"No matter what your story is, there is a place for you in science": Students' Ability to Relate to Scientists Positively Shifts after Scientist Spotlight Assignments, Especially for First-Generation Students and Women

Kelsey J. Metzger,[†]* Molly Dingel,[†] and Ethan Brown[‡]
[†]Center for Learning Innovation, University of Minnesota Rochester, Rochester, MN 55904;
[‡]Research Methodology Consulting Center, College of Education and Human Development, University of Minnesota, St. Paul, MN 55108

ABSTRACT

We evaluate the impact of a low-stakes easy-to-implement course-level intervention, Scientist Spotlight assignments, which feature personal and professional stories of diverse scientists. This work extends previous studies by examining whether shifts in relatability differ across student identities, particularly students who identify as first-generation students, a population that has not been the focus of previous investigations of this intervention. Using paired pre- and postcourse data from four implementations in an introductory biology course, we report a significant, positive shift in undergraduate students' self-reported ability to relate to scientists, and concomitant shifts in how students describe scientists after completing four or six Scientist Spotlight assignments.

Importantly, our data demonstrate a disproportionate, positive shift for first-generation college students and for students who identify as female, a novel contribution to the body of literature investigating the Scientist Spotlight intervention. This study, along with previous reports of similar shifts in varying institutional contexts across different populations of learners, provides a strong argument that instructors interested in diversifying their course content to include representations of diverse scientists to enhance students' ability to identify a range of "types of people" who do science can do so successfully through incorporation of a small number of Spotlight assignments.

INTRODUCTION

Despite increasing racial diversity in the overall population of the United States and in higher education (National Center for Education Statistics, 2019; U.S. Census Bureau, 2020), alarming racial and economic class disparities remain in both pursuit of undergraduate science, technology, engineering, and mathematics (STEM) degrees and in educational outcomes (National Center on Safe Supportive Learning Environments, 2016; Fry *et al.*, 2021; Salehi *et al.*, 2021; Weatherton and Schussler, 2021). Students from certain historically marginalized groups (HMGs), including those who identify as Black, Latinx, and Native American, as well as female students and first-generation college students, depart from undergraduate STEM degree programs more frequently than their White, Asian-American, male, and continuing-generation counterparts (Huang *et al.*, 2000; Higher Education Research Institute, 2010; Ross *et al.*, 2012; Hatfield *et al.*, 2022). Multiple underlying causes contribute to perpetuating achievement and representation disparities, including structural and economic barriers such as access to high quality K–16 STEM education (National Academies of Sciences,

Sehoya Cotner, *Monitoring Editor*
Submitted Jun 8, 2022; Revised Dec 7, 2022; Accepted Dec 23, 2022
CBE Life Sci Educ March 1, 2023 22:ar12
DOI:10.1187/cbe.22-06-0103

*Address correspondence to: Kelsey J. Metzger (kmetzger@umn.edu).

Engineering, and Medicine 2021, 2022). However, even controlling for access issues, research affirms the importance of more local and affective factors such as a sense of belonging to departments and campuses (Seymour and Hewitt 1997; Hurtado and Carter 1997; Hurtado *et al.*, 2010; Murphy and Zirkel, 2015).

Exploring the affective domains of the student learning experience has a long history: Vincent Tinto (1975, 1997, 2000) pioneered the use of sense of belonging—students' academic and social integration into a school—as important for persistence in higher education, and other scholars have expanded on this framework (Hurtado and Carter, 1997; Braxton, 2000; Hoffman *et al.*, 2002; Hausmann *et al.*, 2007, 2009; Dingel and Sage, 2016). Collectively, this research demonstrates that sense of belonging positively correlates with students' success and persistence in higher education (Pascarella and Terenzini, 1980; Hurtado and Carter, 1997; Tinto, 1997, 2000; Braxton, 2000; Hoffman *et al.*, 2002; Hausmann *et al.*, 2007, 2009).

However, scholars have also critiqued the ways Tinto has written about sense of belonging with respect to students from HMGs. Indeed, a large body of research confirms that race and ethnicity fundamentally shape students' experiences in higher education (Gurin *et al.*, 2002; Swail *et al.*, 2003; Einarson and Matier, 2005; Yosso *et al.*, 2009; Ross *et al.*, 2012; Harper, 2015; Lo *et al.*, 2017; Harwood *et al.*, 2018; Masta, 2018; Johnson, 2019), including sense of belonging (Johnson, 2012; Trujillo and Tanner, 2014; Murphy and Zirkel, 2015).

Less research has been done on first-generation college students, but existing research also finds lower sense of belonging in this group (Stebleton *et al.*, 2014; Jantzer *et al.*, 2021), and that having a sense of belonging may be even more important for first-generation college students than continuing-generation students (Gillen-O'Neel, 2021). Research suggests that lower sense of belonging among first-generation college students in science departments is deeply tied to a mismatch in their perception of themselves as a "science person" (Chen *et al.*, 2021), as well as a cultural mismatch between their backgrounds and that of the department (Stephens *et al.*, 2012; Jackson *et al.*, 2016).

Of specific concern with Tinto's original work is an implication that students from HMGs must abandon important aspects of their cultural backgrounds and identities and adopt the mainstream culture in order to develop a sense of belonging to the institution/department (Tierney, 1992; Museus, 2011; Dingel and Sage, 2021). Sense of belonging is influenced by a wide range of factors, including students' relationships with faculty members (Miller *et al.*, 2019; Brooms, 2020; Park *et al.*, 2020), the curriculum (Dewsbury and Brame, 2019), and role models within a field (Drury *et al.*, 2011; Rosenthal *et al.*, 2013), among others. In turn, research has affirmed that, for undergraduate HMG STEM students, the campus and departmental environment influences sense of belonging and persistence (Hurtado, 2007; Taningco *et al.*, 2008; Malone and Barabino, 2009; Trujillo and Tanner, 2014; Winkle-Wagner and McCoy, 2018; O'Brien *et al.*, 2020; Park *et al.*, 2020). Taken together, this body of work provides strong evidence of the importance of cultivating a sense of belonging for students in the classroom and/or in the disciplinary field as an important factor for student persistence and success; this issue becomes more acute for students from HMGs.

A second concept that may work concomitant with sense of belonging in STEM undergraduate education is that of stereotype threat. First coined by Claude Steele (Steele and Aronson, 1995; Steele, 1997), stereotype threat affects a group about whom a negative stereotype exists in a specific situation; in such situations, these stereotypes can cause negative performance to emerge (Steele, 1997; Spencer *et al.*, 2016). Early studies describing stereotype threat included exposing a group to a relevant stereotype (e.g., women not being as good at math). These studies consistently found that the presence of the stereotype created a self-fulfilling prophecy. In other words, women who were told a math test often returned gender-disparate results performed worse than those not given such information (Spencer *et al.*, 1999). Further, explicitly undermining the stereotype caused the differences to disappear—for example, when women were told that a math test did not show gender differences, women scored equally to men (Spencer *et al.*, 1999). Similarly, in exploring the role of the stereotype about Black people's intellectual abilities, Steele and Aronson (1995) found that Black students given a verbal exam that they were told measures intellectual ability performed worse than Black students given the same exam but told it was a problem-solving task unrelated to ability. This work demonstrates the power of stereotypes but also reveals that stereotype threat may be reduced by actively undermining the negative stereotype.

Cultural influences and persistent widely held stereotypes about the identities and other personal characteristics of individuals who belong in or can be successful in science disciplines contribute to stereotype threat (Murphy *et al.*, 2007; Schmader and Beilock, 2012) and perpetuate barriers to establishing a sense of belonging and science identity for students from HMGs in STEM fields (Schinske *et al.*, 2015). As stated by Dewsbury and Brame, "The development of science identity is as crucial a part of the learning experience as the engagement of content" (2019, p. 4). Complicating this barrier, students from HMGs are less likely to see their identities reflected in faculty and administration in higher education or held up as examples of researchers and accomplished individuals, especially in STEM disciplines (Finkelstein *et al.*, 2016; Wood *et al.*, 2020; Simpson *et al.*, 2021). As of 2016, only 37.8% of holders of doctorates in science, engineering, and health were women (National Science Foundation, 2019). Additionally, individuals who identify as African American, Hispanic or Latino, and Native American or American Indian make up just 8.9% of science, engineering, and health doctorates. A recent study of demographic representation in undergraduate biology textbooks revealed an overwhelmingly disproportionate overrepresentation of white male scientists relative to the proportion of white males in the general population of the United States and in the biology student population, and a complete absence of any representation of women scientists of color (Wood *et al.*, 2020). Such a discrepancy can have a vicious feedforward effect: large swaths of students intending to pursue STEM degrees do not see themselves represented in these fields, they experience barriers that prohibit development of a sense of belonging in these academic spaces, and thus they are less likely to persist to degree attainment in STEM fields (Margolis *et al.*, 2000; Park *et al.*, 2020).

Much less work has been done on first-generation college students and stereotype threat (Harackiewicz *et al.*, 2014; Tibbetts *et al.*, 2016). However, consistent with stereotype

TABLE 1. Course-based pre- and postassessment prompts to evaluate impact of Scientist Spotlight assignments (adapted from Schinske *et al.*, 2016)

Assessment prompt	Response type	Instructions for students
Stereotypes prompt	Open-ended	One of the learning goals for this class is to examine the ways in which you think about the process of science and who contributes to scientific knowledge. Based on what you know now, describe the types of people that do science. If possible, refer to specific scientists and what they tell you about the types of people that do science
Relatability prompt	Closed-ended and open-ended	Please provide your opinion regarding the statement below: I know of one or more important scientists to whom I can personally relate 1 = agree 2 = somewhat agree 3 = somewhat disagree 4 = disagree 5 = I don't know Please explain your opinion of the statement: "I know of one or more important scientists to whom I can personally relate."

threat, the small body of existing research confirms the importance of identity toward integration into college departments (Herrmann *et al.*, 2018; Herrmann *et al.*, 2021; Chen *et al.*, 2020), and one study found that first generation-college students were more attentive than continuing-generation students to threatening cues, like comparing their performance with the performance of others (Jury *et al.*, 2015). One set of studies found that using process to help first-generation students affirm their values was successful in improving retention (Harackiewicz *et al.*, 2014; Tibbetts *et al.*, 2016), although another study found the context of the values affirmation to be important in its results (Bayly and Bumpus, 2019). In sum, while research suggests the presence of stereotype threat for first-generation students, there are fewer proposed interventions, and more work is needed in this area.

Curricular Interventions

A broad body of research in higher education indicates that interventions at various institutional, programmatic, curricular, and pedagogical levels can have significant positive effects on combating inequities in course performance and/or stereotype threat. Such interventions can also bolster students' sense of science identity (Schinske *et al.*, 2016; Hammarlund *et al.*, 2022) and sense of belonging (Binning *et al.*, 2020). Within the biology education research community specifically, there has been an increasing focus on the assessment of practices designed to promote the inclusion and sense of belonging of individuals with diverse identities (Schinske *et al.*, 2016; Dewsbury and Brame, 2019; Binning *et al.*, 2020; Hammarlund *et al.*, 2022) or to foster more equitable educational outcomes (Eddy and Hogan, 2014; Weatherton and Schussler, 2021), and to adopt approaches that inform systematic change toward more just institutions of higher education across all institution types (Salehi *et al.*, 2021).

