student left a demographic question blank on one survey, but answered it on the other survey, their answer was filled in to match on both surveys. After matching student responses, we removed student responses who did not rank their level of concern for all 13 items and did not receive a final course grade. This processing left 650 student responses. Demographic information about student participants is included in Table 2.

Top Concern Question Analysis

In the final surveys, students were additionally asked to identify their top concern by submitting a written response. Two authors, AKL and JKS (coders), performed deductive content analysis using the 13 items as the codebook to characterize students' top concern. In cases where responses listed more than one top concern, the response was deleted from the data set since it did not follow the prompt. Some responses contained top concerns that were not from the 13 items. These responses were coded as "other." After familiarizing themselves with the codebook, the coders independently coded 90 responses (15 per survey per institution) and compared results while discussing any differences. This process was done iteratively until all codes reached 90% agreement. JKS coded the remaining responses, consulting with AKL as necessary for unclear cases. Lastly, JKS and AKL returned to all responses coded as "other" to determine if there were common ideas; however, any new concern appeared in only 10 or fewer responses in either the first-week or mid-semester datasets.

Concern Index

To explore general levels of student concern over time, we sought to condense multiple items into one score representing overall student concern. We calculated the correlation coefficients between the 13 items during the final first-week and mid-semester surveys and identified seven items with moderate ($r \geqq 0.3$) correlations. These seven items also loaded onto the first component of a Principal Components Analysis (PCA) for the first-week and mid-semester survey data (Supplemental Appendix C). These items include (1) knowing what to study, (2) the course being too difficult, (3) the pace of the course being too fast, (4) being expected to do too much independent learning outside of class, (5) having the necessary skills/background to succeed in the course, (6) receiving too few in-depth explanations, and (7) being able to get help. Since student scores ranged from 0 (not concerned), 1 (somewhat concerned), or 2 (very concerned) on each of these seven items, we collapsed the seven items into one cumulative score that had a potential range between 0 and 14. This cumulative raw score represented their Concern Index, or CI.

We also explored adjusting student responses to each of the seven items by multiplying each response by its corresponding loading scores onto the first component (Supplemental Appendix C). We summed these adjusted responses to create an adjusted Concern Index (Adj.CI). We found strong correlation between the adjusted and non-adjusted CIs ($r = 0.99$, Supplemental Appendix C), so for ease of interpretation, we used the non-adjusted CI values for further analyses. Finally, we ran reliability analyses of student responses to the seven CI items in R using the psych package (R Core Team 2019; Revelle 2018). Students responded

Table 2 Student demographics from Fall 2018

Total	650
English as first language	
Yes	612 (94%)
No	38 (6%)
First Generation status	
First Generation	176 (27.1%)
Continuing Generation	474 (72.9%)
First-Semester student	
First-Semester	324 (49.8%)
Returning Student	326 (50.2%)
Gender	
Male student	285 (43.8%)
Female student	365 (56.2%)
International	
Domestic	630 (96.9%)
International	20 (3.1%)
Transfer	
Non-Transfer	584 (89.8%)
Transfer	66 (10.2%)
Underrepresented Minority	
URM	93 (14.3%)
Non-URM	557 (85.7%)
Course Size	
Small	17 (2.6%)
Medium	103 (15.8%)
Large	530 (81.5%)
Discipline	
Biology (5 course sections)	193 (29.7%)
Chemistry (2 course sections)	101 (15.5%)
Computer science (4 course sections)	50 (7.7%)
Earth science (1 course section)	11 (1.7%)
Economics (1 course section)	9 (1.4%)
Engineering (1 course section)	5 (0.8%)
Forestry (1 course section)	24 (3.7%)
Mathematics (3 course sections)	28 (4.3%)
Physics (2 course sections)	126 (19.4%)
Statistics (2 course sections)	103 (15.8%)
University	
1	285 (43.8%)
2	295 (45.4%)
3	70 (10.8%)

consistently across these seven items on both the first-week (Cronbach's $\alpha = 0.76$) and mid-semester surveys (Cronbach's $\alpha = 0.81$).

Statistical Analyses and Data Visualization

To assess the impact of demographic characteristics on total levels of student course-based concern, we followed the recommendations outlined in Theobald (2018) and used automated model selection to identify separate models for the first week and mid-semester (Supplemental Appendix D). The demographic variables considered during model selection were English as a first language, first-generation student status, first-semester on a college campus status, gender, international student status, transfer student status, and underrepresented racial/ethnic minority student status.

For the mid-semester model selection, we also tested three additional variables: students' first-week CI as a predictor variable to account for baseline levels of student concern, final standardized course grades as a predictor variable to account for the impact of performance in the course, and the percentage of female students enrolled in students' courses to test for the impact of gender balance within a course. We included students' first-week CI as a predictor because first-week CI and mid-semester CI showed a moderate correlation ($r = 0.58$). We included final standardized grades because in our dataset 20 out of 22 course sections (or 638 out of 650 students) had administered exams by the mid-semester time point and student knowledge of their course performance could impact student concerns. In order to account for the differences in how each course weighted mid-semester exams compared to other course assignments, we included final standardized grades as a predictor variable for overall course performance. Finally, we included the percentage of female students enrolled in students' courses, obtained from each university's registrar, as a predictor because gender balance varied across courses, with the percent of female students ranging between 9% and 64% (Supplemental Figure 2). We also included an interaction between gender and the percentage of female students enrolled to test for potential differential impacts of gender balance on male and female students.

Across the three universities in this study, there were two different GPA scales, with one university using a 4.3 scale and the others using a 4.0 scale. To account for grading scheme differences across universities, we standardized students' final course GPAs using z-scores, which then represented the distance between students' raw final grades and the population mean from each university in units of standard deviation. We calculated z-scores using the formula z-scores $= (X - \mu) / \sigma$, where X is the score of interest, μ is the class mean score, and σ is the standard deviation.

