Scientist Spotlights in Secondary Schools: Student Shifts in Multiple Measures Related to Science Identity after Receiving Written Assignments

Dax Ovid,[17]* L'vannah Abrams,[3] Tess Carlson,[4] Mark Dieter,[5] Paulos Flores,[2] David Frischer,[6] Jolie Goolish,[2,7] Michelle La-Fevre Bernt,[8] Amber Lancaster,[4] Christopher Lipski,[9] Joshua Vargas Luna,[1] Lucy M. C. Luong,[1] Marlene Mullin,[10] Mia Janelle Newman,[11] Carolina Quintero,[1] Julie Reis,[6] Freja Robinson,[12] Allison James Ross,[9] Hilary Simon,[2] Gianne Souza,[6] Jess Taylor,[2,13] Katherine E. Ward,[14] Yvonne Lever White,[15] Emily Witkop,[9] Christine Yang,[16] Aliza Zenilman,[3] Eddie Zhang,[2] Jeffrey N. Schinske,[2] and Kimberly D. Tanner[1]

[1]Department of Biology, San Francisco State University, San Francisco, CA 94132; [2]Department of Biology, Foothill College, Los Altos Hills, CA 94022; [3]San Mateo High School–San Mateo Union High School District, San Mateo, CA 94401; [4]Wallenberg High School–San Francisco Unified School District, San Francisco, CA 94115; [5]Encinal Junior Senior High School–Alameda Unified School District, Alameda, CA 94501; [6]Abraham Lincoln High School–San Francisco Unified School District, San Francisco, CA 94116; [7]Saint Mary's University of Minnesota, Minneapolis, MN 55404; [8]San Marin High School–Novato Unified School District, Novato, CA 94945; [9]Hillsdale High School–San Mateo Union High School District, San Mateo, CA 94403; [10]Eastside College Preparatory School–Ravenswood School District, Palo Alto, CA 94303; [11]K. International School Tokyo–Tokyo Metropolitan, Koto City, Tokyo 135-0021, Japan; [12]Lowell High School–San Francisco Unified School District, San Francisco, CA 94132; [13]Cornell University, Ithaca, NY 14850; [14]Aragon High School–San Mateo Union High School District, San Mateo, CA 94402; [15]Hayward High School–Hayward Unified School District, Hayward, CA 94541; [16]Chaffey High School–Chaffey Joint Union High School District, Ontario, CA 91762; [17]Department of Physiology and Pharmacology, University of Georgia, Athens 30602

ABSTRACT

Based on theoretical frameworks of scientist stereotypes, possible selves, and science identity, written assignments were developed to teach science content through biographies and research of counter-stereotypical scientists—Scientist Spotlights (www.scientistspotlights.org). Previous studies on Scientist Spotlight assignments showed significant shifts in how college-level biology students relate to and describe scientists and in their performance in biology courses. However, the outcomes of Scientist Spotlight assignments in secondary schools were yet to be explored. In collaboration with 18 science teachers from 12 schools, this study assessed the impacts of Scientist Spotlight assignments for secondary school students. We used published assessment tools: Relatability prompt; Stereotypes prompt; and Performance/Competence, Interest, and Recognition (PCIR) instrument. Statistical analyses compared students' responses before and after receiving at least three Scientist Spotlight assignments. We observed significant shifts in students' relatability to and descriptions of scientists as well as other science identity measures. Importantly, disaggregating classes by implementation strategies revealed that students' relatability shifts were significant for teachers reporting in-class discussions and not significant for teachers reporting no discussions. Our findings raise questions about contextual and pedagogical influences shaping student outcomes with Scientist Spotlight assignments, like how noncontent Instructor Talk might foster student shifts in aspects of science identity.

Sehoya Cotner, *Monitoring Editor*

Submitted Aug 1, 2022; Revised Dec 2, 2022; Accepted Feb 13, 2023

CBE Life Sci Educ June 1, 2023 22:ar22

DOI:10.1187/cbe.22-07-0149

*Address corresponding to: Dax Ovid (dax@uga.edu).

INTRODUCTION

> I remember learning about a Latino who immigrated to the United States during the 70s where Latinos were heavily discriminated against, and he became a very high-ranking surgeon.
>
> High School Student, Written Response in Preassessment (PEER, heterosexual cisgender man, 15 years old)

> Based off of the scientists most heavily portrayed and represented in America, scientists tend to be white men, as women and people of color are often discriminated against in STEM fields.
>
> High School Student, Written Response in Preassessment (Non-PEER, bisexual cisgender woman, 16 years old)

The Representation of People in STEM Fields Does Not Yet Reflect Our Society

In the United States, several decades of federal investment and advocacy have taken place to promote the recruitment and persistence of students from underrepresented backgrounds in science, technology, engineering, and mathematics (STEM), and there is still work to do. Fifty-seven years after the Title VI Civil Rights Act of 1964, which prohibits discrimination based on race, color, or national origin by recipients of federal funds, the National Academies of Sciences, Engineering, and Medicine (NASEM, 2023) published a consensus study on Advancing Anti-Racism, Diversity, Equity, and Inclusion in STEMM Organizations. Discrimination still exists. From the overarching findings of NASEM down to the perception of the high school students quoted above, there remains an ongoing need to challenge systems enabling discrimination and to advocate for diversity, equity, and inclusion in STEM fields.

Despite decades of efforts, a disproportionate number of college students who are no longer pursuing STEM majors identify as Black or African American, Hispanic or Latine,[1] and American Indian or Alaska Native, that is, People Excluded due to Ethnicity or Race (PEER; Asai, 2020). Much of this pattern has been attributed to the exclusionary practices perpetuated by the dominant culture in STEM. To intentionally challenge these biases, promote a more inclusive culture, and enable the success of PEER in STEM, how might we bolster students' sense of belonging in the sciences at academic levels *before* college in middle and high school curricula (also known as "secondary schools")? Adolescence is a developmental stage theorized to be when students "develop a notion of work fields and how their self-concept is related to these work fields" (van Tuijl and van der Molen, 2016, p. 171). One source—if not the only source—from which secondary students learn about STEM work fields is their textbooks.

The Representation of Scientists in Textbooks Does Not Yet Reflect Our Society

From discourse and image analysis of science textbooks in the United States, researchers identified that scientists were depicted as predominantly white[2] men in textbooks for grades 1–8 published from 1980 to 1988 (Provenzo *et al.*, 2010) and predominantly European in K–12 science and mathematics textbooks published from 2003 to 2013 (Chacón-Díaz, 2022). This biased representation was also found in the seven most commonly used college biology textbooks published from 2016 to 2019 (Wood *et al.*, 2020). While there has been some effort to include white women and people of color, pre-service teachers noted that information about scientists' backgrounds were boxed off separately from the text, not content focused, and therefore ignored in their teaching (Pringle and McLaughlin, 2014). The images of people from groups marginalized by racism and/or sexism in science textbooks were often not doing science but rather everyday activities like sports or cleaning, and even these depictions were stereotypical in nature (Provenzo *et al.*, 2010). Given the incremental change of scientist representation in textbooks has been slow (Wood *et al.*, 2020), other resources are needed to teach science content through the stories of Black and Indigenous scientists of color, disabled scientists, LGBTQIA+ scientists, international scientists, and scientists from working-class and first-generation college-going backgrounds.

Highlighting Counter-Stereotypical Scientists: College Biology Students Engage in Reflective Writing to Learn Course Content with Scientist Spotlight Assignments

To highlight the work of scientists from backgrounds that have been excluded from science textbooks, Scientist Spotlight assignments were designed to be written reflections that teach science content. The assignments are searchable by content area and Next Generation Science Standards (NGSS) in an online database (www.scientistspotlights.org). Each assignment offers a short introduction, a photograph featuring the scientist, and external links to two types of resources: 1) a content resource about the scientist's work and 2) a biographical resource about the scientist's background. Finally, each Scientist Spotlight assignment has written reflection questions, prompting students to write about what content they learned, what questions they have, and what the assignment tells them about the types of people that do science.

By addressing issues of curricular representation, Scientist Spotlight assignments were shown to support equitable student success in college biology (Schinske *et al.*, 2016). The assignment was developed from research in a 2-year college setting. Schinske *et al.* (2015) laid the groundwork by assessing how scientist stereotypes were described by college-level biology students at an Asian-American and Native American Pacific Islander–Serving community college. They found that students who started the course describing scientists with nonstereotypes (e.g., "Any type of person") had passed the course at significantly higher rates than students who provided more stereotypical descriptors. Subsequently, they designed an intervention to foster nonstereotypical thinking about the types of people that do science—the Scientist Spotlight assignment. The research team found that biology students in sections receiving

[1]We acknowledge the discussion is ongoing for Spanish-speaking communities over gender-neutral terminology. Latine uses morpheme-e, "already existent in the Spanish phonological and lexical systems" (Slemp, 2020, p. ii).