The Current Intervention: Scientist Spotlight Assignments

Previous research (Schinske *et al.*, 2016) has demonstrated that the inclusion of 10 homework assignments featuring personal and professional stories of counterstereotypical scientists ("Scientist Spotlight" assignments) was associated with a positive shift in undergraduate biology students' perceptions of "who does science" and an enhanced sense of personally relating to scientists. The work done by Schinske and colleagues took place in an introductory biology class at a diverse community college in the southwestern United States designated as an Asian-American and Native American, Pacific Islander–serving institution; a total of 464 students were included. Both qualitative and quantitative approaches were used to investigate student stereotypes regarding scientists and students' ability to relate to scientists using pre- and postassessment prompts known as the "stereotypes prompt" and the "relatability prompt" (Table 1).

This initial study of the Scientist Spotlight intervention provided compelling data that shifts in student perceptions of stereotypes of scientists and scientist relatability were maintained 6 months after the intervention took place, but invited subsequent questions, including whether the results would be generalizable to other student populations and teaching contexts, and also whether or not the reported shift in student perspectives was equally distributed across student subpopulations or experienced unequally across student groups with differing identities.

Recent studies have contributed more insight into both questions: Yonas *et al.* (2020) assessed the impact of assigning nine Scientist Spotlight podcast assignments in the context of a large-enrollment undergraduate introductory biology course for nonmajors at a public R1 institution in the Midwest, with a total of 238 students included in the study, which the authors describe as 62% female, 26% non-white, 12% underrepresented minority (URM; defined as African American, Hispanic, Native American, Pacific Islander), and 17% first-generation college students (Yonas *et al.*, 2020, p. 8). Rather than providing an exact replicate study, this study sought to extend the results of Schinske *et al.* (2016) and to investigate "whether a diverse array of students could see their possible selves" by relating to scientists featured in the Spotlight assignments (Yonas *et al.*, 2020, p. 10). Through the inclusion of student demographics coupled with the open-ended responses provided by students following the Spotlight assignments, Yonas and colleagues concluded that not all students responded similarly to each of the scientists: women, LGBTQ students, and politically liberal students self-reported greater level of engagement with many of the podcasts (p. 9). Further, students' ability to relate to scientists was associated with their political perspectives (i.e., liberal, conservative, apolitical), and their majority/URM status. Overall, Yonas *et al.* (2020, p. 10) concluded that "Scientist Spotlight assignments … served to counter several stereotypes students held about scientists," and the authors encouraged the use of such spotlights by instructors seeking to make their courses more inclusive.

Brandt *et al.* (2020) also investigated students' perceptions of who does science and the number of scientists who students indicate they found relatable before and after the completion of eight video/podcast-based Scientist Spotlight assignments. The study took place at the same public R1 institution in the Midwest as Yonas *et al.* (2020), and students in the study were enrolled in either general zoology or a non-majors introductory biology course, with a total of 91 students included in the study. While the prompts presented to students and analyses conducted by Brandt and colleagues did not exactly replicate the approach of Schinske *et al.* (2016), these authors also concluded that the incorporation of Scientist Spotlight assignments increased students' ability to relate to scientists for both male and female students.

While Yonas *et al.* (2020) and Brandt *et al.* (2020) served to provide growing evidence that the Scientist Spotlight intervention was indeed generalizable across multiple student populations and teaching contexts, little continued to be known about how the intervention was impacting demographic subpopulations of students, inviting further investigation of this inclusive teaching pedagogy.

Aranda *et al.* (2021) investigated a modified implementation of the Scientist Spotlight intervention, in which upper-division undergraduate students participated in the creation of new Spotlights to be implemented in biology courses that included introductory and upper-division and majors and non-majors biology courses. The study population was 70% female, 41.5% URM, and 44% Pell Grant eligible. The study included paired pre- and postassessment responses from 752 students, used the same pre- and postassessment prompts as those originally used in Schinske *et al.* (2016), and included analyses of how these responses differed across different demographic student subpopulations after completion of three or four Spotlight assignments.

Aranda *et al.* (2021) reported two particularly interesting findings regarding the efficacy of the Scientist Spotlight intervention at the level of student subpopulations: first, there were no statistically significant differences in the preassessment baseline responses to the relatability prompt between any subpopulations; and second, there were statistically significant increases in students' ability to relate to scientists across all demographic subpopulations included in the study: men and women, URM and non-URM, and Pell Grant–eligible and non–Pell Grant eligible (Aranda *et al.*, 2021, p. 7). Similar results were obtained for the qualitative assessment of shifts in students' descriptions of stereotypes regarding scientists. The results reported by Aranda and colleagues suggest that the Scientist Spotlight intervention has a beneficial impact on students' ability to relate to scientists and perceptions of who does science for all students in the population included in this study, regardless of demographic differences.

The results of the Aranda *et al.* (2021) study also provide two additional important contributions for our understanding of the efficacy of the Scientist Spotlight intervention: first, Spotlight assignments authored by upper-division students can be used effectively in the place of instructor-authored Spotlights for this intervention (and additionally provide positive benefits for the student authors); and second, significant shifts in students' relatability to scientists and stereotypes regarding scientists can be accomplished with fewer assignments than had been previously demonstrated.

The research presented in this paper seeks to extend the results of previous work by investigating whether similar shifts in student perceptions can be achieved by incorporating a reduced number of instructor-authored Scientist Spotlight assignments during a one-semester undergraduate biology class at a predominantly white undergraduate degree–granting institution in the Midwest and to more closely examine the role of student demographics in students' responses to the relatability prompt. It is of value to investigate whether fewer non–student authored Spotlight assignments result in a shift in students' perceptions, as many instructors, particularly of foundational courses, may feel that their courses are already quite "full" with nonnegotiable disciplinary content areas or assignments and would not readily be able to incorporate a large number of additional assignments. Moreover, investigating student perceptions before and after equity-focused interventions at different institution types and in student populations with differing demographic compositions will enable broader understanding of the efficacy and limitations of such interventions.

Paper Overview

We used quantitative approaches to measure students' self-reported scientist relatability at the start and end of term to determine whether students' perspectives changed after completing course work that included personal and professional stories of counterstereotypical scientists. We examined potential shifts in students' self-reported relatability at the level of the whole population and also at the level of student subpopulation to determine whether there were differences in the ways in which students of differing identities and academic performance profiles responded to the intervention. In addition to the quantitative investigation, we used qualitative approaches to investigate the themes present in student open-ended responses regarding scientists at the start and end of term to determine whether the ways in which students describe the people who do science shifted after completing course work that included personal and professional stories of counterstereotypical scientists.

Research Questions

To investigate the impacts of Scientist Spotlight assignments with a novel population of undergraduate health sciences students at a predominantly white institution, the present study addressed the following research questions:

1. How, if at all, do students' self-reported levels of relating to scientists shift after exposure to Scientist Spotlight assignments that are not accompanied by additional classroom instruction?
2. How, if at all, do student demographics or attributes predict relatability to scientists, either at the start or end of the course?
3. How, if at all, do students completing four versus six Scientist Spotlight assignments report different levels of postcourse relatability?

4. How do students describe the types of people who do science before and after the Scientist Spotlight assignments?
5. How do students describe the ways in which they relate (or do not relate) to scientists before and after the Scientist Spotlight assignments?

The first three research questions were investigated through quantitative analyses, while the last two research questions were investigated through qualitative analyses of student responses.

METHODS

This research took place in the context of an undergraduate biology class at a small health sciences–focused undergraduate institution in the Midwest. The student population of the institution is skewed female (77%), white (61%), and originating from in-state (85%). Institutionally, at least 36% of students identify as first-generation college students, while 25% identify as continuing generation. For the remainder of students, their first-generation/continuing-generation status is unknown. The overall student population of the institution grew across all years of the study, from a total student population of 435 in AY2016–17 to 572 in AY2019–20. All students whose data were included in this study consented to be included in the research, and the research was approved by the University of Minnesota Institutional Review Board (protocol no. 0908S71602).

Scientist Spotlight assignments were used in four consecutive annual offerings of a 16-week required introductory biology course taken primarily by first-year health science majors during Spring term. The course was taught by the same instructor for all four offerings, and the only major change in instruction occurred in the Spring of 2020 when the course began with on-campus, in-person instruction and moved to remote, synchronous online instruction for the last two months of the term (approximately March 15–May 15, 2020). The course was taught using a reduced lecture, flipped, and student-centered framework in which students engaged in preclass preparation before attending class meetings and collaborative activities, discussion, and other active-learning strategies were the primary modes of engagement during class meeting times. Instruction took place in active-learning classrooms and labs during in-person instruction.

Four Scientist Spotlight assignments were incorporated in two offerings of the course (Spring 2017, Spring 2018), and six Scientist Spotlight assignments were incorporated during two offerings (Spring 2019, Spring 2020). Beyond the Scientist Spotlight assignments, no explicit instruction regarding scientist stereotypes took place in these classes. At the beginning and end of the class, students were asked to provide responses to two items addressing students' conceptions of scientists ("stereotypes prompt") and students' perceived ability to relate to a scientist ("relatability prompt," consisting of quantitative and qualitative components; Table 1). These items have previously been used to investigate the efficacy of Scientist Spotlight assignments (Schinske *et al.*, 2015, 2016; Aranda *et al.*, 2021). Thus, the experimental design was a pre–post study, repeated across four different implementations of the same course with four different cohorts of students in years 2017, 2018, 2019, and 2020.

Quantitative Analysis

Using the quantitative data resulting from the Likert-scale relatability prompt (Table 1), we were interested in addressing three research related questions:

1. How, if at all, do students' self-reported levels of relating to scientists shift after exposure to Scientist Spotlight assignments that are not accompanied by additional classroom instruction?
2. How, if at all, do student demographics or attributes predict relatability to scientists, either:
   a. at the start of the course?
   b. at the end of the course?
3. How, if at all, do students completing four versus six Scientist Spotlight assignments report different levels of post-course relatability?

As an initial descriptive examination of the quantitative research questions, we performed ordinary least-squares linear regressions to facilitate ease of practical interpretation. This approach uses the general linear model and treats relatability as a ratio variable and is thus comparable to the analyses of covariance in Schinske *et al.* (2016); linear regression allowed us to take a more explicit focus on estimated values of predictors and to compare the amount of variance explained (R^2) between different groups of variables. Note that relatability was a single item measured on a four-point response scale and is therefore ordinal, which violates the assumptions of linear regression. We therefore performed follow-up sensitivity analysis using ordinal logistic regression to determine whether this led to similar conclusions as the linear regression. The Likert-scaled relatability prompt was the outcome variable for all analyses. After much deliberation, we chose to replicate the approach taken by Schinske *et al.* (2016) and Aranda *et al.* (2021) to treat participants who responded "I don't know" as "Disagree" for the quantitative analyses, as the prompt was "I know of one or more important scientists to whom I can personally relate" and those who responded "I don't know" appeared to not know one or more important scientists to whom they could relate given the rationale provided in the paired constructed-response item in which students were asked to explain their quantitative responses.