We compared models with all combinations of fixed effects using the MuMIn R package, which evaluates models using measures of Akaike's information criterion (AIC) (Bartoń 2019). All modeling was conducted using the R statistical software (v.3.3.1) and the lmer and lmerTest packages (Bates et al. 2015; Kuznetsova et al. 2017; R Core Team 2019).

We constructed graphs and figures using the ggplot2, Rcolorbrewer, and ggalluvial packages in R (Brunson 2019; Neuwirth 2014; Wickham 2016). For construction of the Sankey diagrams, we designated students with final standardized grades (i.e., z-scores) between −0.5 - 0.5 SDs from the mean as receiving “medium grades” and students above 0.5 SD or below −0.5 SDs from the mean as receiving “higher grades” and “lower grades,” respectively. In order to categorize CI, we split the data into tertiles and designated students with a CI between 0 and 4 as “low concern,” 5–9 as “medium concern,” and 10–14 as “high concern.”

Results

What Are the Types of Course-Based Concerns Students Have?

In the pilot survey, introductory STEM students responded to an open-ended question during the first week and mid-semester asking them to identify what, if any, concerns they had regarding the differences in how class time was used in their high school and college STEM introductory courses. Inductive content analysis of these answers (first-week $n = 2181$, mid-semester $n = 1920$) identified 13 common course-based concerns, which included concerns related to the structure and content of the course as well as students’ backgrounds and instructional preferences. Nearly all of the survey responses focused on general course-based concerns rather than concerns about the difference in class time between high school and college. The most common course-based concern from the pilot first-week survey was not getting help or knowing where to get help. Students with this concern described worry over not having enough personal access to their instructors, falling behind and not being able to find help, and not getting immediate feedback. One student responded, “Sometimes I feel that I will get behind in class and feel that I won’t have anyone to help,” while another wrote that they were “concerned that we won’t be able to get immediate feedback on practice problems.” Additional representative student responses are included in Supplemental Table 2.

The open-response results from the pilot survey helped to identify types of course-based concerns held by students. However, the open-ended response format resulted in variation in the types of responses we received from students, with some students only submitting one response, and some students submitting multiple concerns. Furthermore, students may have had additional concerns beyond the 1–2 sentences included in their responses. To compare levels of concern held by students across all of the identified course-based concerns, we modified the survey for the Fall 2018 final implementation as described in the methods and in Supplemental Appendix B. Converting the open-response question to a closed-response format allowed us to identify which course-based concerns were the most common and if students changed in their levels of concerns about these items over the course of the semester.

We used stacked bar charts to display the relative frequency of student levels of concern for each of the 13 items. Knowing what to study was an item that students often expressed being “Very Concerned” about, with 34% of students reporting this level during the first week of the semester and 26% at mid-semester (Fig. 1). In addition, knowing what to study was also the item that students most frequently cited as their “top concern” with 10.3% citing it at the start of the semester and 11.1% at the mid-semester time point (Fig. 2).

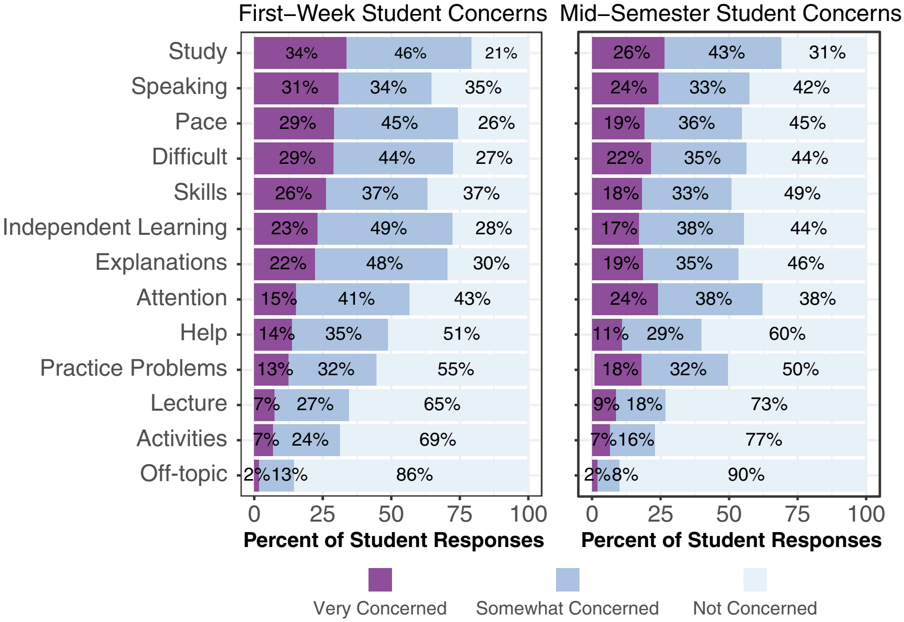

Fig. 1 Percent of student responses at each level of concern for the 13 items, ordered by percent of students citing each level of concern on the final first-week and mid-semester surveys