[2]There are debates about whether the "w" in white should be capitalized as the "b" is for Black when describing socially constructed racial categories (Appiah, 2020). The "w" is lowercase as it is found in previous studies and student written responses and capitalized when describing the label for demographic groups of the present study.

weekly Scientist Spotlight assignments had, on average, *a course grade level higher than* students in sections who received a course reading without written reflections (Schinske *et al.*, 2016). Since these initial studies, Scientist Spotlight assignments have continued to be assessed at the college level (Aranda *et al.*, 2021; Brandt *et al.*, 2020; Macdonald *et al.*, 2019; Ormand *et al.*, 2021; Yonas *et al.*, 2020). However, the impacts of these assignments on a broader population of students—students who might not otherwise choose to enroll in a college-level biology course or even consider pursuing college—have yet to be explored.

Given the findings of previous research, we predicted Scientist Spotlight assignments could be a promising intervention in secondary school science based on three theoretical and phenomenological frameworks, which are introduced in subsequent sections: 1) possible selves, 2) scientist stereotypes, and 3) science identity.

Scientist Spotlight Assignments Aim to Expand Students' Possible Selves to Include Science

As a precursor to promoting a more diverse STEM field, students from backgrounds excluded from the sciences need to be able to conceive the possibility of being and becoming scientists. To better understand "individuals' ideas of what they might become, what they would like to become, and what they are afraid of becoming," Markus and Nurius (1986, p. 954) conceptualized *possible selves* as a theory grounded in self-concept and identity to guide future research. As a dynamic and multifaceted phenomenon, possible selves could affect our behaviors and motivation, or "dispositions" toward possible selves that are hoped for and away from possible selves that are dreaded (Markus and Nurius, 1986; Dunkel and Kerpelman, 2006; Hock *et al*., 2006).

Schinske *et al.* (2016) designed a tool to assess community college students' possible science selves by asking students if they knew of one or more important scientists to whom they could personally relate. Students responded to this Relatability prompt before and after taking a biology course with weekly Scientist Spotlight assignments. After the intervention, students described scientists in relation to themselves by using phrases such as "like me" or "I am also…" followed by students' personal characteristics, suggesting that community college students shifted in their perceptions of their possible science selves (Schinske *et al.*, 2016).

To foster students' possible science selves, educators need broadly accessible and evidence-based interventions to try in their context. While possible selves have been explored in many middle school and high school studies (Oyserman *et al.*, 2004; Hock *et al.*, 2006; Mills, 2014; Shah *et al.*, 2021; Grimalt-Álvaro *et al.*, 2022), such studies have focused on outcomes of a field trip, program, or game that is not readily adaptable or accessible in other contexts. Meanwhile, the Scientist Spotlights Initiative hosts a searchable, online database of adaptable assignments (i.e., an editable Word document) that can readily be used to introduce students to the many types of people that do science. Importantly, Hill *et al.* (2017) suggested that introducing secondary students to examples that counter gender biases could expand their possible selves in science. Would secondary students increase their relatability to scientists—and expand their possible science selves—after receiving Scientist Spotlight assignments?

Scientist Spotlight Assignments Aim to Challenge Scientist Stereotypes

Previous studies show a persisting bias in students' assumptions about *scientist stereotypes* that could impact their perception of the types of people that do science; however, there are limitations to a widely used method to assess the scientist stereotypes perceived by K–12 students. Specifically, students respond to the prompt, "Draw a scientist," and researchers use the Draw-A-Scientist Checklist (DAST-C) to quantify characteristics of the drawings that students provide. A meta-analysis of research using the DAST-C summarized, across study contexts from 2003 to 2018, that students consistently drew scientists as white, older men in lab coats (Ferguson and Lezotte, 2020). DAST was designed to allow young people, with a range of home languages, to express their ideas about scientists without verbal expression (Chambers, 1983). However, there are some notable limitations of DAST. First, students may think the task is to draw a figure that is broadly recognizable as a scientist, leading students to draw from stereotypes instead of representing their own ideas about scientists (Symington and Spurling, 1990; Thomas *et al*., 2006). Second, researchers observed that the details of drawings varied greatly between students in underresourced and well-resourced schools (Chambers, 1983). In this original study, Chambers (1983) suggested these differences could be attributed to students' exposure to various scientific instruments and drawing capabilities. Third, the personal characteristics of student participants could influence the gender and race of the scientists represented in the drawings. Euro-American and African-American students in grades 1–7 mostly cited self-image as a motivating factor behind their choices (Sumrall, 1995). Finally, because students are instructed to simply "draw a scientist," the singular use of the term "scientist" suggests a sole individual. This instruction might not allow students to express the range of personal characteristics that they might have otherwise depicted. Given the limitations of drawing a scientist, how might secondary students describe the types of people that do science in a written prompt, which would allow for an elaboration of students' ideas beyond drawing?

To measure shifts in college students' descriptions of the types of people that do science, Schinske *et al.* (2016) developed the Stereotypes prompt to ask students to describe the types of people that do science. In this previous study, college students received weekly Scientist Spotlight assignments. Relative to students receiving assignments from a course reader, college students receiving Scientist Spotlights had significantly shifted how they described the types of people that do science in their written responses to include more nonstereotypical descriptors (Schinske *et al.*, 2016). Given the findings for college-level biology students, Scientist Spotlight assignments could be an effective intervention to shift secondary students' written descriptions of scientists as well.

Scientist Spotlight Assignments Aim to Invite Students to Reflect on and Develop Aspects of Science Identity

To assess key predictors of student persistence in STEM, we also considered *science identity* in the present study. A growing body of research on science identity has corresponded with

the development of quantitative measures for large-scale studies. Over the past few decades, science identity has been explored through multiple frameworks and across age groups (Simpson and Bouhafa, 2020). Science identity can be shaped by clubs, programs, games, mentoring experiences, culturally relevant pedagogy, and more (Simpson and Bouhafa, 2020). Of relevance for the present study is the PCIR instrument, measuring Performance/Competence, Interest, and Recognition in science (Godwin *et al*., 2016). The development of PCIR was informed by qualitative research on science identity that highlighted these traits as key ingredients for student persistence in science (Brickhouse *et al*., 2000; Carlone, 2004; Carlone and Johnson, 2007; Barton *et al*., 2008). Importantly, it was found that students may disengage from pursuing science if their idea of "science identity" does not align with their cultural and community values. One may wonder whether pushing the definitional boundaries of what it means to be a scientist—through Scientist Spotlight assignments aimed at both expanding students' possible selves and challenging scientist stereotypes—could shift aspects of students' science identity.

To date, PCIR has not been used to compare secondary students' shifts in science identity before and after receiving Scientist Spotlight assignments. Further, to our knowledge, evidence for validity of this instrument had not been gathered in secondary school settings. If there were evidence for validity, this quantitative PCIR instrument could complement the qualitative responses of the Relatability prompt and the Stereotypes prompt. We anticipated that using PCIR in secondary schools, before and after students receive Scientist Spotlight assignments, could assess whether students shift in their performance/competence, interest, and/or recognition in science in a large-scale and scalable manner. Beyond these methodological benefits, Dou *et al.* (2019) showed evidence that higher PCIR scores corresponded to a higher likelihood that students would pursue STEM majors (Dou *et al.*, 2019). Therefore, we anticipated that pre–post shifts in aspects of secondary students' science identity, assessed through PCIR, might be used as a proxy for potential long-term impacts of Scientist Spotlight assignments in secondary school settings.

Research Questions for Implementing Scientist Spotlight Assignments in Secondary Schools

To assess whether Scientist Spotlight assignments could expand students' possible science selves, challenge long-standing scientist stereotypes, and cultivate aspects of science identity, our multi-school study set out to address the following research questions:

1. To what extent are Scientist Spotlight assignments associated with shifts in secondary science students' *relatability to scientists*?
2. To what extent are Scientist Spotlight assignments associated with shifts in secondary science students' *stereotypes about the types of people that do science*?
3. How, if at all, do Scientist Spotlight assignments impact aspects of secondary students' science identity, through a self-assessment of their own *performance/competence, interest, and recognition (PCIR) in science*?

METHODS

To investigate student shifts in multiple measures of science identity, we employed qualitative and quantitative methods. In the following sections, we summarize our recruitment process for teacher-researchers, the assessment design, and implementation procedures. Further, we present our inclusion criteria for student participants and detail our assessment of students' relatability to and stereotypes about scientists through systematic qualitative analysis. Finally, we evaluated the evidence for validity of the quantitative assessment tool for our study context and analyzed shifts in students' responses to the PCIR instrument.

Institutional Research Board

The study was non-exempt and approved by full committee review by the San Francisco State University IRB Office (protocol no. H20-08).

Recruitment of Teacher-Researchers

To evaluate the impacts of Scientist Spotlights in secondary school science classes, we invited secondary school science teachers to collaborate ("teacher-researchers"). To recruit teacher-researchers, we contacted a convenient sample of science teachers who participated in previous Scientist Spotlight and Scientific Teaching workshops who signed up to be contacted for future research opportunities. We asked teachers to complete a survey expressing their interest in implementing Scientist Spotlight assignments in secondary schools. As an incentive, teacher-researchers were offered a $2000 stipend and the opportunity to contribute as coauthors for the resulting paper.