To investigate the extent to which self-reported levels of relating to scientists shift after exposure to Scientist Spotlight assignments (research question 1), we performed a paired *t* test comparing the pre and post relatability scores, interpreting them as interval data similar to Schinske *et al.* (2016). For comparison with the results from Aranda *et al.* (2021) that examined the relatability data as dichotomous, we also performed a Fisher's exact test of the pre and post relatability scores.

To investigate how subgroup membership affected self-reported relatability scores obtained at the beginning of the course (research question 2a), we fit a series of linear regression models with pretest relatability scores as the outcome, treated as a continuous variable. The first model examined the predictive capacity of cohort year; the second model added demographic variables: student of color (yes/no), gender (male/female), first-generation student (yes/no); and the third model added academic achievement variables: ACT composite score and ACT math score.

To investigate how subgroup membership affected the *change* in self-reported relatability (research question 2b), we fit a series of linear regression models with posttest scores as the outcome. The first model examined the predictive capacity of the pretest score, dummy-coded; the second model added the predictive capacity of cohort year (research question 3); the third model added demographic variables: student of color (yes/no); gender (male/female); first -generation student (yes/no); the fourth model added mean-centered academic achievement variables: course grade in percentage points, ACT composite score, and ACT math score. Mean-centering was done for the interpretation of the intercept but did not change statistical or practical conclusions.

Pell Grant eligibility status was available for three of the four cohort years. However, analysis of this variable is not reported in detail here, because data for this variable were not available for all 4 years. Moreover, Pell eligibility status did not show significant predictive power in the models or change any conclusions of other variables when we examined the regression results that included this variable for the three available years only.

Qualitative Analysis

To investigate research question 4 and research question 5, constructed-response questions were asked of students: a stereotypes prompt "Based on what you know now, describe the types of people that do science. If possible, refer to specific scientists and what they tell you about the types of people that do science" and an explanatory prompt regarding the Likert-scaled relatability prompt "Please explain your opinion of the statement: 'I know of one or more important scientists to whom I can personally relate'" (Table 1). To characterize the themes present in student pre- and postcourse responses to these constructed-response items, an iterative approach to coding was implemented, informed by the thematic codes reported in previous research (Schinske *et al.*, 2016). The codes used in Schinske *et al.* (2016) were informed by themes reported in Dikmenli (2010) and Mead and Metraux (1957). Regarding stereotypes, the two main coding categories were Stereotypes and Non-stereotypes. Subcategories of Stereotypes were Positive Stereotypes, Negative Stereotypes, and Stereotypical Scientists. A third-level subcategory was identified for the subcategory Negative Stereotypes, in which students Refute Negative Stereotypes. Subcategories of Non-Stereotypes were Non-stereotypical Descriptions, and Non-stereotypical Scientists. As with previous research, key words that distinguished positive stereotypes included descriptions of scientists as exceptionally smart, dedicated, passionate, curious, motivated, and hardworking. Negative stereotypes included descriptions of scientists as socially awkward, nerdy, selfish, introverted, greedy, crazy, isolated, elitist, and egotistical. Table 2 provides examples of student responses for each of these categories and subcategories.

To distinguish which named scientists would be categorized as Stereotypical Scientists or Non-stereotypical Scientists in student responses, we used a previously established list of stereotypical scientists (Mead and Metraux, 1957; Dikmenli, 2010; Schinske *et al.*, 2015), which was been used in previous studies assessing Scientist Spotlight interventions (Schinske *et al.*, 2015; Aranda *et al.*, 2021). Table 3 provides the list of stereotypical scientists and the original source in the literature.

Additional major coding categories were Field of Science, and Who Does Science. Subcategories of Who Does Science were Everybody Does Science, Many Different Types of People Do Science, and Normal People Do Science. Responses that indicated fields of study in science (biology, chemistry, physics, etc.), rather than describing characteristics of scientists themselves, were coded as Field of Study. Although not used in previous Scientist Spotlight research, due to frequency of prevalence, we added a coding category regarding Who Does Science, using the key words/phrases "everybody," "different kinds of people," and "normal people" as subcategories to identify responses in these subcategories (Table 2). While an argument could be made to include Everybody Does Science, Many Different Types of People Do Science, and Normal People Do Science as Non-stereotypical descriptions, we felt that these student descriptions represented a sentiment that went beyond a position of contradicting canonical stereotypes and embraced an inclusive perspective that merited further granularity in our analysis and consideration of the student thought process.

After initial coding (K.J.M., M.D.), researchers compared a subset of results and discussed discrepancies in how student responses were coded to clarify and refine application of the codebook themes. After two coders had completed the coding of all student responses (pre and post, all 4 years) with the revised codebook, interrater reliability was calculated for all thematic categories using both percent agreement and Cohen's kappa (McHugh, 2012) applied to presence/absence coding in a selection of 240 coded student responses (~25% of all responses). Interrater reliability (McHugh, 2012) resulted in percent agreement of greater than 90% and a Cohen's kappa above 0.8 for all categories, indicating strong level of agreement between coders.

RESULTS

Student Attribute Differences by Year

As we were interested in comparing shifts in students' self-reported relatability across years with different treatments (four assignments vs. six assignments), we investigated potential differences in student attributes, such as academic course performance and demographic characteristics (e.g., first-generation status; gender; student of color status) across years. Academic performance indicated by course grade (out of 100%) showed evidence of differing by year, $R^2 = 0.06$, $F(3, 169) = 9.13$, $p < 0.001$. This appeared to be driven by the lower average grade in 2020, 82.6, compared with 84.3 in 2017 and about 86 in 2018 and 2019. First-generation college student status also showed evidence of differences by year, $\chi^2(6) = 25.84$, $p < 0.001$. This was driven both by differences in non-response and self-reported status across years; non-response steadily decreased from 37% of the 2017 participants to 26% of 2020 participants. Of those who responded, percentages of first-generation college students varied from 48% in 2017 to 83% in 2020. There was little evidence of other demographics varying across years (p values > 0.20).

Overall Pre–Post Comparisons in Relatability Scores

To investigate research question 1, we examined overall changes in precourse to postcourse relatability scores for all students and all years of our study. This overall examination of pre- to postcourse changes enabled us to compare our findings with those of prior studies that took place in different institutional

TABLE 2. Qualitative code categories and examples of student responses to Stereotypes prompt

Code	Subcode	Example student responses
Stereotype	Negative Stereotype	• "Based on what I know, I would say that the type of people that normally do science are very smart intellectually but may not have the best social skills." • "People who are interested in studying science are commonly referred to as nerds." • "I think that most scientists do science to help others and some just try to create profit."
Stereotype	Positive Stereotype	• "Most of the people who do science are really smart." • "I believe the types of people who do science are the people who are willingly dedicating their time to fix, evolve, or try to have a better understanding of issues, health, or diseases in today's world." • "People who do science are dedicated. They will do whatever it takes to solve the problems they face."
Stereotype	Refute Negative Stereotype	• "People who do science aren't all lab coats and mixing chemicals. Sometimes that's what it is, but sometimes it's carrying your poop through an airport and having to explain to security why." • "I learned anyone can be a scientist if they commit to it. The scientist [*sic*] also do not have to be super-geniuses and are normal people." • "I learned that there's a huge stereotype around scientist of being bland, male, "nerdy", social[ly] awkward, etc. However, that's not the case at all. Scientists are passionate, funny, sociable, can be any gender, and most importantly insightful."
Non-stereotype	N/A	"There is a wide range of people who do science. They can be people from any gender, ethnicity, and upbringing. People who grew up poor, rich, or anything in between can do science. People who do science aren't just studious people either. They are passionate individuals, some even with a sense of humor."
Who Does Science	Science Is Not for Everybody	"Science is a typically male dominated field where white people take charge." "I feel that science is something that not everyone can go into, especially medical science." "Science normally does not come easy for most people and you have to really try to learn and understand the concepts."
Who Does Science	Everybody Does Science	• "Any person can and will do science. It doesn't matter what type of person you are, where you come from, what you're like, or your personality. Anyone can do science and it doesn't matter who you are, as long as you have a passion for it." • "Anyone can do science! Literally anybody. Kids can do science, people from disadvantaged backgrounds can do science, rich people can do science, anyone can do science. If you are passionate you can do science."
Who Does Science	Many Different Types of People Do Science	• "There are so many people that do science, all you really need is a passion and interest for it. Each of the scientists we learned about came from all different backgrounds, some from no family support system, some from poor neighborhoods and some with a good education and good support systems."
Who Does Science	Normal People Do Science	• "I think that after learning more about the types of people who do science I have come to realize they are normal people who are devoted to helping people. They are often times very curious and will do anything to solve a problem or to understand something they are working on." • "The types of people who do science can be anyone of any age, race, gender, etc. There was such a wide variety of scientists we looked at, and each one came from different backgrounds and places all over. *People who do science can be ordinary people*, and I think a lot of people forget that."
Field of Science	N/A	• "Researchers, chemists, physicists, sociologists, doctors." • "There are geneticists who study how traits are inherited, epidemiologists who study the spread of diseases, geographers and geologists who study the Earth, marine biologists who study ocean plants and animals, a paleontologist who studies fossils, and a microbiologist who studies microscopic plants and animals" • "I think people that do science include physicians, scientists, engineers.... etc..."

contexts and within different populations of learners. Our results replicated the finding reported by Schinske *et al.* (2016) with strong evidence of increase from precourse to postcourse in relatability scores when interpreted as interval data (by comparing means), $\Delta M = 0.85$, $t(280) = 14$, $p < 0.001$, 95% confidence interval (CI) = [0.73, 0.96]. This corresponded with a Cohen's d of 0.80, 95% CI = [0.66, 0.93], representing a large effect size (Lakens, 2013; Cohen, 1988). Our findings also replicated the finding of Aranda *et al.* (2021) of pre–post differences when relatability is interpreted as a dichotomous categorical variable: Across all years, the percentage of students who responded that they agreed or somewhat agreed with the Relatability prompt at the start of term was a 34%, which increased to 70% at the end of term, odds ratio (OR) = 3.03, Fisher's exact $p < 0.001$, 95% CI = [1.81, 5.24].

The overall percentage of students shifting from either "disagree" or "I don't know" at precourse to "agree" or "somewhat agree" at postcourse (average frequency 42% across all years) is comparable to the percentage reported in previous studies (36% reported in Aranda *et al.*, 2021). Interestingly, the percentage of students who responded "agree" or "somewhat agree" to the Relatability prompt at the start of term was

TABLE 3. Compiled list of stereotypical scientists

Source	Stereotypical scientists
Schinske *et al.*, (2015), Table 2	Thomas Edison Albert Einstein Benjamin Franklin Sigmund Freud Galileo Galilei Leonardo da Vinci James Watson
Dikmenli (2010), Table 2	Archimedes Alexander Graham Bell Francis Crick Charles Darwin Thomas Edison Albert Einstein Antonie van Leeuwenhoek Gregor Mendel Isaac Newton Louis Pasteur Oktay Sinanoğu Ibn Sinna [Avicenna] James Watson Stephen Hawking
Mead and Metraux (1957), p. 387	Marie Curie[a] Albert Einstein J. Robert Oppenheimer Jonas Salk

[a]Note: Marie Curie is listed alongside Einstein, Oppenheimer, and Salk in Mead and Metraux's 1957 examination of how high school students view scientists. However, the inclusion of this single female scientist is in stark contrast to the language of the paper, which uses the masculine pronoun "he" exclusively throughout the paper to describe an individual who does science: "He is a dedicated man who works not for money or fame or self-glory—but, like Madam Curie, Einstein, Oppenheimer, Salk—for the benefit of man-kind and the welfare of his country." (Mead and Metraux, 1957, p. 387). Schinske *et al.* (2015, Table 3) categorized Marie Curie as a non-stereotypical scientist, and we have continued this categorization in the present study.

highest in 2017, the cohort with the lowest percentage of first-generation students (Table 4 and (Figure 1)). The constructed responses provided by students in the follow-up item prompting students to explain their quantitative relatability response add context regarding the rationales used by students in selecting their responses. Table 5 provides examples of paired student pre- and postcourse open-ended responses in each Relatability shift category.