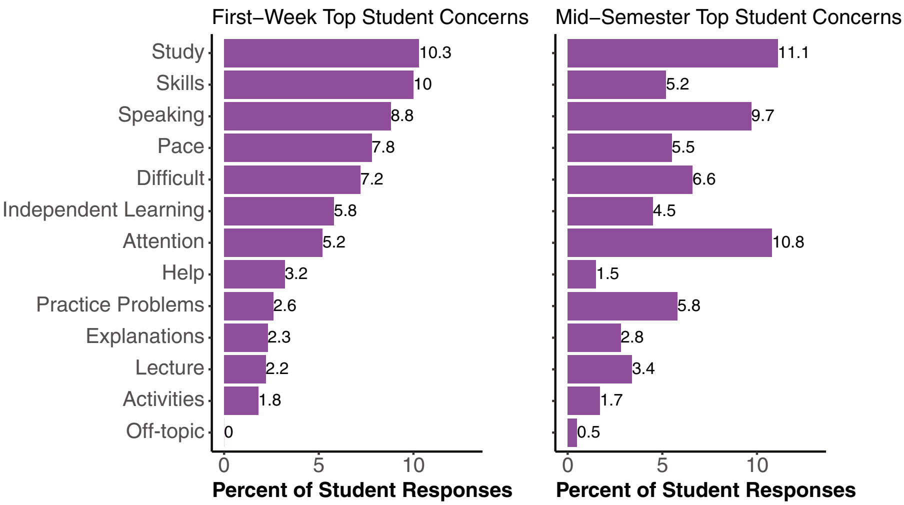

Fig. 2 Top concerns held by students during the final first-week and mid-semester surveys. Percent of student responses who cited each of the 13 items as their top concern, ordered by percent of students citing each level of concern during the first-week survey. Students who reported not having any concerns, students who did not submit a response, students who identified multiple top concerns, and students who wrote responses other than the 13 categories are excluded from the figure. Concern labels are abbreviated, full labels are provided in Supplemental Appendix C

How Do Student Course-Based Concerns Change within a Semester?

We explored how student course-based concerns changed over the semester, across multiple levels of resolution. For approximately half of their selections, students did not change in their levels of reported concern about individual items during the semester, but among students who shifted in their levels, more students decreased in their levels of concern (Table 3). There were two items for which more students reported being very or somewhat concerned at mid-semester compared to during the first-week: being able to pay attention in class and having enough practice problems (Fig. 1). While at the mid-semester time point knowing what to study was still the top concern, being able to pay attention in class was the second most common concern from students (Fig. 2).

We used the Concern Index (CI, see Methods) as a metric to compare if overall course-based concern levels change between the first-week and mid-semester surveys. During the first week, the mean CI reported by students was 6.5 (or 46.4% of total possible concern; Fig. 3). Overall student concerns decreased between the beginning and middle of the semester. Mid-way through the semester, the mean CI reported by students was 5.1 (or 36.5% of total possible concern), which was significantly lower ($p < 0.001$) than mean student CI values at the beginning of the semester.

Do Concerns Differ Based on Student Demographic Characteristics?

Given the range of concern levels held by students during the first-week and mid-semester surveys, we explored if CI findings varied by student-level and course-level demographic variables (Supplemental Appendix D). The best-fitting model for concerns during the first week included first-generation student status, first-semester on a

Table 3 Percent of students who maintained their levels of concern, increased in their levels of concern, or decreased in their levels of concern about the 13 items

Item	Maintained levels of concern	Increased in levels of concern	Decreased in levels of concern
Study	52%	17%	31%
Speaking	55%	18%	27%
Pace	50%	13%	36%
Difficult	54%	14%	32%
Skills	55%	15%	30%
Independent Learning	47%	16%	37%
Explanations	52%	16%	33%
Attention	56%	28%	16%
Help	55%	18%	28%
Practice problems	52%	29%	19%
Lecture	65%	14%	21%
Activities	69%	12%	19%
Off-topic	83%	7%	11%

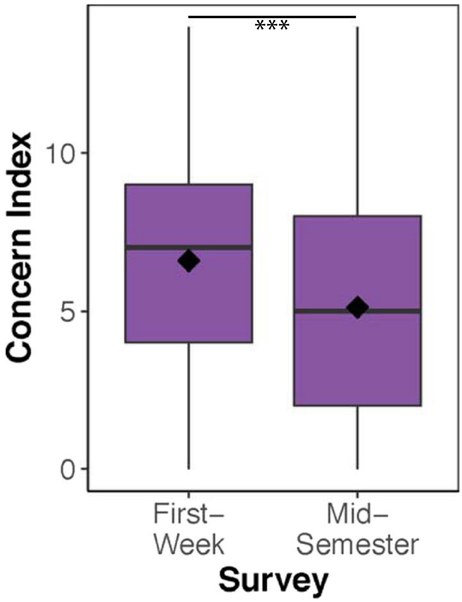

Fig. 3 Boxplot showing distribution of student concern index (CI) during the first-week and mid-semester surveys. Wilcoxon signed rank test for significance with a p value <0.001 indicated by ***. Diamonds represent mean values and horizontal lines on each bar represent median values

college campus status, and gender as predictors, and explained 5% of the variation in the first week CI data (Table 4). While the amount of variation explained by the model is low, this model was significant, indicating that these variables had a significant relationship with the CI of the students during the first week of the semester.

According to the model, first-generation, female, first-semester students had a CI of 8.13. This value is the equivalent of reporting "very concerned" for four out of seven items, or a mix of "somewhat" and "very concerned" for more than four items. The effects of being continuing-generation, male, and a returning student were all negative

Table 4 Estimated coefficients for variables from the best-fitting linear regression that estimates how different predictors impact the Concern Index (CI) from the first week of the semester

Predictors	Estimate	Std. Error	t value	p value[a]	Confidence Interval (CI)
(Mean Intercept)	8.13	0.31	26.18	< 2e-16 ***	7.52–8.74
Continuing-generation student	−0.73	0.29	−2.56	0.011 *	−1.29 - -0.17
Returning student	−0.91	0.26	−3.51	0.00047 ***	−1.41 - -0.40
Male student	−1.28	0.26	−5.0	9.18e-07 ***	−1.79 - -0.77

The intercept represents the CI of a first-generation, female, first-semester student

Residual standard error: 3.22 on 646 degrees of freedom

Multiple R-squared: 0.0556, Adjusted R-squared: 0.0512

F-statistic: 12.66 on 3 and 646 DF, p value: 4.72e-08

Statistical significance is indicated by *, $p < 0.05$; **, $p < 0.01$; *a* and ***, $p < 0.001$

and significant, indicating that these demographic groups were associated with a decrease in first week CI. The best fitting model predicted that a continuing-generation, male, returning student would report a CI of 5.21.