Assessment Design

To assess shifts in multiple measures of students' science identity, the pre- and postassessments consisted of two parts. Part 1 of the assessment included two open-ended prompts: a Relatability prompt and a Stereotypes prompt (adapted from Schinske *et al.*, 2015, 2016; see Table 1). Part 2 of the assessment included 11 closed-ended items to measure students' sense of their own performance/competence, interest, and recognition in science, henceforth PCIR (adapted from Godwin *et al.*, 2016; see Supplemental Table 1).

Scientist Spotlights Intervention in Secondary Schools

Teacher-researchers were asked to give their students a preassessment, at least three Scientist Spotlight assignments, and a postassessment over the course of a term. Before the start of the study, teacher-researchers attended an orientation (facilitated by D.O., K.D.T., and J.N.S.) to explore the database of more than 200 Scientist Spotlight assignments, mostly authored by college students (www.scientistspotlights.org offers examples of these assignments). After selecting at least three Scientist Spotlight assignments to implement in their classes, teacher-researchers could modify the assignments to align with their specific learning goals. Due to the switch to online teaching with the coronavirus pandemic, the vast majority of teacher-researchers gave students the assessments and Scientist Spotlight assignments remotely. Further, given the "excruciatingly difficult conditions brought about by COVID school closures and the lack of preparedness by school districts and community infrastructure" (as described by a coauthor), we encouraged teacher-researchers to

TABLE 1. Open-ended assessment prompts (adapted from Schinske *et al.*, 2016)

Assessment	Response type	Instructions for student respondents
Relatability prompt	Closed-ended and open-ended	"I know of one or more important scientists to whom I can personally relate." (select one) Strongly disagree – Somewhat disagree – Somewhat agree – Strongly agree Please explain your opinion of the statement.
Stereotypes prompt	Open-ended	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.

implement the assignments as they saw fit for their context with the intention of following up post facto to ask about implementation.

Inclusion Criteria for Student Participants

To enable a pre and post comparison, we included secondary science students who completed both the pre- and postassessment sequentially. Only students who agreed to participate, responded to the closed-ended Relatability prompt, and completed the 11 PCIR questions were included in the study. To ensure students received at least three Scientist Spotlight assignments, teacher-researchers provided students' homework responses. The Scientist Spotlight assignments that were implemented at each school are summarized in Supplemental Table 2.

Identifying Participant Demographics

To evaluate whether and how Scientist Spotlights would correspond to student outcomes across demographic groups, an optional form was included at the end of the postassessment, inviting participants to share their self-identified gender, race, and ethnicity. Only students who completed the relevant aspect of the demographic form were included in the disaggregated analysis. Self-identified gender included women, men, and nonbinary. Race and ethnic identities were disaggregated in two ways: by PEER and by SOC. PEER included students who identified as Hispanic or Latine, Black or African American, and mixed ethnicity and races with these designations (Asai, 2020). SOC included all students who did not identify only as white, which included and is not limited to students who identified as Asian, Middle Eastern, and mixed race.

Assessing Shifts in Secondary Students' Relatability to Scientists following the Scientist Spotlights Intervention

To explore the extent to which students in secondary school shift in their relatability to scientists, we used a Relatability prompt: "I know of one or more important scientists to whom I can personally relate" (Schinske *et al.*, 2015; Table 1). After responding to a closed-ended agreement scale (strongly agree, somewhat agree, somewhat disagree, strongly disagree), students were asked to elaborate on their choice with an open-ended prompt: "Please explain your opinion of the statement." To maximize statistical power, we grouped students as "agree" or "disagree."

To validate the assessment prompt in a novel context, J.N.S. and D.O. independently reviewed students' written responses to the open-ended prompt and investigated whether agreement could be deduced. To minimize coding bias, the pre- and postassessment responses randomly sorted. Each student's written response was coded as "agree," "disagree," or not codable. J.N.S. and D.O. coded all written responses to consensus. From this process, we created a researcher-adjusted code to analyze shifts in students' relatability to scientists. If a student's written response was not codable, we used the student's closed-ended response as the researcher-adjusted code. To explore evidence for the validity of the prompt, we checked the alignment between the researcher-adjusted code (from students' open-ended responses) and students' selection on the closed-ended agreement scale (grouped as "agree" or "disagree"). The rate of alignment between the researcher-adjusted code and students' closed-ended response is summarized in the *Results*.

Next, to analyze the researcher-adjusted code and evaluate whether there was a statistically significant shift overall in the students who selected "disagree" in the preassessment to "agree" in the postassessment, we conducted a cross-tabulation analysis and then McNemar's chi-square tests in RStudio using the mcnemar.test function from the stats package (v. 3.6.2; R Core Team, 2019). To investigate the extent to which students across demographics shifted in their relatability to scientists, we disaggregated responses by students' self-identified gender and by race and ethnicity in two ways: 1) PEER and non-PEER and 2) SOC and White students. Bonferroni-corrected McNemar's chi-square tests were conducted on cross-tabulations across demographic groups. For seven comparisons (pre and post shifts for students overall, women, men, PEER, non-PEER, SOC and White students), Bonferroni's adjustment for significance was calculated to be $p < 0.00714$.

Analyzing Secondary Students' Stereotypes about the Types of People That Do Science following the Scientist Spotlights Intervention

To assess secondary students' perceptions of the types of people that do science and shifts in scientist stereotypes before and after completing at least three Scientist Spotlight assignments, students responded to the open-ended 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" (Schinske *et al.*, 2015; Table 1).

For systematic qualitative analysis, a team of undergraduate, graduate, postdoctoral, and faculty researchers employed the coding rubric originating in Schinske *et al.* (2015) and adapted in Aranda *et al.* (2021). For each student's written response, we counted the frequency of names, words, and phrases in the following categories: 1) Descriptors, 2) Scientists, and 3) Fields of Science. For Descriptors, we coded for three categories: Positive Stereotypes (e.g., curious, interested in science), Negative Stereotypes (e.g., only people from specific backgrounds), and Nonstereotypes (e.g., anyone, regardless of background). To be

consistent with the coding practices of Schinske *et al.* (2015) and Aranda *et al.* (2021), both Positive and Negative Stereotypes were informed by previous research on scientist stereotypes (Dikmenli, 2010; Mead and Metraux, 1957), and Nonstereotypical Descriptors extend beyond these previously described stereotypes. For *Scientists*, we coded for two categories: Stereotypical Scientists (e.g., names listed in Dikmenli, 2010) and Nonstereotypical Scientists (e.g., names from Scientist Spotlights). For Fields of Science, we coded for professional titles (e.g., "cardiologist"), technical fields (e.g., "cardiology"), and everyday vernacular (e.g., "studied the heart"). This allowed us to code for diction used across a range of age groups found in the present study (see Supplemental Table 3 for the coding rubric and student examples for each category, adapted from Aranda *et al.*, 2021).

To minimize coding bias, all pre- and postassessment written responses to the Stereotypes Prompt were anonymized and randomly sorted. The coding team (P.F., J.G., J.V.L., C.Q., H.S., J.T., E.Z., D.O., K.D.T.) conducted systematic qualitative analysis on students' written responses, adapting a previously developed coding rubric for our participant population (Supplemental Table 3). We tallied the number of examples per category for each student response. Each response was discussed in pairs to reach consensus in weekly meetings. Discrepancies in interpretation that were not addressed by the existing coding rubric were brought to the larger group to discuss.

After tallying the number of codes for Descriptors (Positive Stereotypes, Negative Stereotypes, Nonstereotypical), Scientists (Stereotypical, Nonstereotypical), and Fields of Science for all written responses, we calculated both 1) the proportion of students who provided at least one example in each category and 2) the mean number of examples in each category. To compare the pre–post shift in the proportion of students providing at least one example, we used McNemar's chi-square tests (Bonferroni's adjustment for significance set at $p < 0.00714$). To compare the pre–post shift in the mean number of examples in each category, we used two-tailed paired *t* tests (Bonferroni's adjustment for significance is $p < 0.00714$). Because the *n* value is greater than 200, the parametric test (i.e., paired *t* test) is robust despite deviations from normality of the data (Fagerland, 2012; Fagerland and Sandvik, 2009). Finally, to compare to a previous analysis conducted in Schinske *et al*. (2016) using the Stereotypes prompt, we analyzed the normalized percentage of Stereotypical descriptors (Positive and Negative) relative to Nonstereotypical descriptors. By dividing by the total number of descriptors, we could normalize for the various lengths of students' open-ended responses.