Linear Regression Models

In addition to examining pre- to postcourse changes globally in the whole study population, we sought to investigate whether there were differences in self-reported relatability among subpopulations of students at both the pre- and postcourse time points. Our further analyses extend prior findings by examining how student demographic and academic performance variables jointly predict relatability scores through a series of linear regression models, treating the relatability scores as interval data following Schinske *et al.* (2016).

Regression Models Predicting Pre-intervention Relatability Scores

To investigate research question 2a, linear regression models were used to investigate whether subgroup memberships predicted relatability scores at the beginning of the course–-before the Scientist Spotlight intervention. The first model examined the predictive capacity of cohort year; the second model added demographic variables: student of color (yes/no), gender (male/female), first-generation student (yes/no); and the third model added academic preparation variables: ACT composite score and ACT math score.

Linear regression results using the precourse relatability scores are shown in Table 6. The results of our linear regression models found that subgroup membership did not explain much of the variation in pre-intervention relatability scores overall. However, we did find evidence of differences in precourse relatability across our four cohorts, evidence of male–female differences, and evidence that first-generation student status predicted lower relatability at the start of the term after accounting for other subgroup variables.

Differences in Precourse Relatability Scores across Different Cohort Years of Study. Model 1, with cohort year alone, accounted for about 4% of the variance in precourse relatability scores, $\Delta R^2 = 0.04$, $F(3, 435) = 6.09$, $p < 0.001$. There was evidence of a difference between 2017, the year with the highest precourse relatability scores, and 2018, the year with the lowest, $\beta = 0.55$, $t = 3.83$, $p < 0.001$, 95% CI = [0.27, 0.84], but little evidence of differences between 2018, 2019, and 2020.

Differences in Precourse Relatability Scores across Demographic Subgroups. Model 2 included demographic variables, and the percentage of variance accounted for increased to 8%,

TABLE 4. Frequency (percentage) of types of shifts in students' pre-to-post Relatability prompt responses

Students who…[a]	2017 N = 101	2018 N = 135	2019 N = 116	2020 N = 100	All years N = 452
Always disagreed	4%	16%	4%	7%	8%
Shifted from IDK to disagree	6%	5%	10%	7%	7%
Shifted from agree to disagree	5%	4%	3%	5%	4%
Shifted from disagree to IDK	2%	3%	2%	5%	3%
Always IDK	1%	7%	8%	7%	6%
Shifted from agree to IDK	0%	1%	1%	3%	1%
Shifted from disagree to agree	20%	21%	27%	30%	24%
Shifted from IDK to agree	18%	18%	21%	15%	18%
Always agreed	45%	24%	25%	21%	28%

[a]Students who responded "agree" or "somewhat agree" to the relatability prompt were combined into the category "agree," while students who responded "disagree" or "somewhat disagree" to the relatability prompt were combined into the category "disagree." IDK, I don't know.

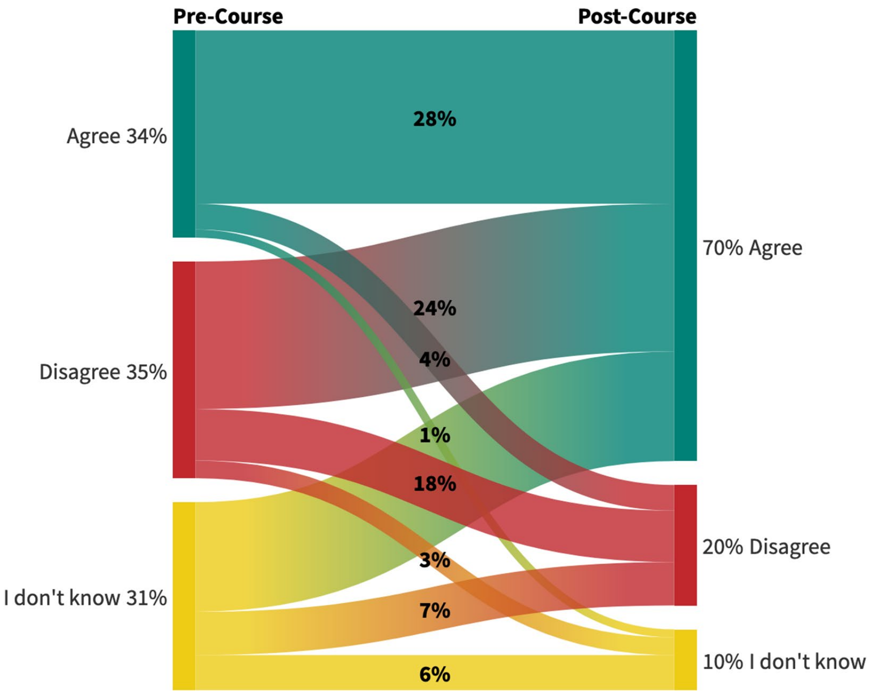

FIGURE 1. Shifts in Relatability prompt responses pre- to postcourse for all years. Students who responded "agree" or "somewhat agree" to the relatability prompt were combined into the category "agree," while students who responded "disagree" or "somewhat disagree" to the relatability prompt were combined into the category "disagree."

$\Delta R^2 = 0.04$, $F(4, 431) = 4.95$, $p < 0.001$. Students who either self-reported first-generation status ($\beta = -0.46$, $t = -3.26$, $p = 0.001$, 95% CI = [–0.74, –0.18]) or who declined to identify their first-generation status ($\beta = -0.42$, $t = -2.83$, $p = 0.005$, 95% CI = [–0.71, –0.13]) were predicted to have about a half-point lower precourse relatability score than students who did not self-report identifying as first-generation students. Male-identified students had a predicted pretest relatability score of a quarter-point higher than female-identified students, though the statistical evidence was fairly weak, $\beta = 0.26$, $t = 2.00$, $p = 0.046$, 95% CI = [0, 0.51].

Differences in Precourse Relatability Scores Predicted by Academic Preparation. Model 3 added ACT comprehensive scores and ACT math scores to assess the contributions of prior academic ability. These ACT variables did not appear to predict precourse relatability scores beyond the variables in prior models, $\Delta R^2 = 0.00$, $F(2, 429) = 1.17$, $p = 0.311$.

Regression Models Predicting Post-intervention Relatability Scores

To investigate how subgroup membership affected the *change* in self-reported relatability from precourse to postcourse (research question 2b), we fit a second series of linear regression models with posttest scores as the outcome and precourse relatability scores and demographic and academic variables as predictors. The first model examined the predictive capacity of the pretest score, dummy-coded; the second model added the predictive capacity of cohort year (research question 3); the third model added demographic variables: student of color (yes/no), gender (male/female), first -generation student (yes/no); and the fourth model added mean-centered academic achievement and academic preparation variables: course grade in percentage points, ACT composite score, ACT math score. These models assessed the degree to which demographic and academic variables predicted the change in relatability scores (research question 2b), as postcourse relatability scores were corrected for precourse score differences by their inclusion in the models. Further, model 2 which included cohort year as a predictive variable, enabled us to examine the impact of the number of Scientist Spotlight assignments (research question 3), because the number of assignments differed from four to six in different offerings of the course.

Linear regression results for postcourse relatability scores are shown in Table 7. Overall, we did not find substantial contributions of any demographic or academic variables to the change in relatability.

Precourse Relatability as a Predictor of Postcourse Relatability Shift. Model 1, with only precourse scores included as a categorical predictor, accounted for about 12% of the variability in postcourse scores, $\Delta R^2 = 0.12$, $F(3, 428) = 19.29$, $p < 0.001$. Participants who responded "somewhat agree" on precourse relatability received predicted postcourse relatability scores about a half-point higher on the 1–4 scale than those who responded "disagree," $\beta = 0.57$, $t = 4.87$, $p < 0.001$, 95% CI = [0.34, 0.81], which nearly doubled to a full point for participants responding "agree" precourse, $\beta = 1.00$, $t = 6.8$, $p < 0.001$, 95% CI = [0.71, 1.29].

Number of Scientist Spotlights as a Predictor of Postcourse Relatability Shift. Model 2 of this series of linear regression analysis enables us to evaluate whether the number of Scientist Spotlight assignments (four or six) is associated with a significant difference in the extent to which students' relatability response shifts from pre- to postcourse through the addition of the cohort variable (research question 3). Model 2 shows that there was little evidence of any cohort year effects in postcourse scores after controlling for precourse scores, $\Delta R^2 = 0.01$, $F(3, 425) = 1.39$, $p = 0.246$.

Demographic Variables as Predictors of Postcourse Relatability Shift. Model 3 found little evidence of a predictive effect of being a student of color, first-generation status, or gender for the change in relatability from pre- to postassessment.

Academic Preparation and Performance Variables as Predictors of Postcourse Relatability Shift. Model 4 found little evidence of academic variables such as course grade or ACT scores being predictive. Model 4, with all subgroup predictors, only

TABLE 5. Examples of paired pre–post student responses to open-ended Relatability prompt