We also determined if changes in mid-semester CI were similar for students from different demographic groups. We used model selection to identify if any demographic variables explained variation in mid-semester student CI and included first-week CI to control for students' initial levels of concern (Supplemental Appendix D). Additionally, we included final standardized grades (z-scores, see Methods) from students as a predictor for each model to account for differences in course performance.

The best-fitting model for mid-semester CI explained 41% of the variation in student responses. According to the model, a female, first-generation, domestic student who grew up speaking English at home, with an average CI from the first week of the semester, who received the average final grade at their institution, and who was enrolled in a course with the average percentage of female students (39% female) would report a CI of 5.84 (Table 5). Although first-semester on a college campus was

Table 5 Estimated coefficients for variables from the best fitting linear regression that estimates how different predictors impact the Concern Index (CI) mid-way through the semester

Predictors	Estimate	Std. Error	*t* value	*p* value[a]	Confidence Interval (CI)
(Mean Intercept)	5.84	0.24	24.30	< 2e-16 ***	5.36–6.31
Initial CI	0.58	0.03	16.95	< 2e-16 ***	0.51–0.64
Final Standardized Grade	−0.78	0.11	−7.01	5.91e-12 ***	−1.00 - -0.56
Percentage of Female students in a course	−0.048	0.01	−4.74	2.61e-06 ***	−0.07 - -0.28
Male student	−0.66	0.24	−2.80	0.00534 **	−1.12 - -0.20
Continuing Generation student	−0.43	0.25	−1.72	0.086	−0.93 – 0.06
International student	1.15	0.66	1.73	0.085	−0.16 – 2.45
Student who did not speak English growing up	−0.74	0.49	−1.52	0.13	−1.71 – 0.22
Male student: Percentage of Female students in a course (Interaction)	0.033	0.013	2.49	0.013 *	−1.71 – 0.22

First week CI and the percentage of female students enrolled in their course were included as predictor variables and were centered. Students' standardized final course grades were also included. The intercept represents a female, first-generation, domestic student who spoke English at home growing up enrolled in a course with the average percentage of female students (39%), with an average CI from the first week of the semester and who received the average final grade at their university

Residual standard error: 2.767 on 641 degrees of freedom

Multiple R-squared: 0.4155, Adjusted R-squared: 0.4083

F-statistic: 56.97 on 8 and 641 DF, p value: < 2.2e-16

Statistical significance is indicated by *, $p < 0.05$; **, $p < 0.01$; and ***, $p < 0.001$

significant during the first week, it was not part of the best-fitting model at the mid-semester point. A one unit increase in initial CI is associated with a higher (0.58 unit) mid-semester CI, reflecting that while overall student concerns decrease during the semester, students' initial levels of concern can predict their levels of concern midway through the semester. Similarly, a one-standard deviation increase in final grade was associated with a 0.78 unit decrease in mid-semester CI. In other words, students who ultimately performed better than their peers in the course were less concerned than their peers midway through the semester.

Four demographic variables were included in the best-fitting model for the mid-semester time point after controlling for initial CI and final standardized grades (Table 5). English spoken at home and international student status were included but were not significant. The effect of being a student who spoke English at home lowered CI relative to the intercept, while the effect of being an international student raised CI relative to the intercept. The low number of students who spoke English at home ($n = 38$) as well as international students in our sample ($n = 20$) and large confidence intervals for these students requires further investigation with a larger sample size. First-generation status was similarly included in the mid-semester model, but it was also not a significant predictor of CI. One possible explanation for this result is that the effect of being a first-generation student was masked by differences in initial concern or final standardized grades. In our dataset, we found that on average first-generation students received lower final standardized grades than their peers (Supplemental Figure 3).

On the other hand, the effect of being a female student was significant after controlling for higher initial concern, final standardized grades, and percentage of female students in the course. A female student who had reported the same level of initial concern as a male peer, received the same final standardized grade, and was enrolled in a course with the average percentage of female students (39%) reported a 0.66 unit higher level of mid-semester concern than a male peer (Table 5). This outcome is notable because female students had no significant differences in raw or standardized final grades from male students (Supplemental Figure 3). Taken together, the results from these models suggest that initial course-based concern and final course performance impact mid-semester student course-based concern, but even with controlling for these variables, female students still have significantly more total concern than their peers mid-way through the semester.

The models identified that gender was the only significant demographic variable with an effect on mid-semester CI. Therefore, we used Sankey diagrams to provide a visual representation of students' changes in initial CI to mid-semester CI disaggregated by gender (Fig. 4a). The Sankey diagrams confirmed that male students hold consistently lower levels of concern during both the beginning and middle of the semester than their female peers. While the overall percent of students increasing in their levels of concern is smaller than the percent of students decreasing in their concern (indicated by flow size in Fig. 4a), a larger percentage of female students increased in their concern compared to male students.