Quantifying Secondary Students' Shifts in Performance, Competence, Interest, and Recognition in Science following the Scientist Spotlights Intervention

To evaluate shifts in students' self-assessment of latent variables connected to science identity, we included 11 closed-ended items from the PCIR instrument (adapted from Godwin *et al.*, 2016; Supplemental Table 1). For each item, students responded to a five-option agreement scale (strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree). We calculated the mean composite score for each construct as Likert-scale items (Boone and Boone, 2012) for each student, distinguishing from pre- and postassessment responses collected before and after receiving at least three Scientist Spotlights.

While similar items have been used for first-year college students in prior studies (Dou *et al.*, 2019), we opted to conduct exploratory and confirmatory factor analysis (EFA and CFA, as described in Knekta *et al.*, 2019) to gather validity evidence and confirm the model fits the data in our context and for the population of interest—secondary science students. To do so, the PCIR responses from the preassessment were randomly sorted and divided into a learning sample for EFA ($n = 396$) and CFA ($n = 387$). Blank responses made up less than 5% of the data and were excluded.

Multivariate normality—skewness and kurtosis—was evaluated with the Mardias test (mvn R package v. 5.9; Korkmaz *et al.*, 2014). The Kaiser-Meyer-Olkin (KMO) measure of factorability was used to test sampling adequacy (psych R package v. 2.2.9; Revelle, 2022). An inter-item correlation was used to check for the strength of correlations among PCIR items. The ideal number of factors was evaluated from scree plot and parallel analysis (nFactors R package v. 2.4.1; Raiche *et al.*, 2020). Adequate model fit was determined using previously described cutoff values for four indices (Hu and Bentler, 1999; Brown and Moore, 2012; Knekta *et al*., 2019): Comparative fit index (CFI) > 0.95, Tucker-Lewis index (TLI) > 0.95, standardized root-mean-square residual (SRMR) < 0.08, and root-mean-square error of approximation (RMSEA) < 0.06 (lavaan R package v. 0.6-9; Rosseel *et al*., 2017). Coefficient alpha was computed to assess reliability, and acceptable values were greater than 0.70 (Knekta *et al*., 2019).

To evaluate how, if at all, students shift in Performance/Competence, Interest, and Recognition in science, we calculated pre- and postassessment composite scores for each construct and analyzed shifts with the paired *t* test (Bonferroni's adjustment for significance is $p < 0.0167$).

Disaggregating Secondary Student Responses by Implementation Strategies Reported by Teacher-Researchers

Because we opted to encourage teacher-researchers to implement Scientist Spotlight assignments as they saw fit for their classes, we decided to follow up about differences in implementation that we anticipated could lead to variable student outcomes. While not a part of the original design of the research study, and therefore not a driving research question, we observed variation in how teachers reported implementing Scientist Spotlight assignments based on this programmatic evaluation. Subsequently, to explore how, if at all, teacher implementation corresponded to student outcomes across multiple science identity measures, we grouped students based on teachers' self-reported implementation strategies from a retrospective reflection survey. The survey was given to teachers after they shared student responses for the pre- and postassessments.

In the survey, teachers reported on engaging students with In-Class Discussions about Scientist Spotlight assignments and also assignment types (i.e., when students worked on the assignments). Specifically, teacher-researchers responded to the following survey questions: 1) Did you engage students in a discussion of the purpose of Scientist Spotlights *before* giving the assignment? 2) Did you engage students in a discussion of their experiences of doing Scientist Spotlights *after*

giving the assignment? 3) When did students work on Scientist Spotlights?

From the survey responses reported by teachers, we deduced four variations of engaging students with In-Class Discussions: Before giving the assignment; After, Both before and after; or Neither (i.e., the teacher reported not having any In-Class Discussions, neither before nor after implementing Scientist Spotlight assignments). Further, we identified three variations of Scientist Spotlight Assignment Types reported by teachers: In-Class, as Homework (HW), or a mix of both In-Class & HW.

For analyzing student shifts in their Relatability to scientists, we disaggregated students based on the four variations of In-Class Discussions and the three variations of Assignment Types reported by teachers. For statistical comparisons of disaggregated student responses to the Relatability prompt, Bonferroni-corrected McNemar's chi-square tests were conducted (Bonferroni's adjustment for significance set at $p < 0.00714$). We determined that if In-Class Discussions and/or Assignment Types corresponded to variable shifts in students' Relatability, we would focus on that implementation strategy to analyze data for the remaining measures—the Stereotypes prompt and PCIR. As discussed in the *Results*, variations of In-Class Discussions became the foci for subsequent analyses. First, we compared shifts in student responses for the Stereotypes prompt across the four variations of In-Class Discussions. We conducted McNemar's chi-square tests to assess pre–post shifts in the proportion of students including at least one Nonstereotype (Bonferroni's adjustment for significance set at $p < 0.0125$). Second, we analyzed students' self-assessment for PCIR using paired t test (12 comparisons, Bonferroni's adjustment for significance set at $p < 0.00417$). All statistical tests were conducted in R Studio (R Core Team, 2019).

RESULTS

In the following sections, we summarize the outcomes for our recruitment of teacher-researchers, the inclusion criteria and personal characteristics of student participants, and students' shifts in multiple measures of science identity. We disaggregated by students' self-identified gender, race, and ethnicity as well as by teacher-researchers' reported implementation strategies of Scientist Spotlight assignments in their classes.

Teacher-Researchers and Schools

All 18 teachers who completed the recruitment survey were invited to participate. Overall, the teacher-researchers worked in eight school districts and 12 different schools. The demographics of the student body for each participating school are summarized in Supplemental Table 4. Geographically, 11 of the 12 schools were based in a western, urban region of the United States, and the remaining school was based in an urban region of Japan.[3]

Student Participation and Demographics

A total of 51 classes, in which 1534 students were enrolled, met our inclusion criteria and were given at least three Scientist Spotlights assignments. Within these classes, 799 students agreed to participate, completed both the preassessment and postassessment, and responded to the closed-ended Relatability prompt. Two students left all the PCIR questions blank and were excluded, leaving our final participant number at $N = 797$, which is 52% of student enrollment in participating classes. Based on the optional form in the postassessment, students reported ages ranging from 13 to 19 years old. Further, 54% ($n = 426$) self-identified as cisgender women and 80% ($n = 637$) identified as SOC. Based on Asai's (2020) definition of PEER, 14% ($n = 112$) of students self-identified as Black, Hispanic, Latine, Indigenous, and/or mixed race. Variance across measures of science identity was visually checked with box plots by teachers, by classes, and by schools.

Relatability to Scientists: Evaluating the Relatability Prompt for Secondary School Students and Measuring Pre–Post Shifts in Students' Agreement.

To investigate shifts in secondary science students' relatability to scientists, we evaluated the validity of the Relatability prompt (Table 1) in our study context then compared students' pre–post relatability to scientists. In the following sections, we describe our process for evaluating validity by coding students' open-ended responses, comparing this to students' closed-ended responses, and developing a "researcher-adjusted code."

Evaluating the Relatability Prompt: Analyzing Secondary Students' Open-Ended Responses for a Researcher-Adjusted Code. To create a researcher-adjusted code (see *Methods*), we randomly sorted and analyzed 1594 written responses, which included 797 student responses from both the pre- and postassessments. From the preassessment, 693 (87%) of the written responses could be coded as "agree" or "disagree" based on what the student wrote. Of the postassessment responses, 710 (89%) were codable as well, indicating a high proportion of meaningful written responses for both pre- and postassessments.

To evaluate the validity of the Relatability prompt, we checked the alignment between the researcher-adjusted code from students' written responses and students' closed-ended responses. For example, we anticipated some students would choose "agree" for their closed-ended response and then write a disagreeing statement in their open-ended response (e.g., "I don't know any scientists I can relate to."). The researcher-adjusted code would more accurately reflect the student's written response.

Importantly, we found that the vast majority of students' open-ended, written responses aligned with their closed-ended responses. Less than 10% of each assessment's responses had changed with the researcher-adjusted code. For the preassessment, 6% ($n = 48$ students) chose "disagree" for the closed-ended response, while the researcher-adjusted code suggested agreement, and less than 2% ($n = 13$ students) chose "agree" while the researcher-adjusted code suggested disagreement. For the Postassessment, 8% ($n = 60$ students) chose "disagree" while the researcher-adjusted code suggested agreement, and less than 2% ($n = 14$ students) chose "agree" while the researcher-adjusted code suggested disagreement. Given the vast majority (92%) of students' written responses aligned with their closed-ended responses, we considered the Relatability prompt to be reliable for this participant population.

[3]Although in a different country from other participating schools, the school in Japan used English as the primary language of science instruction as a part of an internationally recognized curricular program (International Baccalaureate). We observed no marked differences in the kinds of responses that students provided from this context, so we opted to include them in our overarching analyses.