Students who…[a]	Pre	Post	Students[b]
Always disagreed	"I have heard of many scientists, but there aren't any that I personally relate to, mainly because I do not know enough about one individual enough to be able to say that I relate to them."	"I know the scientists from scientist spotlights, but there is none that I can personally relate to. I think each of their stories is very unique and interesting, but there was no story that stood out to me as something I have experienced."	Kaitlyn White female, FGEN
Shifted from IDK to disagree	"I chose I don't know since I really am unsure of what important scientist I can personally relate to."	"I don't personally know any scientist."	Kao Asian female, FGEN, PELL
Shifted from agree to disagree	"I can relate to past and present teachers and professors who I look up to and aspire to have their knowledge of the subject."	"I don't feel like I know any important scientists that I can personally relate to."	Sam American Indian, male
Shifted from disagree to IDK	"I don't know any important scientists that I can relate my life, motivations, and morals to. I know there are well known scientists and their major accomplishments, but I can not personally relate to them."	"Even though we went over many scientists throughout this semester, there wasn't a certain scientists [*sic*] that I can actually relate to and look up to as a role model. Everyone has there [*sic*] own life experiences to get where they are at now but not one connected to be as much as I hoped so."	Jasmine Asian female, FGEN, PELL
Always IDK	"I don't personally know any scientists. I know the names of some well known scientists and the broad area of what they're doing research on and studying about. But, I can't for sure [*sic*] that I can personally relate to any of them."	"All of the scientists in the spotlights were interesting. However I can only relate to the fact that we are all passionate about the STEM field."	Savannah White female, FGEN, PELL
Shifted from agree to IDK	"I said somewhat agree because I know of scientists who I understand, but I wouldn't really say I can personally relate with them."	"I don't know because I don't really know a scientist who I can relate with, Maybe in the near future I will find someone."	Amy Asian female, FGEN, PELL
Shifted from disagree to agree	"As of now I am not aware of many scientists because in the past I have taken biology classes but rarely focused on the scientists themselves or it just didn't seem as important at the time. However, I am willing to learn more about what they do and who they are."	"I can relate because many of them struggled to where they are now and even though I am not where I want to ultimately be, I can relate in my struggles too."	Mariana Hispanic female, FGEN
Shifted from IDK to agree	"I have learned about a lot of different scientist[s] throughout my high school years and last semester, however I cannot personally relate to them because I never had to really connect with them. All I was required to do was understand theories that they came up with or concepts that we now have in science because of them. Other than that, I did not learn much about the scientist lives and what they did that lead [*sic*] them to their discoveries or what consequence they faced in life due to this. These are things that I feel I need to know before I have a connection to a scientist."	"I felt like I could relate to most of the women we did scientific spotlights on which is why I know one or more important scientist I can relate to otherwise I don't because aside from this project most of the scientist we learn about throughout are educations are white males and since I am neither white nor male it's hard for me to personally relate to them."	Faduma Black female, PELL
Always agreed	"I feel that I know a good amount of older scientists who laid the groundwork for what we know now about science. I feel that I can relate to them based on their curiosity and my curiosity to learn."	"Many of the scientists we talked about in class are very relatable because they all came from struggles and were able to follow their dream the best."	Tyler White Male, CGEN

[a]IDK, I don't know.
[b]FGEN, first-generation college student; CGEN, continuing-generation college student; PELL, Pell Grant–eligible student.

explained 3% more variation than model 1, with precourse relatability scores alone, $\Delta R^2 = 0.03$, $F(10, 418) = 1.71$, $p = 0.076$.

Excluding Precourse Scores as a Predictive Variable of Postcourse Relatability Shift. To gain more insight into the null findings of the variables in models 2–4, we used model 5 to

TABLE 6. Linear regressions predicting pretest relatability scores[a]

	Dependent variable: Relatability: Pre		
	(1)	(2)	(3)
Year: 2017	0.55***	0.51***	0.51***
	(0.27, 0.84)	(0.22, 0.79)	(0.23, 0.80)
Year: 2019	0.01	−0.02	−0.03
	(−0.27, 0.28)	(−0.29, 0.25)	(−0.30, 0.24)
Year: 2020	0.09	0.11	0.11
	(−0.20, 0.37)	(−0.17, 0.40)	(−0.17, 0.39)
Student of color		0.16	0.13
		(−0.07, 0.38)	(−0.10, 0.36)
Male		0.26**	0.30**
		(0.01, 0.51)	(0.04, 0.56)
First generation: unknown		−0.40***	−0.42***
		(−0.69, −0.12)	(−0.71, −0.13)
First generation: yes		−0.43***	−0.46***
		(−0.69, −0.16)	(−0.73, −0.18)
Composite ACT			0.01
			(−0.04, 0.06)
ACT Math			−0.03
			(−0.07, 0.01)
Constant	1.78***	2.02***	2.04***
	(1.60, 1.97)	(1.73, 2.30)	(1.74, 2.33)
Observations	439	439	439
R^2	0.04	0.08	0.09

[a]Model 1 contains cohort year only; model 2 adds demographic variables; model 3 adds ACT scores as a proxy for prior academic ability.
* $p < 0.10$.
**$p < 0.05$.
***$p < 0.01$.

further examine how these same variables predicted postcourse scores in a model that *excluded* the precourse scores. The estimates of subgroup variables were similar between model 4 and model 5. In particular, there was still little evidence that first-generation students performed differently on the postcourse than non–first generation students, $\beta = -0.10$, $t = -0.76$, $p = 0.446$, 95% CI = [−0.36, 0.16], even though first-generation status had predicted precourse scores, and in turn precourse scores predict postcourse scores. If there is a mediated relationship between first-generation status, precourse scores, and postcourse scores, it may be obscured by the low levels of explained variation throughout all models—only 9% for the precourse relatability models and 15% for the postcourse relatability models.

Qualitative Analysis Results: Stereotypes and Stereotypical Scientists

Research question 4 sought to characterize how undergraduate students describe the types of people who do science before and after the Scientist Spotlight assignments. In examining the pre–post change in frequency of thematic codes averaged across all years, the categories with the largest change are Field of Science and Different Kinds of People Do Science. The theme Field of Science, for which students name a particular field of science, exhibits a decrease from appearing in nearly 35% of all student responses at the beginning of the term to appearing in only approximately 7% of student responses at the end of term. Conversely, the theme Different Kinds of People Do Science initially appears in 7% of student responses, but increases to appear, on average, in 30% of all end of term student responses across all 4 years of the study. Supplemental Table 1 reports average frequency for qualitative codes across years, as well as the change in frequency from pre- to postcourse.

Student Callouts of Stereotypical Scientists in Precourse and Postcourse Responses

In their responses, some students listed multiple scientists and other students listed no scientists. As stated previously, to distinguish which named scientists would be categorized as stereotypical scientists or non-stereotypical scientists in student responses, we used a previously established list of stereotypical scientists (Table 3) that has been used in previous studies assessing Scientist Spotlight interventions (Schinske *et al.*, 2015; Aranda *et al.*, 2021). Using that established reference list, stereotypical scientists were named in nearly 20% of all student responses at the start of term; by the end of term, names of stereotypical scientists appeared in fewer than 2% of all student responses across all years. Reciprocally, as the frequency of naming stereotypical scientists decreased, a concomitant increase in naming non-stereotypical scientists was observed from pre- to postassessment: At the start of term, students named non-stereotypical scientists at a frequency of approximately 12%, but at the end of term, the name of a non-stereotypical scientist appeared in 23.5% of all responses, representing a nearly twofold increase in frequency.

TABLE 7. Linear regressions predicting posttest relatability scores[a]

	Dependent variable: Relatability: Post				
	(1)	(2)	(3)	(4)	(5)
Relatability: Pre = somewhat disagree	0.35**	0.34**	0.31**	0.32**	
	(0.07, 0.62)	(0.06, 0.61)	(0.04, 0.59)	(0.04, 0.59)	
Relatability: Pre = somewhat agree	0.57***	0.55***	0.56***	0.57***	
	(0.34, 0.80)	(0.32, 0.79)	(0.33, 0.80)	(0.33, 0.80)	
Relatability: Pre = agree	1.00***	0.94***	0.98***	0.95***	
	(0.71, 1.29)	(0.65, 1.24)	(0.68, 1.28)	(0.65, 1.25)	
Year: 2017		0.26**	0.24*	0.28**	0.45***
		(0.005, 0.51)	(−0.01, 0.50)	(0.02, 0.54)	(0.18, 0.72)
Year: 2019		0.12	0.11	0.12	0.12
		(−0.12, 0.36)	(−0.13, 0.36)	(−0.12, 0.37)	(−0.14, 0.37)
Year: 2020		0.08	0.08	0.13	0.17
		(−0.18, 0.33)	(−0.17, 0.34)	(−0.13, 0.39)	(−0.11, 0.44)
Student of color			0.002	0.03	0.08
			(−0.20, 0.20)	(−0.17, 0.24)	(−0.14, 0.30)
Male			−0.22*	−0.23*	−0.14
			(−0.45, 0.01)	(−0.46, 0.01)	(−0.39, 0.10)
First generation: unknown			−0.001	0.02	−0.10
			(−0.26, 0.26)	(−0.24, 0.29)	(−0.38, 0.17)
First generation: yes			−0.02	0.03	−0.10
			(−0.26, 0.23)	(−0.22, 0.28)	(−0.36, 0.16)
Course grade (out of 100)				0.02*	0.02**
				(−0.00, 0.03)	(0.00, 0.04)
ACT composite				0.04*	0.04*
				(−0.004, 0.08)	(−0.003, 0.09)
ACT math				−0.02	−0.03
				(−0.06, 0.02)	(−0.07, 0.01)
Constant	2.48***	2.39***	2.44***	2.38***	2.68***
	(2.36, 2.60)	(2.21, 2.57)	(2.16, 2.72)	(2.10, 2.66)	(2.40, 2.96)
Observations	432	432	432	432	432
R^2	0.12	0.13	0.13	0.15	0.05

[a]Model 1 contains precourse relatability score only; model 2 adds cohort year; model 3 adds demographic variables; model 4 adds ACT scores and course grades; model 5 removes precourse relatability scores to assess their impact on other coefficients.
*$p < 0.10$.
**$p < 0.05$.
***$p < 0.01$.

Table 8 reports the consistently most frequently named scientists in the precourse survey across all years in raw counts. There were eight scientists who were named consistently in each of the 4 years: four of the eight scientists frequently named in the precourse survey were Stereotypical Scientists (Charles Darwin, Albert Einstein, Isaac Newton, Thomas Edison), whereas four were not (Marie Curie, Neil deGrasse Tyson, Rosalind Franklin, Bill Nye). Perhaps not surprisingly for an undergraduate biology course, Charles Darwin was the most frequently named scientist in the precourse survey across all years.

Table 9 reports the most frequently named scientists in the postcourse survey across all years. In all cases, the most frequently named scientists in the postcourse survey were those scientists who had been featured in the Scientist Spotlight assignments; instances of the well-known stereotypical scientists who overwhelmingly dominated the precourse survey were virtually absent from the postcourse survey responses. Scientist Spotlight assignments were assigned on a schedule that aligned the research area of the scientist with the topical content area of the course, so the same order of assignments was used in each year (Francis Kelsey, Shinya Yamanaka, Agnes Day, Clare Feiseler in 2017 and 2018; Francis Kelsey, Shinya Yamanaka, Agnes Day, Lawrence David, Flossie Wong-Staal, and Juan Perilla, Clare Fieseler in 2019 and 2020). Interestingly, but perhaps not surprisingly, the scientist who was the focus of the most recently assigned Spotlight assignment (Clare Fieseler) was the most frequently named scientist on the postcourse survey in every year, with a corresponding pattern of recency and frequency observed for the other featured scientists.

Student Descriptions of Scientists in Precourse and Postcourse Responses

Although the overwhelming majority of student responses (>75% on average across all years) included descriptions of positive stereotypes associated with scientists (intelligent, curious, motivated, etc.) with very little change in frequency pre- to postassessment, the ways in which students discussed *negative*

TABLE 8. Scientists frequently named in precourse Stereotypes prompt across years

Scientist	2017	2018	2019	2020	Total
Charles Darwin	6	12	14	6	38
Albert Einstein	11	8	9	6	34
Isaac Newton	5	8	7	2	22
Neil deGrasse Tyson	2	3	1	5	11
Marie Curie	1	3	5	2	11
Rosalind Franklin	2	2	3	1	8
Bill Nye	1	4	1	2	8
Thomas Edison	3	2	1	1	7

stereotypes (nerdy, isolated, exclusive, etc.) did change in frequency from pre- to postcourse. At the start of term, student responses including negative stereotypes of scientists did so in a way that affirmed the negative stereotype. However, by the end of term, the majority of student responses that included reference to a negative stereotype associated with scientists did so to *refute* the stereotype rather than affirm its accuracy or truthfulness. At the end of the term, one student wrote:

> "I really liked how you showed that scientists aren't just nerds that sit in a lab, they are cool people that can do research outside of the lab."