The percentage of female students enrolled in a course also affected students' CI. A 1 % increase in the percentage of female students enrolled in a course was associated with a -0.048 unit decrease in CI for female students (Table 5). However, the interaction of gender with the percentage of female students showed

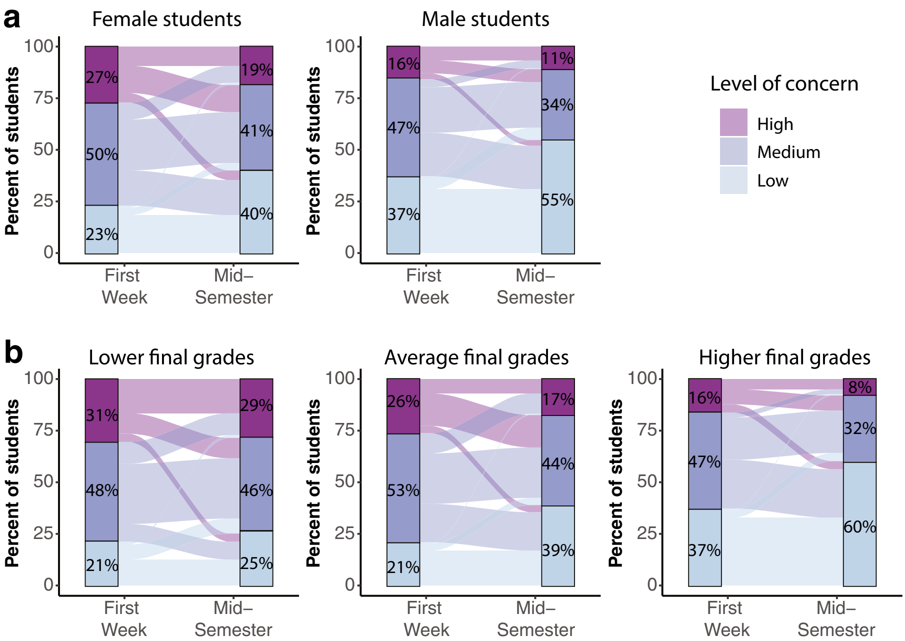

Fig. 4 Sankey diagrams displaying individual student shifts in CI (low = 0–5, medium = 5–9, high = 10–14), disaggregated by (**a**) gender or (**b**) final grades. The size of each flow is proportional to the percent of students who reported similar CIs to one another during the first-week and mid-semester surveys

that female student concerns decreased more sharply than male student concerns as the percent of female students in a course increased. The difference in rate is calculated by starting with the interaction (0.033) and adding the effect of the percentage of female students enrolled in a course (−0.048), which results in a − 0.015 unit decrease in CI for male students per 1 % increase in female student enrollment. The model predicts a steeper decrease in CI for female students when compared to male students, as enrollment of percent female students increases (Supplemental Figure 4).

Our regressions identified that course performance had a significant effect on mid-semester CI. The Sankey diagrams also revealed interesting patterns for levels of concern based on final standardized grades (Fig. 4b). Students who received lower final standardized grades held higher levels of concern at the beginning of the semester, and the majority of them did not decrease in their levels of concern. On the other hand, students who received higher final standardized grades in the course were initially less concerned than their peers and a larger proportion of them decreased in their levels of concern.

We also examined whether any individual concern was more likely to correlate with course performance, because studies have shown that students' perceived course difficulty is inversely related to course performance (England et al. 2019). Each of the concerns in our study is negatively correlated with final standardized grades but the overall CI has the strongest correlation (Supplemental Table 3), indicating that no single concern is driving the connection with course performance in our dataset.

Discussion

We described concerns held by undergraduate students in introductory STEM courses, observed if those concerns changed over the course of the semester, and assessed if there were differences in concerns based on student characteristics such as gender or course grades. We identified a broad range of course-based issues that may be leveraged to improve undergraduates' experiences in introductory STEM courses (Figs. 1 and 2). Students' total concern level, measured using the Concern Index (CI), was used to investigate demographic differences in concern and the relationship between student concern and course performance (Tables 4 and 5, Fig. 4).

What Concerns Do Students Hold about their Introductory Courses?

Students in introductory STEM courses expressed a variety of course-based concerns, which we further quantified using closed-ended survey items. For every item, there were students who reported being concerned indicating that these items resonated with students across three campuses and in many different introductory STEM courses (Fig. 1). The list of concerns included topics that may be actionable by the instructor, such as those related to course structure or pedagogy (e.g., pace of the class). There were also concerns related to students incoming preparation and skills (e.g., having the necessary skills/background to succeed in this course), which may indicate areas that could be covered early in the semester to help students with the transition from high school to college STEM courses or that could be provided through supplemental instruction programs.

How Do Students' Concerns in Introductory Courses Change within a Semester?

We used CI as a metric for students' overall level of concern and found that student concern went down from the first week to mid-semester (Table 3 and Fig. 3). When investigating the range of concerns (Fig. 1) and students' top concerns (Fig. 2), the highest reported concern during the first week and mid-semester time point was not knowing what to study. These results may indicate that students were not developing study skills over the course of the semester. Providing students with study strategies early in the course as well as directing them towards specific resources to help them build their study skills may reduce this concern over the semester as well as improve student grades.

Interestingly, at the mid-semester time point, not being able to pay attention for the entire class period rose to become a commonly reported concern (Figs. 1 and 2). Given that undergraduate STEM courses predominantly use lecture (Stains et al. 2018), one way to increase student attention could be through the introduction of more active learning (Lane and Harris 2015). Another potential issue related to attention is that students can use electronic devices for a wide array of non-academic purposes, even though they report knowing that technology use can distract them from learning (McCoy 2013; Sana et al. 2013; Tindell and Bohlander 2012). One way to help students pay attention is to couple the use of technology in the classroom to active learning strategies, such as discussion boards and online problem solving (e.g., Barak et al. 2006).

Do Concerns Differ Based on Student Demographic or Course Characteristics?