Quantifying Pre–Post Shifts in Students' Agreement with the Relatability Prompt. To assess the extent to which students shifted in their relatability to scientists, we conducted comparative statistical analyses to measure pre–post shifts in the proportion of students who agreed with the Relatability prompt. In Figure 1 and Table 2, we show the proportion of students who agreed and disagreed with the Relatability prompt, including students overall, by self-identified gender, and by race and/or ethnicity: PEER and SOC. Twenty-seven percent of students overall (n = 213/797) shifted from disagree in the preassessment to agree in the postassessment. The majority of students (65%) did not shift in their response, i.e., 29% (n = 228/797) always agreed, and 36% (n = 283/797) always disagreed. Only 9.2% (n = 73/797) shifted from agree in the preassessment to disagree in the postassessment. Importantly, 1.3% of students (n = 11) identified as trans* or nonbinary. Given the low n value, we did not run statistics for pre–post shifts, though a similar number of students who identified as trans* or nonbinary agreed with the Relatability prompt both pre (n = 7) and post (n = 6). For the remaining demographic groups, McNemar's chi-square tests indicated that pre–post shifts in the proportion of students who agreed with the Relatability prompt significantly increased across groups (Bonferroni's adjustment for significance set at $p < 0.00714$). Because the shifts were modest and the n values for each demographic varied (as described in the *Discussion*), we felt this data set was not conducive to a comparative analysis of differences in shifts for each demographic. The results of these seven statistical comparisons are illustrated in Figure 1 and summarized in Table 2.

In Table 3, there are representative examples of students' written responses to explain their selections on the Relatability prompt. While the themes emerging from the Relatability prompt were too varied for systematic analysis, these examples serve to offer a range of the kinds of responses that students provided. They include students who shifted from disagree to agree, agree to disagree, or did not shift in their response to the Relatability prompt. Even though the vast majority of students did not shift after receiving at least three Scientist Spotlight assignments, the nature of some students' explanations in the postassessment included reference to Scientist Spotlight assignments as a promising avenue to support their relatability to scientists in the future, for example:

> "I somewhat disagree with the statement because I feel as though we haven't learned about, nor have I met, any important scientist (outside of the scientists and doctors in my family) that I can relate to. I suppose that over the course of this next unit, with assignments like the Scientist Spotlight, I will learn about more scientists and eventually find one whom I can personally relate to."

> "I don't really relate to any of the scientists I learned about through the scientist spotlights. Their stories are really cool, but I don't see myself in them. Maybe Vivien because he did stuff with the heart, and I think hearts are interesting because of its major role in survival."

The implications of students' written responses to the Relatability prompt are noteworthy, and we explore additional interpretations of students' responses to this prompt in the *Discussion* section.

Stereotypes about the Types of People That Do Science: Systematic Qualitative Analysis of Students' Written Responses and Comparative Statistical Tests

The resulting coding rubric to evaluate whether and how students shifted in their descriptions of the types of people that do science was adapted from Aranda *et al.* (2021) and can be found in Supplemental Table 3. Categories in the rubric included Descriptors (Nonstereotypes, Positive Stereotypes, and Negative Stereotypes), Scientists (Stereotypical and Nonstereotypical), and Fields of Science. We focus subsequent analyses on shifts based on 1) the proportion of students providing at least one example and 2) the mean number of Descriptors and Scientists.

Comparing the Proportion of Students Who Provided at Least One Example of Descriptors and Scientists in the Pre- and Postassessments. From our analysis of the proportion of students offering at least one Descriptor or Scientist (shown in Table 4), we found significant shifts in all categories (for seven comparisons, Bonferroni's adjustment for significance was set at $p < 0.00714$). McNemar's chi-square tests showed a significant increase in the proportion of students who provided at least one Descriptor, from 74% (n = 591) in the preassessment to 93% (n = 714) in the postassessment ($\chi^2 = 73$, $p < 0.00714$). The proportion of students who provided at least one Nonstereotype in their written response significantly shifted from 24% (n = 191) in the preassessment to 45% (n = 359) in the postassessment ($\chi^2 = 88$, $p < 0.00714$). Further, a significantly higher proportion of students provided at least one Positive stereotype following the Scientist Spotlights intervention, increasing from 63% (n = 502) in the preassessment to 70% (n = 558) in the postassessment ($\chi^2 = 11$, $p < 0.00714$). While few students included a Negative stereotype to begin with in the preassessment (8.8%, n = 70), this significantly decreased for the postassessment (4.1%, n = 33).

Also included in Table 4 is the proportion of students who included at least one Scientist in their written responses. Additionally, Figure 2 illustrates the shift in the proportion of students including Stereotypical and Nonstereotypical Scientists in their responses. There was a significant decrease in the proportion of students naming a scientist, with 25% (n = 203) including at least one Scientist in the preassessment and 19% (n = 153) in the postassessment ($\chi^2 = 9.8$, $p < 0.00714$). Further, there was a significant increase in the proportion of students naming Nonstereotypical Scientists—11% (n = 88) included at least one Nonstereotypical Scientist in the preassessment, while 16% (n = 124) included at least one in the postassessment ($\chi^2 = 7.4$, $p < 0.00714$). The proportion of students naming Stereotypical Scientists significantly decreased from 16% (n = 130) in the preassessment to 4.0% (n = 32) in the postassessment ($\chi^2 = 69$, $p < 0.00714$).

Quantifying the Mean Number of Descriptors and Scientists in the Pre- and Postassessments. Table 5 summarizes the mean values, standard error, effect size, and two-tailed paired t tests for each category (Bonferroni's adjustment for significance set at $p < 0.00714$). Similar to Table 4, we observed significant shifts across each category of Descriptors and

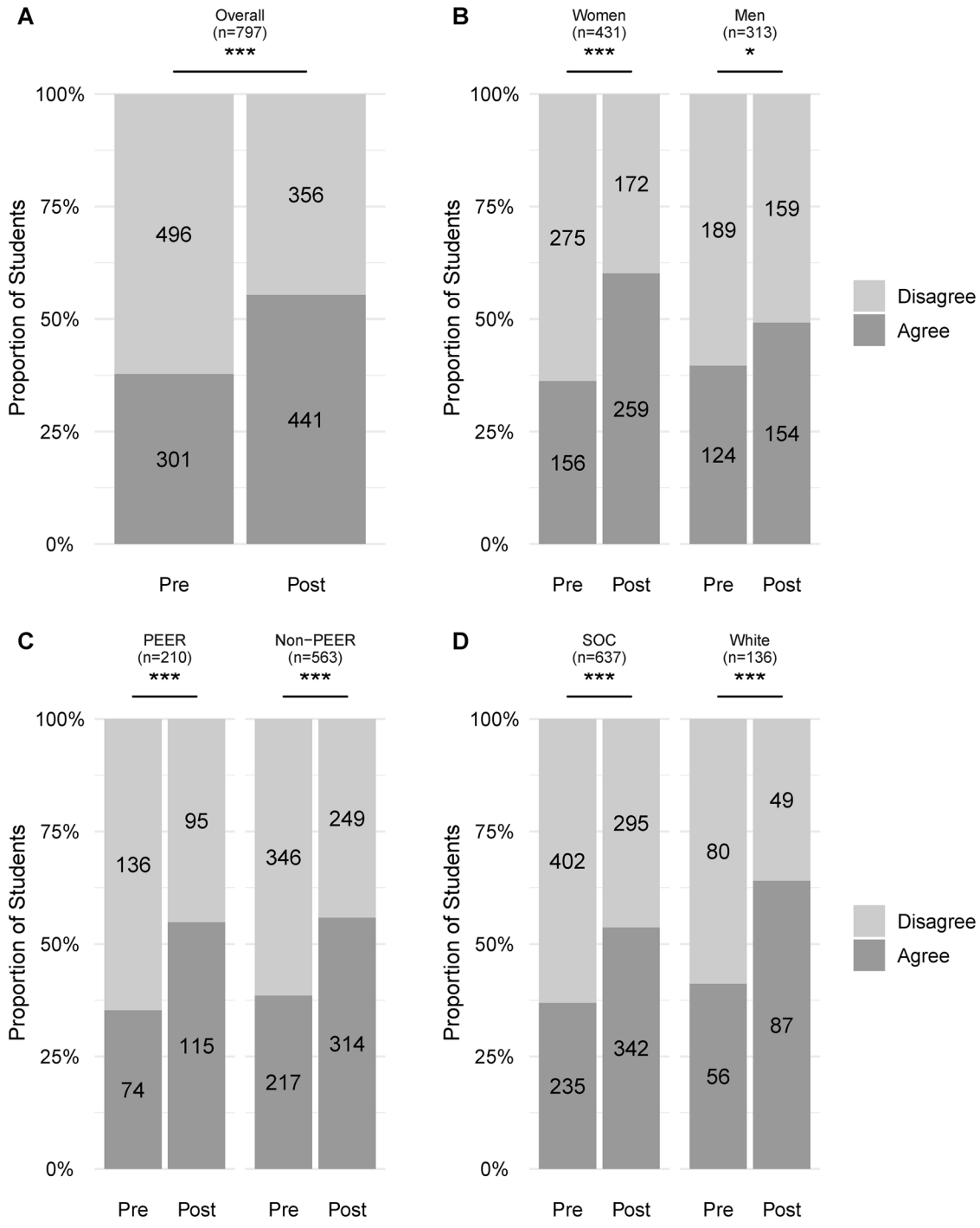

FIGURE 1. Overall shifts and demographic disaggregation of secondary science students' relatability to scientists before and after Scientist Spotlight assignments, based on researcher-adjusted code. The proportion of students who agreed ("strongly agree" and "somewhat agree") and disagreed ("strongly disagree" and "somewhat disagree") with the Relatability prompt, including (A) students overall, (B) by self-identified gender, and (C and D) by race and/or ethnicity: PEER and SOC. McNemar's chi-square tests show pre–post differences are significant at *$p < 0.01$ and ***$p < 0.0001$, with Bonferroni's adjustment for significance set at $p < 0.00714$.