Another student wrote:

> "I learned that there's a huge stereotype around scientist of being bland, male, 'nerdy', social[ly] awkward, etc. However, that's not the case at all. Scientists are passionate, funny, sociable, can be any gender, and most importantly insightful."

Several students explained their change in perspective as resulting from the exposure to scientists presented in the Spotlight assignments:

> "Any one [*sic*] can do science, regardless of their age or sex. Our view that scientist[s] are only white men have been defied throughout scientific spotlights."

> "I think that my new understanding of people who do science comes from the scientist spotlights. I always thought that scientists were people who came from the top of their class and knew what they wanted to do from day one on. However, the articles I've read have exposed me to the fact that scientists may arise from ordinary people who never knew they'd end up researching and finding such extraordinary things."

Taken together, these changes in thematic prevalence suggest that students who experience Scientist Spotlight assignments are developing a more nuanced understanding and appreciation of the individuals who participate in scientific endeavors in our society. At the start of term, students describe the Types of People Who Do Science in overwhelmingly positive ways, which is congruent with their own trajectories and identities as health sciences students. Beyond the perception of positive stereotypes associated with scientists, students early in the term tend to refer to well-known, stereotypical scientists or fail to refer to a scientist by name. Students early in the term also frequently describe or list and define different fields of science, rather than focusing on the individuals who engage in scientific work. At the end of the term, students continue to describe the Types of People Who Do Science with positive stereotypes, but include statements that reflect a more inclusive view of who belongs or who can be successful in scientific roles. Rather than simply contradicting a canonical stereotype, students described a perspective of science as an inclusive realm in which *anybody* and *everybody* has a rightful place. Often, students would use the specific scientists highlighted in the Scientist Spotlight assignments during the term to exemplify the role of diverse individuals in science and affirm that many different kinds of people do science.

One student end of term response sums it up thusly: "After doing scientist spotlights, it has become clear to me that many different people with many interests can do science. Science is a topic that has been stereotyped into explosions and flasks of mysterious liquids. But this is not the case. Science is a giant umbrella that has many different categories that people can go into, making it adaptive for many."

As another student succinctly concluded: "No matter what your story is, there is a place for you in science."

How Students Talk about Relatability

Our final research question (research question 5) sought to characterize how undergraduate students describe the ways in which they relate (or do not relate) to scientists, before and after the Scientist Spotlight assignments. A closer examination of the ways in which individual students explain their Likert-scale responses to the Relatability prompt reveals that different aspects of the Scientist Spotlight assignments are influential. In the example quotes provided here, student names have been changed to pseudonyms to preserve anonymity, although demographic characteristics have been preserved.

For example, Gabby, who identifies as a first-generation Hispanic female student, shifted from a response of "somewhat disagree" at the start of term to "somewhat agree" at the end of term. At the start of term, she wrote:

> "I have heard of many of scientists but I mostly know about the work that they've done and nothing personal that can make me relate to them."

In contrast, at the end of term, she provided this explanation:

> "I can relate to all of the female scientists, as females there are stereotypes that we have to push past. I also related more to the women of color."

TABLE 9. Scientists frequently named in postcourse Stereotypes prompt across years

Scientist	2017	2018	2019	2020	Total
Clare Fieseler	16	26	10	11	63
Agnes Day	8	15	3	6	32
Shina Yamanaka	6	10	5	7	28
Francis Kelsey	2	10	5	9	26
Lawrence David	(Not assigned as a Spotlight)	(Not assigned as a Spotlight)	9	9	18
Flossie Wong-Staal	(Not assigned as a Spotlight)	(Not assigned as a Spotlight)	2	5	7
Juan Perilla	(Not assigned as a Spotlight)	(Not assigned as a Spotlight)	1	5	6

For Gabby, therefore, having examples of scientists who shared her identity as a woman of color in science was particularly important to shifting her perception of being able to relate to scientists.

For Mark, who identifies as a white male, and who also shifted from a response of "somewhat disagree" at the start of term to "somewhat agree" at the end of term, having a shared identity with the scientists was not an impetus for his shift in perspective. At the start of term, Mark wrote:

> "It's hard for me to personally relate to an important scientist because, first, I don't personally know any, and second, I am not a scientist myself, so it's hard to know how I am supposed to relate to them."

At the end of term Mark wrote:

> "Learning about the scientists from the scientist spotlights, I know and can relate to a few of them. Knowing that they have repeatedly failed, endured hardships, and were criticized helps me to power through when things are getting tough."

In this instance, the narrative of not getting everything right on the first try and having to persevere to achieve success is what resonated with Mark as relatable and provided a source of motivation.

In their explanation of their Likert-scale responses to the Relatability prompt at the start of term, many students indicated that, although they could think of specific scientists, because they had only been exposed to the scientists' work, and not necessarily who they were as people, they did not think they could relate to those scientists.

For example, Kalia, who identifies as an Asian female student, shifted from "I don't know" at the start of term to "somewhat agree" at the end of term. She wrote:

> "I was never really exposed to scientists throughout high school. Usually we only talked about a few briefly because we were learning something that certain scientists have proven or brought attention to. Therefore, I am unsure whether there are any important scientists that I can personally relate to."

At the end of term, her perspective had shifted; she wrote:

> "I think I could personally relate to Clare Fieseler when she brought up the theme of empowering female scientists because they are underrepresented in science and society. As a woman, myself, this is relatable… men seem to be much more valued than women as men usually hold higher authoritative positions than woman, and they also get paid more."

Similarly, Sarah, a white female student shifted from "somewhat disagree" at the start of term to "agree" at the end of term, writing:

> "I know of scientists, but I do not personally relate to any of them that I know. I just cannot draw many similarities."

At the end of term, Sarah wrote:

> "While doing the scientist spotlights, I discovered scientists who are funny and also have a passion for the earth and helping it thrive. I related to this, for I also have a sense of humor and I also have a passion for helping the earth."

Alyssa, a white female first-generation student provides an interesting example of a student who always agreed, yet shifted in the ways she related to scientists, as seen in her explanation of her ability to relate to scientists.

In the beginning of term, Alyssa wrote:

> "I remember reading a children's book about Marie Curie in second grade and for some reason really identifying with her seemingly knowledge-starved brain in regard to science—I think the book was something about 'the value of learning' and that always stuck with me. I also remember watching a documentary about Rosalind Franklin in high school and just really finding the fact that a woman conducted a lot of the research that led to the discovery of the DNA structure to be cool. I am planning on going into medical lab science and/or toxicology, so having female idols along with the typically discussed fathers of science is important to me."

At the end of the term, Alyssa's ability to relate to a scientist moved beyond the shared identify of being a woman in science to include more personal aspects. She wrote:

> "In the beginning of the semester I remember putting 'Marie Curie' for this answer and I honestly don't really relate to Marie Curie at all (except for being a female), though she was a cool scientist to learn about. Through the scientist spotlights—especially the last one—I found that there are tons of contemporary scientists that I genuinely relate to. I really related to Clare Fieseler with wanting to showcase women in science and show that you can be a scientist and a multi-faceted person at the same time. I guess it just really resonated with me because I like so many things that aren't science and yet [I'm planning

to pursue a career in science] … because I like science the most. For a while this year I thought I was supposed to kind of sacrifice the other things in my life … While that is true for some things, I can still enjoy being creative in non-scientific areas or doing non-scientific things and be a science person at the same time—it isn't a death sentence."

In Alyssa's response, we see that even when students' quantitative responses do not change, these students may still be deepening their understanding of, and relationship to, science.

DISCUSSION AND CONCLUSIONS

Encouraging and supporting a broad spectrum of students to pursue study of and careers in STEM fields is a high priority in our increasingly diverse society. While there are many reasons for racial, class, and gender disparities in who pursues a STEM degree, this paper focuses on the ways that a course-level intervention might influence students' perceptions of scientists, their sense of relating to individuals defined as scientists, and by extension, their sense of identity as congruent with their academic trajectory. Specifically, stereotype threat—not seeing oneself represented as a scientist—might influence a student's sense of belonging to a STEM discipline, and sense of belonging is correlated with persistence and success in college (Trujillo and Tanner, 2014; O'Brien *et al.*, 2020).

As Few as Four Scientist Spotlight Assignments Significantly Shift Students' Self-Reported Ability to Relate to Scientists and Descriptions of the Types of People Who Do Science

This study, along with a growing number of other studies, provides a strong argument for the incorporation of low-stakes assignments in classes that seek to promote inclusive representations of scientists. The research presented in this paper corroborates and extends previous research demonstrating that a small number of Scientist Spotlight assignments can have a significant impact (Aranda *et al.*, 2021). Here, we report the finding that as few as four Scientist Spotlight assignments in a 16-week semester may significantly shift students' self-reported ability to relate to scientists as well as student descriptions of the types of people who do science. This finding is consistent regardless of the number of Spotlights assigned (four or six). This result, taken along with previously published reports of similar shifts in varying institutional contexts and across different populations of learners, indicates that instructors who are interested in diversifying their course content to include representations of diverse scientists to enhance students' ability to identify a range of possible "types of people" who do science can do so successfully through the incorporation of a small number of Scientist Spotlight assignments.

First-Generation Students Experienced a Greater Increase in Their Average Ability to Relate to a Scientist from Pre- to Postcourse Assessment, Holding Other Factors Equal

Our study is unique in adding information about how relatability in a course using the Scientist Spotlight intervention is differently associated with subpopulations of students, most especially students who identify as first-generation college students, a subpopulation that has not been specifically examined in previous studies of Scientist Spotlight assignments. Importantly, our data provide tentative evidence that Scientist Spotlight assignments may reduce the differences in self-reported relatability between first-generation students and continuing-generation students: At the start of the term, compared with continuing-generation students, students who identify as first-generation or who chose not to disclose their first-generation/continuing-generation status were *significantly less likely* to report that there was a scientist to whom they could personally relate ("relatability"). At the end of term, first-generation/continuing-generation status was no longer predictive of relatability, indicating that students who identify as first-generation students, or whose first-generation/continuing generation status is unreported, experienced a greater increase in their average ability to relate to a scientist from pre- to postcourse assessment, holding other factors equal. This finding is particularly important, given that first-generation students may struggle with fitting in, and therefore represent students who should be cultivated by instructors and institutions to instill a sense of belonging. This finding is also noteworthy because this is the first study examining the Scientist Spotlight intervention to report a significant difference between demographic subpopulations before the Scientist Spotlight intervention: Aranda *et al.* (2021) reported no significant difference in baseline relatability between men and women, URM students and non-URM students, or between Pell Grant–eligible and non–Pell Grant eligible students (p. 7).