Our work adds student concern to the growing list of differences between how female and male students experience STEM courses. In our study, female students reported higher levels of concern than male students at the beginning of the semester (Table 4), and their concerns remained higher than male students at the mid-semester point even when controlling for initial concerns, course performance, and the percentage of female students enrolled in their courses (Table 5 and Fig. 4a). This pattern of concern is consistent with that seen for course-related anxiety as previous work has shown that female students have increased anxiety and test anxiety (e.g., Chapell et al. 2005; Harris et al. 2019; Núñez-Peña et al. 2016; Salehi et al. 2019). Additionally, female students can have lower participation in STEM classes and have different views of success and failure when compared to male students with comparable grades (e.g., Eddy et al. 2014; Freedman et al. 2018; Grunspan et al. 2016; Lowe, 2015; Marshman et al. 2018; Robnett and Thoman 2017). However, our work additionally reveals that as courses increase in their percentage of female students, levels of mid-semester concern decrease for all students and at a higher rate for female students (Table 5, Supplemental Figure 4). This variable was not identified in the first-week model, which could indicate that as the semester progresses, classroom environments differ in courses based on the percentages of female students enrolled. Programs designed to aid and increase representation of women in STEM may want to consider measuring concern levels of female students and ascertain whether targeted interventions can lower female concerns and impact subsequent performance.

Studying the change in concern across two time points also revealed patterns in the relationship between students' final standardized grades and course-based concerns. Students who ultimately performed below the mean held higher levels of concerns at the first-week and mid-semester time points (Fig. 4b). Since students with lower performance enter the course with higher concern, it may be possible to ask students about their concerns, including the commonly identified concerns in Figs. 1 and 2, and instructors could talk to their students and make adjustments to alleviate top concerns. Instructors may also consider using a concern metric similar to a CI for identifying students who could benefit from early intervention, such as peer tutoring or supplemental instruction (e.g., Batz et al. 2015; Deslauriers et al. 2012; Lizzio and Wilson 2013; Stanich et al. 2018). Studies about these kinds of supplemental resources often report evidence of effectiveness for improving student performance and retention. However, do such supplemental resources also impact student anxiety and concern? Combining course performance with anxiety and concern measures may also help advance research on course interventions for struggling students.

Conclusions and Future Directions

In this study, we measured course-based concerns, but there are many other factors that may affect students. For example, students likely also hold concerns outside of the course that have potential to impact their performance, such as those about financial resources. Furthermore, our work focused on undergraduates at research institutions, and students at other institution types (e.g., community colleges, regional comprehensives, or minority serving) may hold different concerns. It is important for future work to explore the concerns of students at other institution types to capture the variety of concerns and assess if different interventions are required to provide those students with positive course experiences.

Future work should also investigate the concerns of international students and students who did not speak English at home growing up. While international student status and growing up speaking English at home were included in the mid-semester model, they were not significant (Table 5). Low numbers of international students and students who are non-native English speakers in our sample could be one reason for this unclear result and indicates that more data are needed to explore these trends. However, combining these results with Eddy et al.'s (2015) finding that international students report higher levels of anxiety related to peer discussions, suggests that the concerns and anxiety of international students is an important future direction.

Further investigation is also required to understand the complex interplay between student characteristics, concerns, and success. For example, the relationship between first-generation status, concerns, and grades is complex. First-generation status was identified in the best-fitting model but was not significant when accounting for final standardized grades (Fig. 3), which may be because first-generation students had lower performance than continuing generation students in our sample thereby conflating the measures of performance and first-generation status (Supplemental Figure 3). More investigation is required to empirically study first-generation students' experiences and disentangle grades and first-generation status when it comes to concerns.

Our work focused on identifying course-based concerns held by students across STEM disciplines, but future work could compare variation of course-based concerns within disciplines. To do this, student concern data could be collected across numerous courses within the same STEM disciplines. In addition, the relationship between gender and course gender representation with CI in a discipline should also be investigated. Previous work revealed that female students have lower levels of confidence in their mathematical abilities than male students in calculus and are 1.5 times more likely to leave STEM after college calculus (Ellis et al. 2016). Similar trends are seen in physics and computer science (Marshman et al. 2018).

Taken together, this work focused on identifying student concerns across several demographic groups. There are many studies that aim to describe sources of student anxiety or concern, but to date there are few that use measurements of anxiety or concern to assess the impact of specific interventions. Future investigations of the links between student concerns and instructional practices would be beneficial for educators. For example, autonomy-supporting teaching practices such as providing students with decision-making opportunities related to classroom management, choice in how to present their ideas, and opportunities to evaluate work (Reeve 2016; Stefanou et al. 2004; Williams and Deci 1996) aim to provide students with learning environments that support their daily autonomy. Higher student autonomy support from teachers is correlated with higher student self-efficacy and lower drop-out intention among first-year students (Girelli et al. 2018). Interventions that focus on supporting student self-direction may alleviate a number of course-based student concerns. While we can speculate on practices that may help lessen anxiety and concerns for students, their effectiveness would require further evidence. When studying these interventions, it will be important to look at any differing impact based on student demographics, especially gender. Addressing the concerns of female students in particular could be an important step to reducing disparities in STEM majors and future careers.

Acknowledgements We thank the members of the Cornell Physics Education Research Group for their feedback on this article. We also greatly appreciate the faculty and students who participated in this study.

Funding Information This material is based on work supported by the National Science Foundation under grants DUE-1712074, DUE- 1712060, and DUE-1347814; the Provost's Gateway Initiative and the Center for Teaching Innovation at Cornell University; and the Nebraska Science Program of Excellence at the University of Nebraska-Lincoln.

Compliance with Ethical Standards

Competing Interests The authors declare no competing interests.

Ethical Approval This research was considered exempt from institutional review: University of Maine protocol 20170512, University of Nebraska-Lincoln protocol 20170617341, and Cornell University protocol 1806008047.

References

Akiha, K., Brigham, E., Couch, B. A., Lewin, J., Stains, M., Stetzer, M. R., et al. (2018). What types of instructional shifts do students experience? Investigating active learning in science, technology, engineering, and math classes across key transition points from middle school to the university level. *Frontiers in Education, 22*, 1–18. https://doi.org/10.3389/feduc.2017.00068.