Scientists, with the exception being Nonstereotypical Scientists, which was not significant; $t(796) = -2.5$, $p = 0.011$. However, the mean number of Nonstereotypes significantly increased and more than doubled from 0.37 in the preassessment to 0.86 in the postassessment; $t(796) = -8.8$, $p < 0.00714$.

Finally, to analyze our results in a manner that is comparable to previous work (i.e., Schinske *et al.*, 2016), we calculated the normalized percentage of Stereotypical descriptors (Positive and Negative) and Nonstereotypes. Because there was a large difference in the number of students providing at least one

TABLE 2. Pre- and postassessment agreement of secondary students' relatability to scientists[a]

		Preassessment	Postassessment	Disagree to agree	McNemar's chi-square	
	n	Agreement % (*n*)	Agreement % (*n*)	Pre–post shift % (*n*)	χ^2	*p* value
Overall	797	38% (301)	55% (441)	27% (213)	68	**<2.2e-16**
Women	431	36% (156)	60% (259)	32% (137)	61	**6.2e-15**
Men	313	40% (124)	49% (154)	21% (65)	8.4	**0.0037**
Trans* and nonbinary[b]	11	64% (7)	55% (6)	9% (1)	–	–
PEER	210	35% (74)	55% (115)	29% (61)	20.0	**8.8e-06**
Non-PEER	562	38% (217)	56% (314)	26% (147)	47	**4.8e-12**
SOC	637	37% (235)	54% (342)	26% (167)	49	**2.0e-12**
White	136	41% (56)	64% (87)	30% (41)	18	**2.7e-05**

[a]Based on researcher-adjusted code of closed-ended responses and disaggregated by self-reported demographics. Bold values indicate significant shifts using McNemar's chi-square tests. For seven comparisons, Bonferroni's adjustment for significance is $p < 0.00714$.
[b]Statistical comparison not conducted due to low *n* value; numbers and percentages are presented for inclusive purposes.

TABLE 3. Examples of students' pre and post written responses to the Relatability Prompt: "I know of one or more important scientists to whom I can personally relate"[a]

Preassessment Relatability response	Postassessment Relatability response
"I don't really know any scientists." (disagree)	"All of the scientists I read about came from humble beginnings and worked really hard to pursue a career in science, despite obstacles. That is the level of hard work and perseverance I aspire to have." (agree)
"I know of many scientists that have made discoveries in a variety of fields. However, I can't say I totally agree because I haven't related very much to them. I understand their point of view & their curiosity to know more but I don't relate too much to it." (disagree)	"I put I strongly agree because I now know about important scientists such as Dr. Black and Ms. Tran. I can relate to them because they studied medicine, which is something I want to do in the future. Not only that but their background is like mine, since I am also a minority." (agree)
"I don't really think that I can recall any scientists by name, and I feel like the science world is sometimes a bit of a mystery to me, even though scientists are great about publishing their findings and their work. The scientists that I know of are probably all doctors, and I don't know any scientists that work in labs, and so I don't know if I could relate to them." (disagree)	"In all three of the scientists here I feel like I see that they are really curious about the world around them, and they are excited to learn more, and I feel very similarly. They all care about people and want to help them as best as they can, which I also want to do." (agree)
"My dad is a mechanical engineer at Genentech which is a pharmaceutical company that creates medications, there some people create or try out new formulas." (agree)	"I was moved by participating in this experiment and learning that biracial females like myself can flourish and become extremely educated and successful in what I believed was a male dominant career." (agree)
"I somewhat agree with the statement because as stated earlier, Mark Rober has taught me a lot about science, and I feel like I can relate to him, however, the lack of actually knowing a proper scientist in person is there, hence why I do not strongly agree." (agree)	"I somewhat disagree with the statement because I feel as though we haven't learned about, nor have I met, any important scientist (outside of the scientists and doctors in my family) that I can relate to. I suppose that over the course of this next unit, with assignments like the Scientist Spotlight, I will learn about more scientists and eventually find one whom I can personally relate to." (disagree)
"While there's people in science that I admire, I don't really have someone that I really identify with." (disagree)	"While I admire many of the individuals we have done spotlights on, it's hard for me to personally relate to them because we have different life experiences. I haven't gone through college or the process of working in a lab, so I don't really understand the setbacks or moments of growth they had in these environments." (disagree)
"I don't personally know any scientists so I can't relate to them and I also don't know any famous scientists other than Albert Einstein, Gregor Mendel, Charles Darwin, and a few other biologists and I do not relate." (disagree)	"I don't really relate to any of the scientists I learned about through the scientist spotlights. Their stories are really cool, but I don't see myself in them. Maybe Vivien because he did stuff with the heart, and I think hearts are interesting because of its major role in survival." (disagree)

[a]Each row represents the same student, pre and post. The researcher-adjusted code is provided in parentheses after the quote.

TABLE 4. Percentage of students offering Descriptors and Scientists for Stereotypes prompt[a]

	% Students Offering ≥ 1 Example			McNemar's chi-square	
N = 797	Pre		Post	χ^2	*p* value
Descriptors	74% (591)	⇧	93% (714)	73	**<2.2e-16**
Nonstereotypes	24% (191)	⇧	45% (359)	88	**<2.2e-16**
Positive Stereotypes	63% (502)	⇧	70% (558)	11	**1.1e-03**
Negative Stereotypes	8.8% (70)	⇩	4.1% (33)	18	**1.4e-04**
Scientists	25% (203)	⇧	19% (153)	9.8	**1.7e-03**
Nonstereotypical Scientists	11% (88)	⇧	16% (124)	7.4	**6.6e-03**
Stereotypical Scientists	16% (130)	⇩	4.0% (32)	69	**<2.2e-16**

[a]Arrow direction indicates the significant increase (up) or decrease (down) in the percentage of students offering at least one example of Descriptors or Scientists in their written response to the Stereotype Prompt. Bold values indicate significant shifts using McNemar's chi-square tests. For seven comparisons, Bonferroni's adjustment for significance is $p < 0.00714$.

Descriptor in the preassessment (n = 591) and the postassessment (n = 714), we were unable to conduct paired statistical analyses on these data. The results are shown in Supplemental Figure 1.

In Table 6, we include a range of examples of students' written responses to the Stereotypes prompt. We observed a pattern of students focusing more on Nonstereotypes in their written responses for the postassessment (e.g., "All types of people," "Anyone, no matter gender, age, or race," or "There isn't a singular type"). As seen in previous studies, students' responses to the Stereotypes prompt in the preassessment tended to focus more on fixed attributes (e.g., "dedicated," "interested," or "smart") or fields of study ("biologist," "chemist," etc.).

Secondary Students' Performance/Competence, Interest, and Recognition in Science

As an additional measure of science identity, we evaluated the evidence of validity for the PCIR instrument with secondary students with EFA and CFA. In the following sections, we describe the results of EFA/CFA and comparative statistical tests for each construct, including performance/competence, interest, and recognition in science.

EFA and CFA of PCIR Instrument. The constructs and item prompts for the PCIR instrument are summarized in Supplemental Table 1 (adapted from Godwin *et al*., 2016), and the factors loaded as expected (see Supplemental Table 5 for factor loadings and coefficient alpha). Specifically, to conduct EFA/CFA for PCIR in our study context, we first tested whether various assumptions were met. For univariate normality, the Mardias test evaluates skewness and kurtosis and both had values less than |2.0| for all items (Bandalos and Finney, 2010; Knekta *et al*., 2019). Further, KMO showed measure of sampling adequacy to be greater than 0.89 for all items (cutoff > 0.70 indicates good factorability; Knekta *et al*., 2019). The inter-item correlations were greater than 0.4 among PCIR items, which exceeds the 0.3 cutoff expected for similar factors (Knekta *et al*., 2019). The scree plot recommended two factors, and Horn's parallel analysis retained three factors. Because the PCIR instrument was previously used and validated with three factors, the EFA was conducted with three factors and promax rotation. Finally, adequate model fit was confirmed: CFI > 0.95, TLI > 0.95, SRMR < 0.08, and RMSEA < 0.06 (Hu and Bentler, 1999; Brown and Moore, 2012; Knekta *et al*., 2019).