Academic Performance in Course Only Weakly Correlates with Postcourse Relatability

We also add to the literature by demonstrating a correlation between students having a scientist to whom they could personally relate and a measure of academic preparation as well as academic performance, but the relationship is only weakly supported and is only associated with postcourse relatability. While academic preparation (i.e., ACT score) did *not* predict relatability at the beginning of the term, ACT composite score (although not ACT math) was weakly predictive at the postcourse time point, with students expressing higher relatability also having had higher composite ACT scores. Additionally, academic performance in the course (final course grade expressed as a percent) weakly predicted relatability at the end of the term: Students who performed better in the course had higher postcourse relatability. This result complements the finding of Schinske *et al.* (2016) that students who completed Scientist Spotlight assignments had statistically higher course grades as compared with students in control sections who completed a similar activity that lacked connections with diverse scientists (pp. 12–13), although there is ambiguity regarding how to interpret the relationship between the academic preparation measure (i.e., ACT score) and post relatability scores following this curricular intervention. Further, in our population, the addition of academic preparation and performance variables did little to increase the explanatory power of linear regression models beyond models that did not include these variables (see R^2 in Tables 6 and 7).

Students Who Identify as Male Are Less Likely to Increase Their Perception of Relatability to a Scientist from Pre- to Postassessment

Finally, we found descriptively that students who identify as male report a relatability approximately half a point lower than students who identify as female at the postcourse time point.

This finding indicates that males in our population are less likely to increase their perception of relatability to a scientist from pre- to postassessment. While it may be intuitive to interpret this result as male students being negatively impacted by the intervention, it is important to note that the level of relatability for students who identify as male does not change from pre- to postassessment, but rather it is the response of students who identify as female that is primarily driving the change in gender comparisons from pre- to postassessment. Being a male-identified student was weakly predictive of preassessment relatability scores, with male-identified students reporting a preassessment relatability of a quarter-point higher than female-identified students (preassessment linear regression model 2), but there was little evidence of gender predicting postassessment scores in our regression analyses (postassessment linear regression model 3).

A Growing Body of Evidence Indicates That the Scientist Spotlight Intervention is an Effective and Efficient Pathway for Instructors to Move toward More Inclusive Teaching Practices

Similar to previous studies (Brandt *et al.*, 2020), open-ended responses provided by students at the end of the course revealed descriptions of a sense of affinity with scientists featured in the assignments whose revealed identities were congruent with the students' own identities, circumstances, or habits of being. The present study reflects data collected across four offerings of a course, with rigorous assessment of the impact of assigning four to six Scientist Spotlights in an undergraduate introductory biology course for health sciences students at a small primarily undergraduate institution (PUI) in the Midwest. This study confirms the positive significant association of a course using Scientist Spotlight assignments, as has been demonstrated in other institutional contexts such as 2-year colleges (Schinske *et al.*, 2016), R1 public research institutions (Brandt *et al.*, 2020; Yonas *et al.*, 2020), and a PUI Hispanic-serving institution in an urban area on the West Coast (Aranda *et al.*, 2021), demonstrating the transferability of the intervention across multiple institution types and undergraduate populations. This study also affirms the recent finding of Aranda *et al.* (2021) that fewer Spotlight assignments (approximately four) is associated with significant shifts in students' ability to relate to scientists and the ways in which students describe scientists and the type of people who do science.

With multiple studies at different institutions demonstrating that fewer assignments can produce a significant impact, a growing body of evidence indicates that the Scientist Spotlight intervention is an effective and efficient pathway for instructors to move toward more inclusive teaching practices, without the barrier of having to construct new assignments, take additional training, or devote additional instruction time or time providing feedback to students on these assignments. In short, any instructor who has recognized the need to cultivate a more inclusive STEM environment should be persuaded and empowered to implement Scientist Spotlights as a part of their courses.

LIMITATIONS AND FUTURE WORK

There are many facets of the introductory-level course in which these assignments were situated that are designed to support students of diverse backgrounds and identities in feeling welcomed, supported, and integrated as full and valued members of the learning community. Taken together, the design of the course, the pedagogical approaches used, and the incorporation of the Scientist Spotlight assignments, as well as factors outside the class context, may all contribute to the change in students' reported perceptions within the pre- to postassessment time frame. The nature of this study precludes the determination that the effect observed is due solely to the Scientist Spotlight assignments. Additionally, it should be acknowledged that measuring relatability based on a single Likert-type item is unlikely to be a reliable, complete measure (e.g., Diamantopoulos *et al.*, 2012). However, evaluation of a random subset (~5–10% of responses from each year) of students' open-ended responses to the Relatability prompt did concur with and lend support to the validity of their Likert-scale response to the closed-ended Relatability prompt. Further, previous studies have examined the clarity of the Relatability prompt using informal interviews with undergraduate students (Aranda *et al.*, 2021). Future measurements of relatability can incorporate best practices in assessment design that provide evidence for the reliability and validity of the measure (American Educational Research Association *et al.*, 2014). This study also does not include a longitudinal component to determine whether the increase in students' relatability to scientists and more inclusive descriptions of who does science is maintained beyond the end of the semester in which the assignments were completed. A likely avenue for future study is to investigate the extent to which these observed changes are persistent in this study population.

Other avenues of future work may include the evaluation of modified approaches to the Scientist Spotlight that have been described as effective elsewhere, including alternative formats of Scientists Spotlights (e.g., podcasts) as described by Yonas *et al.* (2020), student-authored Scientist Spotlights as described by Aranda *et al.* (2021), or investigating the impact of as-yet-unexplored modifications such as allowing students to choose their own scientists to investigate rather than providing instructor-chosen Spotlights. Finally, this work leaves unanswered the question of how instructor identities (visible and/or revealed) may mediate the impact of such equity-focused interventions.

Scientist Spotlight assignments can be downloaded from the Scientist Spotlights Initiative: https://scientistspotlights.org.

ACKNOWLEDGMENTS

The authors would like to express appreciation to Dr. Robert Erdmann and Dr. Molly Ubbesen for providing thoughtful feedback at various stages of the analysis of this project and during the drafting and revision of this article. The authors would also like to thank two anonymous reviewers for helpful feedback and suggestions leading to an improved final paper for publication. This work was financially supported by start-up and departmental funds provided to K.J.M.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Aranda, M. L., Diaz, M., Mena, L. G., Ortiz, J. I., Rivera-Nolan, C., Sanchez, D. C., ... & Tanner, K. D. (2021). Student-authored Scientist Spotlights: Investigating the impacts of engaging undergraduates as developers of inclusive curriculum through a service-learning course. *CBE—Life Sciences Education*, *20*(4), ar55. https://doi.org/10.1187/cbe.21-03-0060

Bayly, B. L., & Bumpus, M. F. (2019). An exploration of engagement and effectiveness of an online values affirmation. *Educational Research and Evaluation*, *25*(5–6), 248–269. https://doi.org/10.1080/13803611.2020.1717542

Binning, K. R., Kaufmann, N., McGreevy, E. M., Fotuhi, O., Chen, S., Marshman, E., ... & Singh, C. (2020). Changing social contexts to foster equity in college science courses: An ecological-belonging intervention. *Psychological Science*, *31*(9), 1059–1070.

Brandt, S., Cotner, S., Koth, Z., & McGaugh, S. (2020). Scientist Spotlights: Online assignments to promote inclusion in ecology and evolution. *Ecology and Evolution*, *10*(22), 12450–12456. https://doi.org/10.1002/ece3.6849

Braxton, J. M. (2000). *Reworking the student departure puzzle* (1st ed.). Nashville, TN: Vanderbilt University Press.

Brooms, D. R. (2020). Helping us think about ourselves: Black males' sense of belonging through connections and relationships with faculty in college. *International Journal of Qualitative Studies in Education*, *33*(9), 921–938.

Chen, S., Binning, K. R., Manke, K. J., Brady, S. T., McGreevy, E. M., Betancur, L., ... & Kaufmann, N. (2021). Am I a science person? A strong science identity bolsters minority students' sense of belonging and performance in college. *Personality and Social Psychology Bulletin*, *47*(4), 593–606. https://doi.org/10.1177/0146167220936480

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dewsbury, B., & Brame, C. J. (2019). Inclusive teaching. *CBE—Life Sciences Education*, *18*(2), fe2. https://doi.org/10.1187/cbe.19-01-0021

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Dikmenli, M. (2010). Undergraduate biology students' representations of science and the scientist. *College Student Journal*, *44*(2), 579–588. Retrieved July 3, 2020, from http://login.ezproxy.lib.umn.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=s3h&AN=51456663&site=ehost-live

Dingel, M. J., & Sage, S. (2016). Dimensions of difference, sense of belonging, and fitting in: Tensions around developing peer groups, student body diversity, and academic culture. *Learning Communities Journal, 8*(1), 131–156.

Dingel, M., & Sage, S. (2021). Habitus congruence and college student experiences in social, academic, and racial domains. *International Journal of Qualitative Studies in Education*, 1–17. https://doi.org/10.1080/09518398.2021.1885072

Drury, B. J., Siy, J. O., & Cheryan, S. (2011). When do female role models benefit women? The importance of differentiating recruitment from retention in STEM. *Psychological Inquiry*, *22*(4), 265–269. https://doi.org/10.1080/1047840X.2011.620935

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, *13*(3), 453–468. https://doi.org/10.1187/cbe.14-03-0050

Einarson, M., & Matier, M. W. (2005). Exploring race differences in correlates of seniors' satisfaction with undergraduate education. *Research in Higher Education*, *46*(6), 641–675.

Finkelstein, M. J., Conley, V. M., & Schuster, J. H. (2016). *Taking the measure of faculty diversity* (p. 3). Washington, DC: TIAA Institute.

Fry, R., Kennedy, B., & Funk, C. (2021, April 1). *STEM jobs see uneven progress in increasing gender, racial and ethnic diversity*. Washington DC: Pew Research Center. Retrieved February 21, 2022, from www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity

Gillen-O'Neel, C. (2021). Sense of belonging and student engagement: A daily study of first- and continuing-generation college students. *Research in Higher Education*, *62*(1), 45–71. https://doi.org/10.1007/s11162-019-09570-y

Gurin, P., Dey, E. L., Hurtado, S., & Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review*, *72*(3).

Hammarlund, S. P., Scott, C., Binning, K. R., & Cotner, S. (2022). Context matters: How an ecological-belonging intervention can reduce inequities in STEM. *BioScience*, *72*(4), 387–396. https://doi.org/10.1093/biosci/biab146

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology*, *106*(2), 375–389. https://doi.org/10.1037/a0034679

Harper, S. R. (2015). Black male college achievers and resistant responds to racist stereotypes at predominantly white colleges and universities. *Harvard Educational Review*, *85*(4), 646–674.

Harwood, S. A., Mendenhall, R., Lee, S. S., Riopelle, C., & Huntt, M. B. (2018). Everyday racism in integrated spaces: Mapping the experiences of students of color at a diversifying predominantly white institution. *Annals of the American Association of Geographers*, *108*(5), 1245–1259.

Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus*, *1*(4), pgac167. https://doi.org/10.1093/pnasnexus/pgac167

Hausmann, L. R. M., Schofield, J. W., & Woods, R. L. (2007). Sense of belonging as a predictor of intentions to persist among African American and white first-year college students. *Research in Higher Education*, *48*(7), 803–839. https://doi.org/10.1007/s11162-007-9052-9

Hausmann, L. R. M., Ye, F., Schofield, J. W., & Woods, R. L. (2009). Sense of belonging and persistence in white and African American first-year students. *Research in Higher Education*, *50*(7), 649–669. Retrieved October 11, 2021, from www.jstor.org/stable/40542320

Herrmann, S. D., Varnum, M. E. W., Straka, B. C., & Gaither, S. E. (2021). Social class identity integration and success for first-generation college students: Antecedents, mechanisms, and generalizability. *Self and Identity*, *20*(5), 553–587. https://doi.org/10.1080/15298868.2021.1924251

Higher Education Research Institute. (2010). *Degrees of success: Bachelor's degree completion rates among initial STEM majors*. Los Angeles, CA: Higher Education Research Institute, University of California.

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention*, *4*(3), 227–256.

Huang, G., Taddese, N., & Walter, E. (2000). *Entry and persistence of women and minorities in college science and engineering education* (NCES 2000-601). Washington, DC: National Center Educational Statistics, U.S. Department of Education.

Hurtado, S. (2007). The study of college impact. In Gumport, P. J. (Ed.), *Sociology of higher education: Contributions and their contexts* (pp. 94–112). Baltimore, MD: Johns Hopkins University Press.

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college students' sense of belonging. *Sociology of Education*, *70*(4), 324–345.

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research*, *2010*(148), 5–15. https://doi.org/10.1002/ir.357

Jackson, M. C., Galvez, G., Landa, I., Buonora, P., & Thoman, D. B. (2016). Science that matters: The importance of a cultural connection in underrepresented students' science pursuit. *CBE—Life Sciences Education*, *15*(3). https://doi.org/10.1187/cbe.16-01-0067

Jantzer, J., Kirkman, T., & Furniss, K. L. (2021). Understanding differences in underrepresented minorities and first-generation student perceptions in the introductory biology classroom. *Journal of Microbiology and Biology Education*, *22*(3). https://doi.org/10.1128/jmbe.00176-21

Johnson, A. M. (2019). "I can turn it on when I need to": Pre-college integration, culture, and peer academic engagement among Black and Latino/a engineering students. *Sociology of Education*, *92*(1), 1–20.

Johnson, D. R. (2012). Campus racial climate perceptions and overall sense of belonging among racially diverse women in STEM majors. *Journal of College Student Development*, *53*(2), 336–346.

Jury, M., Smeding, A., Court, M., & Darnon, C. (2015). When first-generation students succeed at university: On the link between social class, academic performance, and performance-avoidance goals. *Contemporary Educational Psychology*, *41*, 25–36. https://doi.org/10.1016/j.cedpsych.2014.11.001

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lo, C. C., McCallum, D. M., Hughes, M., Smith, G. P. A., & McKnight, U. (2017). Racial differences in college students' assessments of campus race relations. *Journal of College Student Development*, *58*(2), 247–263.

Malone, K. R., & Barabino, G. (2009). Narrations of race in STEM research settings: Identity formation and its discontents. *Science Education*, *93*(3), 485–510.

Margolis, J., Fisher, A., & Miller, F. (2000). The anatomy of interest: Women in undergraduate computer science. *Women's Studies Quarterly*, *28*(1/2), 104–127. Retrieved March 10, 2022, from www.jstor.org/stable/40004448

Masta, S. (2018). Strategy and resistance: How Native American students engage in accommodation in mainstream schools. *Anthropology & Education Quarterly*, *49*(1), 21–35. https://doi.org/10.1111/aeq.12231

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Mead, M., & Metraux, R. (1957). Image of the scientist among high-school students. *Science*, *126*(3270), 384–390. https://doi.org/10.1126/science.126.3270.384

Miller, A. L., Williams, L. M., & Silberstein, S. M. (2019). Found my place: The importance of faculty relationships for seniors' sense of belonging. *Higher Education Research & Development*, *38*(3), 594–608.

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, *18*(10), 879–885. https://doi.org/10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., & Zirkel, S. (2015). Race and belonging in school: How anticipated and experienced belonging affect choice, persistence, and performance. *Teachers College Record*, *117*(12), 1–40.

Museus, S. D. (2011). Generating Ethnic Minority Student Success (GEMS): A qualitative analysis of high-performing institutions. *Journal of Diversity in Higher Education*, *4*(3), 147–162.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2021). *Call to action for science education: Building opportunity for the future*. Washington, DC: National Academies Press. https://doi.org/10.17226/26152

NASEM. (2022). *Educational pathways for black students in science, engineering, and medicine: Exploring barriers and possible interventions: Proceedings of a workshop*. Washington, DC: National Academies Press. https://doi.org/10.17226/26391

National Center for Education Statistics. (2019). *Status and trends in the education of racial and ethnic groups 2018*. Retrieved May 25, 2022, from https://nces.ed.gov/programs/raceindicators/

National Center on Safe Supportive Learning Environments (NCSSLE). (2016). *Advancing diversity and inclusion in higher education*. Retrieved December 3, 2021, from https://safesupportivelearning.ed.gov/resources/advancing-diversity-and-inclusion-higher-education

National Science Foundation. (2019). *Women, minorities, and persons with disabilities in science and engineering: 2019*. Arlington, VA: National Center for Science and Engineering Statistics. Retrieved March 4, 2020, from https://ncses.nsf.gov/pubs/nsf19304

O'Brien, L. T., Bart, H. L., & Garcia, D. M. (2020). Why are there so few ethnic minorities in ecology and evolutionary biology? Challenges to inclusion and the role of sense of belonging. *Social Psychology of Education*, *23*(2), 449–477.

Park, J. J., Kim, Y. K., Salazar, C., & Eagan, M. K. (2020). Racial discrimination and student–faculty interaction in STEM: Probing the mechanisms influencing inequality. *Journal of Diversity in Higher Education*, *15*(2), 218–229. https://doi.org/10.1037/dhe0000224

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *Journal of Higher Education*, *51*(1), 60–75.

Rosenthal, L., Levy, S. R., London, B., Lobel, M., & Bazile, C. (2013). In pursuit of the MD: The impact of role models, identity compatibility, and belonging among undergraduate women. *Sex Roles*, *68*, 464–473.

Ross, T., Kena, G., Rathbun, A., KewalRamani, A., Zhang, J., Kristapovich, P., & Manning, E. (2012). *Higher education: Gaps in access and persistence study*. Washington, DC: U.S. Department of Education. Retrieved October 12, 2021, from https://nces.ed.gov/pubs2012/2012046.pdf

Salehi, S., Berk, S. A., Brunelli, R., Cotner, S., Creech, C., Drake, A. G., ... & Ballen, C. J. (2021). Context matters: Social psychological factors that underlie academic performance across seven institutions. *CBE—Life Sciences Education*, *20*(4), ar68. https://doi.org/10.1187/cbe.21-01-0012

Schinske, J., Cardenas, M., & Kaliangara, J. (2015). Uncovering scientist stereotypes and their relationships with student race and student success in a diverse, community college setting. *CBE—Life Sciences Education*, *14*(3), ar35. https://doi.org/10.1187/cbe.14-12-0231

Schinske, J. N., Perkins, H., Snyder, A., & Wyer, M. (2016). Scientist Spotlight homework assignments shift students' stereotypes of scientists and enhance science identity in a diverse introductory science class. *CBE—Life Sciences Education*, *15*(3), ar47. https://doi.org/10.1187/cbe.16-01-0002

Schmader, T., & Beilock, S. (2012). An integration of processes that underlie stereotype threat. In Inzlicht, M., & Schmader, T. (Eds.), *Stereotype threat: Theory, process, and application* (pp. 34–50). New York, NY: Oxford University Press.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Simpson, D. Y., Beatty, A. E., & Ballen, C. J. (2021). Teaching between the lines: Representation in science textbooks. *Trends in Ecology & Evolution*, *36*(1), 4–8. https://doi.org/10.1016/j.tree.2020.10.010

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, *67*, 415–437.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4–28.

Stebleton, M. J., Soria, K. M., & Huesman, R. L., Jr. (2014). First-generation students' sense of belonging, mental health, and use of counseling services at public research universities. *Journal of College Counseling*, *17*(1), 6–20. https://doi.org/10.1002/j.2161-1882.2014.00044.x

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(3), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test-performance of African-Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811.

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C. S., & Covarrubias, R. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, *102*(6), 1178–1197. https://doi.org/10.1037/a0027143

Swail, W. S., Redd, K. E., & Perna, L. W. (2003). *Retaining minority students in higher education: A framework for success* (ASHE-ERIC Higher Education Report, Vol. 30, No. 2, Adult education). San Francisco, CA: Jossey-Bass.

Taningco, M. T. V., Mathew, A. B., & Pachon, H. P. (2008). *STEM profession: Opportunities and challenges for Latinos in science, technology, engineering, and mathematics*. San Francisco, CA: Tomás Rivera Policy Institute. Retrieved October 12, 2021, from https://eric.ed.gov/?id=ED502063

Tibbetts, Y., Harackiewicz, J. M., Canning, E. A., Boston, J. S., Priniski, S. J., & Hyde, J. S. (2016). Affirming independence: Exploring mechanisms underlying a values affirmation intervention for first-generation students. *Journal of Personality and Social Psychology*, *110*(5), 635–659. https://doi.org/10.1037/pspa0000049

Tierney, W. G. (1992). An anthropological analysis of student participation in college. *Journal of Higher Education*, *63*(6), 603–618.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*(1), 89–125. https://doi.org/10.2307/1170024

Tinto, V. (1997). Classrooms as communities. *Journal of Higher Education*, *68*(6), 599–623. https://doi.org/10.1080/00221546.1997.11779003

Tinto, V. (2000). Linking learning and leaving: Exploring the role of the college classroom in student departure. In Braxton, J. M. (Ed.), *Reworking the student departure puzzle*. Nashville, TN: Vanderbilt University Press.

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE—Life Sciences Education*, *13*(1), 6–15.

U.S. Census Bureau (2020). *Demographic turning points for the United States: Population projections for 2020 to 2060*. Retrieved May 25, 2022, from www.census.gov/library/publications/2020/demo/p25-1144.html

Weatherton, M., & Schussler, E. E. (2021). Success for all? a call to re-examine how student success is defined in higher education. *CBE—Life Sciences Education*, *20*(1), es3. https://doi.org/10.1187/cbe.20-09-0223

Winkle-Wagner, R., & McCoy, D. L. (2018). Feeling like an "alien" or "family"? Comparing students and faculty experiences of diversity in STEM disciplines at a PWI and an HBCU. *Race Ethnicity and Education*, *21*(5), 593–606.

Wood, S., Henning, J. A., Chen, L., McKibben, T., Smith, M. L., Weber, M., ... & Ballen, C. J. (2020). A scientist like me: Demographic analysis of biology textbooks reveals both progress and long-term lags. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1929), 20200877. https://doi.org/10.1098/rspb.2020.0877

Yonas, A., Sleeth, M., & Cotner, S. (2020). In a "Scientist Spotlight" intervention, diverse student identities matter. *Journal of Microbiology & Biology Education*, *21*(1), 21.1.15. https://doi.org/10.1128/jmbe.v21i1.2013

Yosso, T. J., Smith, W. A., Ceja, M., & Solórzano, D. G. (2009). Critical race theory, racial microaggressions, and campus racial climate for Latino/a undergraduates. *Harvard Educational Review*, *79*(4), 659–690.