Alexander, C., Chen, E., & Grumbach, K. (2009). How leaky is the health career pipeline? Minority student achievement in college gateway courses. *Academic Medicine, 84*(6), 797–802.

Alting, A., & Walser, A. (2007). Retention and persistence of undergraduate engineering students: "What happens after the first year?". *ASEE, 2007*, 9.

Barak, M., Lipson, A., & Lerman, S. (2006). Wireless laptops as means for promoting active learning in large lecture halls. *Journal of Research on Technology in Education, 38*(3), 245–263.

Barrows, J., Dunn, S., & Lloyd, C. A. (2012). Anxiety, self-efficacy, and college exam grades. *Universal Journal of Educational Research, 1*(3), 204–208.

Bartoń, K. (2019). MuMIn: Multi-Model Inference. *R package version, 1*(43), 6 https://CRAN.R-project.org/package=MuMIn.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Batz, Z., Olsen, B. J., Dumont, J., Dastoor, F., & Smith, M. K. (2015). Helping struggling students in introductory biology: A peer-tutoring approach that improves performance, perception, and retention. *CBE—Life Sciences Education, 14*(2), ar16.

Beasley, M. A., & Fischer, M. J. (2012). Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education, 15*(4), 427–448. https://doi.org/10.1007/s11218-012-9185-3.

Bellinger, D. B., DeCaro, M. S., & Ralston, P. A. (2015). Mindfulness, anxiety, and high-stakes mathematics performance in the laboratory and classroom. *Consciousness and Cognition, 37*, 123–132.

Brown, T. L., Brazeal, K. R., & Couch, B. A. (2017). First-year and non-first-year student expectations regarding in-class and out-of-class learning activities in introductory biology. *Journal of Microbiology & Biology Education, 18*(1), 1–9.

Brunson, J. C. (2019). ggalluvial: Alluvial diagrams in 'ggplot2'. R package version 0.10.0. https://CRAN.R-project.org/package=ggalluvial.

Cataldi, E. F., Bennett, C. T., Chen, X., & Simone, S. A. (2018). First-generation students: College access, persistence, and postbachelor's outcomes (NCES 2018-421). Retrieved from National Center for Education Statistics.

Chang, M. J., Cerna, O., Han, J., & Sàenz, V. (2008). The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *Review of Higher Education: Journal of the Association for the Study of Higher Education, 31*(4), 433–464. https://doi.org/10.1353/rhe.0.0011.

Chapell, M. S., Benjamin Blanding, Z., Takahashi, M., Silverstein, M. E., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268–274. https://doi.org/10.1037/0022-0663.97.2.268.

Cooper, K. M., Downing, V. R., & Brownell, S. E. (2018). The influence of active learning practices on student anxiety in large-enrollment college science classrooms. *International Journal of STEM Education, 5*(1). https://doi.org/10.1186/s40594-018-0123-6.

Deslauriers, L., Harris, S. E., Lane, E., & Wieman, C. E. (2012). Transforming the lowest-performing students: An intervention that worked. *Journal of College Science Teaching, 41*(6), 80–88.

Eagan, K., Hurtado, S., Figueroa, T., & Hughes, B. (2014). Examining STEM pathways among students who begin college at four-year institutions. *Paper Prepared for the Committee on Barriers and Opportunities in Completing 2- and 4- Year STEM Degrees*, (May), 1–27.

Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education, 13*(3), 478–492.

Eddy, S. L., Brownell, S. E., Thummaphan, P., Lan, M. C., & Wenderoth, M. P. (2015). Caution, student experience may vary: Social identities impact a student's experience in peer discussions. *CBE—Life Sciences Education, 14*(4), ar45.

Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave stem pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS One, 11*(7), 1–14. https://doi.org/10.1371/journal.pone.0157447.

England, B. J., Brigati, J. R., & Schussler, E. E. (2017). Student anxiety in introductory biology classrooms: Perceptions about active learning and persistence in the major. *PLoS One, 12*(8), 1–17. https://doi.org/10.1371/journal.pone.0182506.

England, B. J., Brigati, J. R., Schussler, E. E., & Chen, M. M. (2019). Student anxiety and perception of difficulty impact performance and persistence in introductory biology courses. *CBE—Life Sciences Education, 18*(2), ar21. https://doi.org/10.1187/cbe.17-12-0284.

Foley, A. E., Herts, J. B., Borgonovi, F., Guerriero, S., Levine, S. C., & Beilock, S. L. (2017). The math anxiety-performance link: A global phenomenon. *Current Directions in Psychological Science, 26*(1), 52–58. https://doi.org/10.1177/0963721416672463.

Freedman, G., Green, M. C., Flanagan, M., Fitzgerald, K., & Kaufman, G. (2018). The effect of gender on attributions for Women's anxiety and doubt in a science narrative. *Psychology of Women Quarterly, 42*(2), 178–191. https://doi.org/10.1177/0361684318754528.

Geisinger, B. N., & Raman, D. R. (2013). Why they leave: Understanding student attrition from engineering majors. *International Journal of Engineering Education, 29*(4), 914.

Girelli, L., Alivernini, F., Lucidi, F., Cozzolino, M., Savarese, G., Sibilio, M., & Salvatore, S. (2018). Autonomy supportive contexts, autonomous motivation, and self-efficacy predict academic adjustment of first-Year University students. *Frontiers in Education, 3*, 95. https://doi.org/10.3389/feduc.2018.00095.

Grunspan, D. Z., Eddy, S. L., Brownell, S. E., Wiggins, B. L., Crowe, A. J., & Goodreau, S. M. (2016). Males under-estimate academic performance of their female peers in undergraduate biology classrooms. *PLoS One, 11*(2), e0148405. https://doi.org/10.1371/journal.pone.0148405.