Comparative Statistical Tests of PCIR Constructs before and after Scientist Spotlights. In Table 7, we summarize comparative statistics for the three PCIR constructs in students' pre- and postassessment responses. From the pre- to the postassessment, we observed a significant increase for two of the three construct means—Recognition and Performance/Competence, but not Interest. As described in the *Methods*, after taking the mean value across items for each construct (see Supplemental Table 1 for item prompts), we conducted three, two-tailed paired t tests with Bonferroni's adjustment for significance set at $p < 0.0167$. We found that the construct mean (M) for the Recognition score (e.g., "I see myself as a 'science person'") significantly increased from the preassessment to the postassessment; $M_{Pre} = 3.13$, $SEM_{Pre} = 0.035$, $M_{Post} = 3.25$, $SEM_{Post} = 0.035$, $t(796) = -4.9$, $p = 1.4e\text{-}06$. Further, we calculated that the Performance/Competence score (e.g., "I can do well on exams relating to scientific concepts and ideas") significantly increased from pre to post; $M_{Pre} = 3.76$, $SEM_{Pre} = 0.026$, $M_{Post} = 3.84$, $SEM_{Post} = 0.025$, $t(796) = -4.2$, $p = 2.6e\text{-}05$. However, there was no significant shift in the construct mean of the Interest score (e.g., "I enjoy learning new scientific concepts and ideas"); $t(796) = 0.83$, $p = 0.41$.

Implementation Strategies of Scientist Spotlights Reported by Teacher-Researchers: Do They Correspond to Shifts in Multiple Measures of Science Identity?

While we did not set out to study implementation strategies, we were able to collect evidence to assess potential variation in how teachers reportedly implemented Scientist Spotlight assignments in their classes. Even though the aforementioned shifts in multiple measures of science identity were significant, they were not as large as we expected based on previous studies (Aranda *et al*., 2021; Schinske *et al*., 2016). In a retrospective reflection survey we gave to teacher-researchers after sharing their data, we noticed differences in how teachers reported their implementation strategies. Therefore, we predicted that differences in implementation strategies might correspond to varying degrees of students' shifts across science identity measures. To test this prediction, we grouped students based on teachers' self-reported implementation strategies for the following analyses, and we evaluated students' outcomes on multiple measures of science identity.

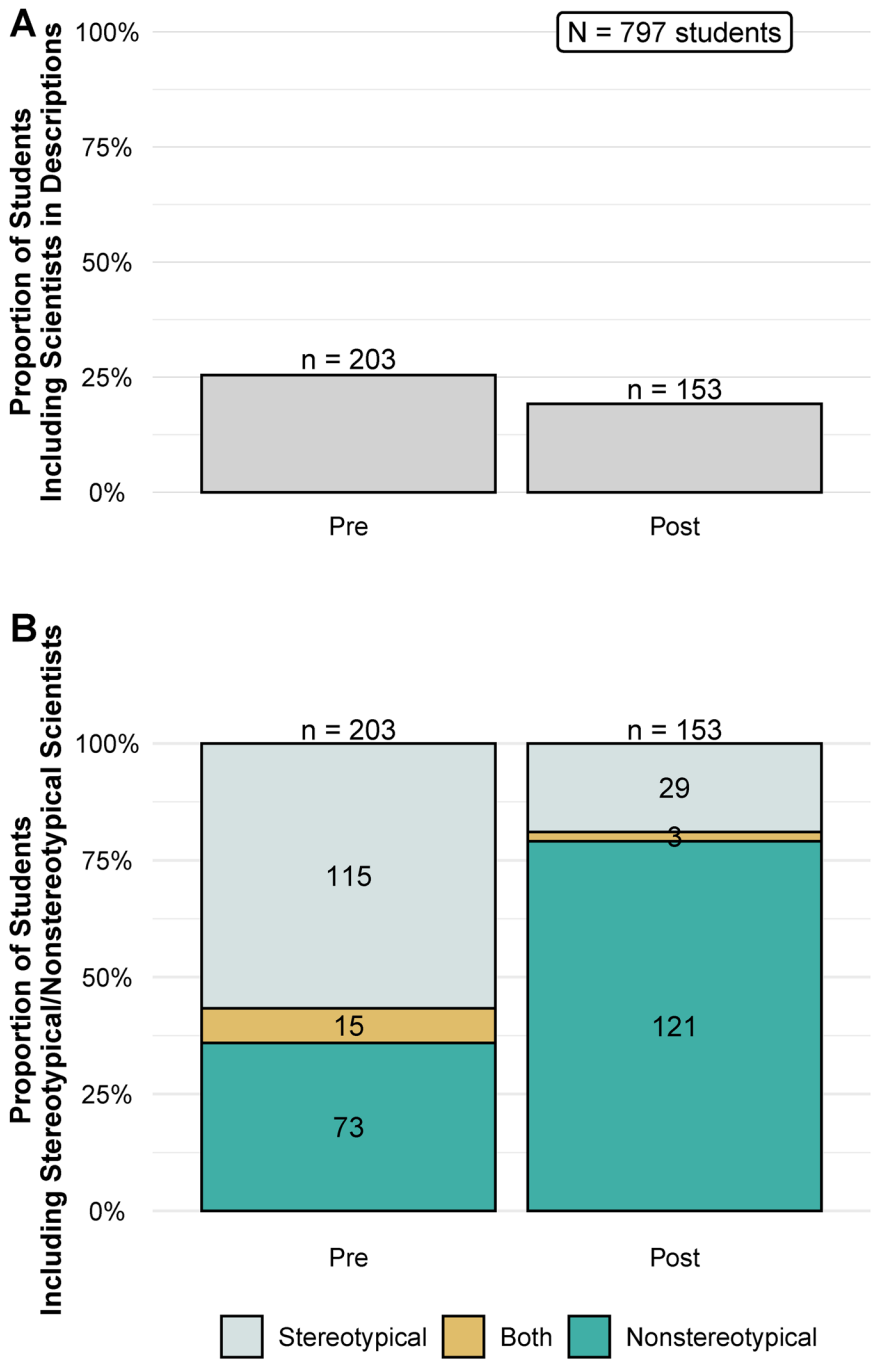

FIGURE 2. Scientists in students' responses about the types of people that do science. Scientists include Stereotypical, Nonstereotypical, and Both (see *Methods* for details). (A) Proportion of students out of the total ($N = 797$) who included at least one Scientist in their written responses to the Stereotypes prompt. (B) Proportion of students including Scientists that were Stereotypical (gray/top), Nonstereotypical (blue/bottom), or Both (gold/middle) among students who included at least one Scientist (pre = 203 students, post = 153 students).

Identifying Variations in Scientist Spotlight Implementation Strategies from Teacher-Researchers' Responses to a Retrospective Reflection Survey. To assess whether variations in implementation strategies corresponded to student outcomes, we analyzed the breakdown of In-Class Discussions and Assignment Types for each teacher from a retrospective teacher-researcher reflection survey, as described in the *Methods*. Based on teacher-researcher responses, we identified four variations of In-Class Discussions: Neither before nor after implementing Scientist Spotlight assignments ($n = 121$ students), just Before ($n = 297$ students), just After ($n = 156$ students), or Both before and after ($n = 223$ students). Further, we could group students into three variations of Assignment Types: Scientist Spotlights were completed In-Class ($n = 172$ students), as Homework (HW; $n = 167$ students), or both In-Class & HW ($n = 458$ students). Students for each teacher are grouped accordingly in the subsequent sections.

Investigating the Relationship between Students' Relatability to Scientists and Implementation Strategies: In-Class Discussions and Assignment Types Reported by Teachers. In Figure 3 and Table 8, we show the proportion of students who agreed with the Relatability prompt, disaggregated by implementation strategies reported by teacher-researchers. McNemar's chi-square tests showed a significant increase in the proportion of students who agreed in all groups except the Neither group, which reportedly did not have In-Class Discussions (Bonferroni's adjustment for significance set at $p < 0.00714$). The proportion of students who switched from disagree to agree from the pre- to the postassessment corresponded to the timing and frequency of In-Class Discussions: Neither before nor after (19%, $n = 23/121$); Before (23%, $n = 67/297$); After (31%, $n = 49/156$); and Both (33%, $n = 74/223$). However, the proportion of students who shifted from disagree to agree as grouped by Assignment Type did not have as clear a relationship: In-Class (25%, $n = 42/167$); HW (30%, $n = 52/172$); In-Class & HW (26%, $n = 119/458$).

Given that In-Class Discussions uniquely corresponded to differences in students' pre–post shifts in Relatability, and Assignment Type did not, we focused the following analyses on students grouped by In-Class Discussions reported by teachers.