Harris, R. B., Grunspan, D. Z., Pelch, M. A., Fernandes, G., Ramirez, G., & Freeman, S. (2019). Can test anxiety interventions alleviate a gender gap in an undergraduate STEM course? *CBE—Life Sciences Education, 18*(3), ar35.

Hughes, B. E. (2018). Coming out in STEM: Factors affecting retention of sexual minority STEM students. *Science Advances, 4*(3), eaao6373.
Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26.
Lane, B. E. S., & Harris, S. E. (2015). A new tool for measuring student behavioral engagement in Large University classes. *Journal of College Science Teaching, 44*(6), 83–91.
Lizzio, A., & Wilson, K. (2013). Early intervention to support the academic recovery of first-year students at risk of non-continuation. *Innovations in Education and Teaching International, 50*(2), 109–120.
Lowe, P. A. (2015). Should test Anxiety be measured differently for males and females? examination of measurement bias across gender on measures of test anxiety for middle and high school, and college students. *Journal of Psychoeducational Assessment, 33*(3), 238–246.
Marshman, E. M., Kalender, Z. Y., Nokes-Malach, T., Schunn, C., & Singh, C. (2018). Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm? *Physical Review Physics Education Research, 14*(2), 20123. https://doi.org/10.1103/PhysRevPhysEducRes.14.020123.
McCoy, B. R. (2013). Digital distractions in the classroom: Student classroom use of digital devices for non-class related purposes. *Journal of Media Education, 4*, 5–14.
Meaders, C. L., Toth, E. S., Lane, A. K., Shuman, J. K., Couch, B. A., Stains, M., et al. (2019). "What will I experience in my college STEM courses?" an investigation of student predictions about instructional practices in introductory courses. *CBE—Life Sciences Education, 18*(4), 1–14. https://doi.org/10.1187/cbe.19-05-0084.
Mervis, J. (2011). Undergraduate science: Weed-out courses hamper diversity. *Science, 334*(6061), 1333. https://doi.org/10.1126/science.334.6061.1333.
National Science Foundation, National Center for Science and Engineering Statistics (2015). *Science and Engineering Degrees:* 1966–2012. Detailed statistical tables NSF 15–326. Arlington, VA. Available at http://www.nsf.gov/statistics/2015/nsf15326/.
Neuwirth, E. (2014). RColorBrewer: ColorBrewer palettes. R package version 1.1–2. https://CRAN.R-project.org/package=RColorBrewer
Núñez-Peña, M. I., Suárez-Pellicioni, M., & Bono, R. (2016). Gender differences in test anxiety and their impact on higher education students' academic achievement. *Procedia-Social and Behavioral Sciences, 228*, 154–160.
Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13–36). Amsterdam: Academic.
President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: U.S. Government Office of Science and Technology.
R Core Team (2019). R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. URL https://www.R-project.org/.
Reeve, J. (2016). Autonomy-supportive teaching: What it is, how to do it. In W. Liu, J. Wang, & R. Ryan (Eds.), *Building autonomous learners*. Singapore: Springer.
Revelle, W. (2018). Psych: Procedures for personality and psychological research. R package version 1.8.12. https://CRAN.R-project.org/package=psych.
Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher, 48*(3), 133–144.
Robnett, R. D., & Thoman, S. E. (2017). STEM success expectancies and achievement among women in STEM majors. *Journal of Applied Developmental Psychology, 52*, 91–100.
Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., et al. (2019). Gender performance gaps across different assessment methods and the underlying mechanisms: The case of incoming preparation and test anxiety. *Frontiers in Education, 4*, 107.
Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24–31.
Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27–41.
Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder: Westview Press.
Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., et al. (2018). Anatomy of STEM teaching in north American universities. *Science, 359*(6383), 1468–1470.

Stanich, C. A., Pelch, M. A., Theobald, E. J., & Freeman, S. (2018). A new approach to supplementary instruction narrows achievement and affect gaps for underrepresented minorities, first-generation students, and women. *Chemistry Education Research and Practice, 19*(3), 846–866.

Stefanou, C. R., Perencevich, K. C., DiCintio, M., & Turner, J. C. (2004). Supporting autonomy in the classroom: Ways teachers encourage student decision making and ownership. *Educational Psychologist, 39*(2), 97–110.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*(4), 1–19.

Suresh, R. (2006). The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention: Research, Theory & Practice, 8*(2), 215–239.

Theobald, E. (2018). Students are rarely independent: When, why, and how to use random effects in discipline-based education research. *CBE—Life Sciences Education, 17*(3), rm2.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237–246.

Tindell, D. R., & Bohlander, R. W. (2012). The use and abuse of cell phones and text messaging in the classroom: A survey of college students. *College Teaching, 60*(1), 1–9.

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching, 42*(5), 36–41.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis. Springer-Verlag*. New York.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology, 70*, 767–779.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Clara L. Meaders[1] · A. Kelly Lane[2] · Anya I. Morozov[2] · J. Kenny Shuman[2] · Emma S. Toth[3] · Marilyne Stains[4] · MacKenzie R. Stetzer[5,6] · Erin Vinson[6] · Brian A. Couch[2] · Michelle K. Smith[1]

1 Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA

2 School of Biological Sciences, University of Nebraska–Lincoln, 204 Manter, Lincoln, NE 68588-0118, USA

3 Noble Middle School, Berwick, ME 03901, USA

4 Department of Chemistry, University of Virginia, Charlottesville, VA 22904, USA

5 Department of Physics and Astronomy, University of Maine, Orono, ME 04469, USA

6 Maine Center for Research in STEM Education, University of Maine, Orono, ME 04469, USA