Exploring Students' Nonstereotypes of Scientists in Relation to In-Class Discussions Reported by Teachers. In Figure 4, we show the proportion of students including at least one Nonstereotype in their written response to the Stereotypes

TABLE 5. Summary statistics for Stereotypes prompt[a]

	No. of descriptors (M ± SEM)		Kendell's W	Paired *t* test	
N = 797	Pre	Post	Effect size	*t* (796)	*p* value
Descriptors	1.8 ± 0.060	2.4 ± 0.061	0.067 *(small)*	−8.8	**<2.2e-16**
Nonstereotypes	0.37 ± 0.028	0.86 ± 0.046	0.12 *(small)*	−10.0	**<2.2e-16**
Positive Stereotypes	1.3 ± 0.050	1.5 ± 0.051	0.0094 *(small)*	−3.2	**1.3e-03**
Negative Stereotypes	*0.14* ± 0.018	*0.049* ± 0.0090	0.023 *(small)*	4.5	**7.9e-06**
Scientists	*0.38* ± 0.028	*0.26* ± 0.021	0.014 *(small)*	3.7	**2.3e-04**
Nonstereotypical Scientists	0.14 ± 0.016	0.20 ± 0.018	0.012 *(small)*	−2.5	0.011
Stereotypical Scientists	*0.24* ± 0.022	*0.056* ± 0.011	0.014 *(small)*	7.5	**1.2e-13**

[a]Mean (M) ± standard error of the mean (SEM) of the number of Descriptors (Nonstereotypes, Positive Stereotypes, and Negative Stereotypes) and Scientists (Nonstereotypical and Stereotypical). Italic values indicate significant decrease in mean value from pre to post. Bold values indicate significant shifts using a two-tailed paired *t* test. For seven comparisons, Bonferroni's adjustment for significance is $p < 0.00714$.

prompt, grouped by In-Class Discussions reported by teachers. Surprisingly, all groups showed significant shifts in the proportion of students including at least one Nonstereotype (McNemar's chi-square tests, Bonferroni's adjustment for significance $p < 0.0125$), suggesting that In-Class Discussions did not correspond to variable shifts in students' stereotypes about the types of people that do science.

Examining Students' Performance/Competence, Interest, and Recognition in Science in Relation to In-Class Discussions Reported by Teachers. In Table 9, based on two-tailed, paired *t* tests, we show that students who experienced In-Class Discussions both before and after receiving Scientist Spotlights—Both—significantly increased in their Recognition and Performance/Competence scores from pre–post assessments.

TABLE 6. Examples of students' pre and post written responses to the 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"[a]

Preassessment Stereotypes responses	Postassessment Stereotypes responses
"One type of person who does science is probably those who interested in science. One example of a scientist is a biologist. Another example of a scientist could be a pathologist, an archaeologist, or a chemist."	"A wide variety of people are doing science and there isn't a 'singular type' of person in the field of science."
"Based off of the scientists most heavily portrayed and represented in America, scientists tend to be white men, as women and people of color are often discriminated against in STEM fields. However, this portrayal of scientists as white men is not based off of intellect only; white men tend to be more privileged and have more opportunities to pursue science, easier access to education, and face less discrimination."	"Scientists are people who study science, regardless of their race, gender, sexual orientation, etc. While scientists in the past/throughout history have been predominantly white, male, and heterosexual, people with other identities are able to do science as well, although they face more obstacles entering the STEM field."
"There are many different types of people that are all classified under 'scientists.' When people speak of scientists, they could be referring to biologists, geologists, astrologists, or many other types of people of those professions."	"After learning about the different types of people that do science, I know that not all scientists come from genius families or have science that has run in their families for decades. These scientists pursue their jobs because of their interest in it and some even go against their parent's wishes of them doing a specific career to do so. I have learned that anyone, no matter gender, age, or race can become a scientist which shows the diversity in the system and how far science as a whole has come."
"People that do science are people who are interested in a study of something, whether it be about the environment or about space, almost anything you can think of, there can be a scientist in that field. They will create and perform experiments to find the answer to their question using science."	"All types of people do science. They come from all backgrounds, races and ethnicities. There are many different types of scientists out there, it isn't just white people in lab coats."
"A lot may be due to stereotypes, but I feel like people who are scientists are old white men like Albert Einstein and Isaac Newton. I feel like scientist are very enthusiastic about learning and finding new information. They are very dedicated to the work they do and are always trying to be more recognized for their work like other popular scientists."	"People who are passionate about wanting to learn by and discover. Seeing Blake Riggs also allowed me to see that it can also be people who are passionate about helping others get to where he is in his career in order to achieve more diverse representation in the STEM field. Hearing about Erika Zavaletas['s] story and seeing a woman of color with immigrant parents just like me allowed me to see the representation of women of color who are passionate in science as well as see that people from various back grounds do science and have the talent and dedication to do so."

[a]Each row represents same student, pre and post.

TABLE 7. PCIR summary statistics overall for pre- and postassessment Construct mean (M) ± standard error of the mean (SEM)[a]

	Construct mean (M ± SEM)		Kendell's W	Paired *t* test	
N = 797	Pre	Post	Effect size	*t*(796)	*p* value
Recognition score	3.13 ± 0.035	3.25 ± 0.035	0.038 *(small)*	–4.9	**1.4e-06**
Interest score	*4.14* ± 0.029	*4.12* ± 0.030	N/A	0.83	0.41
Performance/competence score	3.76 ± 0.026	3.84 ± 0.025	0.014 *(small)*	–4.2	**2.6e-05**

[a]Bold values indicate significant shifts using the two-tailed paired *t* test. Italics indicate mean values that decreased from pre to post. For three comparisons, Bonferroni's adjustment for significance is $p < 0.0167$.

Given the 12 comparisons (for four In-Class Discussion groups and three constructs), Bonferroni's adjustment for significance was $p < 0.00417$. Even though other groups had some measures that were close to significant, it appeared that the students grouped in Both (i.e., teacher-researchers who reported discussing Scientist Spotlight assignments before and after implementing them) were the only group that showed significant pre–post shifts in students' mean construct scores for Recognition and Performance/Competence.

DISCUSSION

This study addresses a gap in the literature by exploring whether evidence-based curricular supplements used in college biology, Scientist Spotlight assignments, might also support shifts in multiple measures of science identity for secondary students. We found that secondary students receiving Scientist Spotlight assignments increased in their relatability to and descriptions about scientists, as well as their sense of recognition and performance/competence in science, based on the measures we selected for our study context. Additionally, this study advances our understanding of how teacher implementation of Scientist Spotlight assignments could lead to variable student outcomes.

Here, we discuss key takeaways from these results, compare our findings with those of previous studies in college-level biology, and offer recommendations for future research on Scientist Spotlight assignments.

Key Takeaway: As Few as Three Scientist Spotlight Assignments Can Lead to Significant Shifts for Secondary Science Students

To explore whether Scientist Spotlight assignments could be an effective curricular intervention in secondary school settings, we evaluated their impact using assessment tools developed in previous studies for college students. Scientist Spotlight assignments have been shown to be an effective intervention in the context of college-level biology; however, the outcomes of Scientist Spotlight assignments for secondary students have yet to be determined. Our findings showed pre to post shifts in the proportion of secondary students who related to the types of people that do science, their descriptions of scientists, and measures of science identity. In the following sections, we elaborate on these main findings and implications for future research.

Possible Science Selves: More Secondary Students Found Scientists to Be Relatable after the Scientist Spotlights Intervention Than Before. After receiving at least three Scientist Spotlight assignments, a significantly higher proportion of secondary students reported knowing important scientists to whom they could personally relate. Even though there was a significant pre–post increase, we explore possible explanations for why this increase might not have been as large as expected in subsequent parts of the *Discussion*.

Beyond whether or not students' agreed with the Relatability prompt in closed-ended responses, we observed in the open-ended responses that secondary students made connections between scientists and themselves by using phrases "like me" following the Scientist Spotlights intervention:

> "Now, I am familiar with scientists who have experienced similar struggles to struggles which I face. Prior to this experience, I wasn't really familiar with scientists who were like me."

> "I think that everyone automatically assumes that people who do science have to be old white men, but I think in actuality, there are young peoples [*sic*] and young women like me that do science and are successful in their fields."

> "I've seen many women in STEM through the scientist spotlight and have thoughts of my own on who I relate to. Kizzmekia Kizzy Corbett interested me the most because of her heavy involvement in virology and immunology responses during the Pandemic. I'm just glad to see people who look like me do what I want to see myself doing."

> "Like Mercedes Lopez, I am and immigrant and will be the first in my family to attend a 4 year college."

In responses like these, we observed students making connections between where they came from and how they envision themselves in the future while referencing Scientist Spotlight assignments.

We also noted that some students did not shift in their Relatability to scientists following the Scientist Spotlights intervention:

> "I relate to the fact that they are minorities just like me, but I still don't know if I will become a scientist like them in the future."

> "There are certain female scientists whom I've learned about who inspire me and make me think I could end up in situations similar to theirs, such as Rosalind. However, I also know that there are heavily fewer female scientists than male scientists, so I don't know that many that I can relate to. But perhaps through more reading and studying, I can learn about more scientists who are like me."

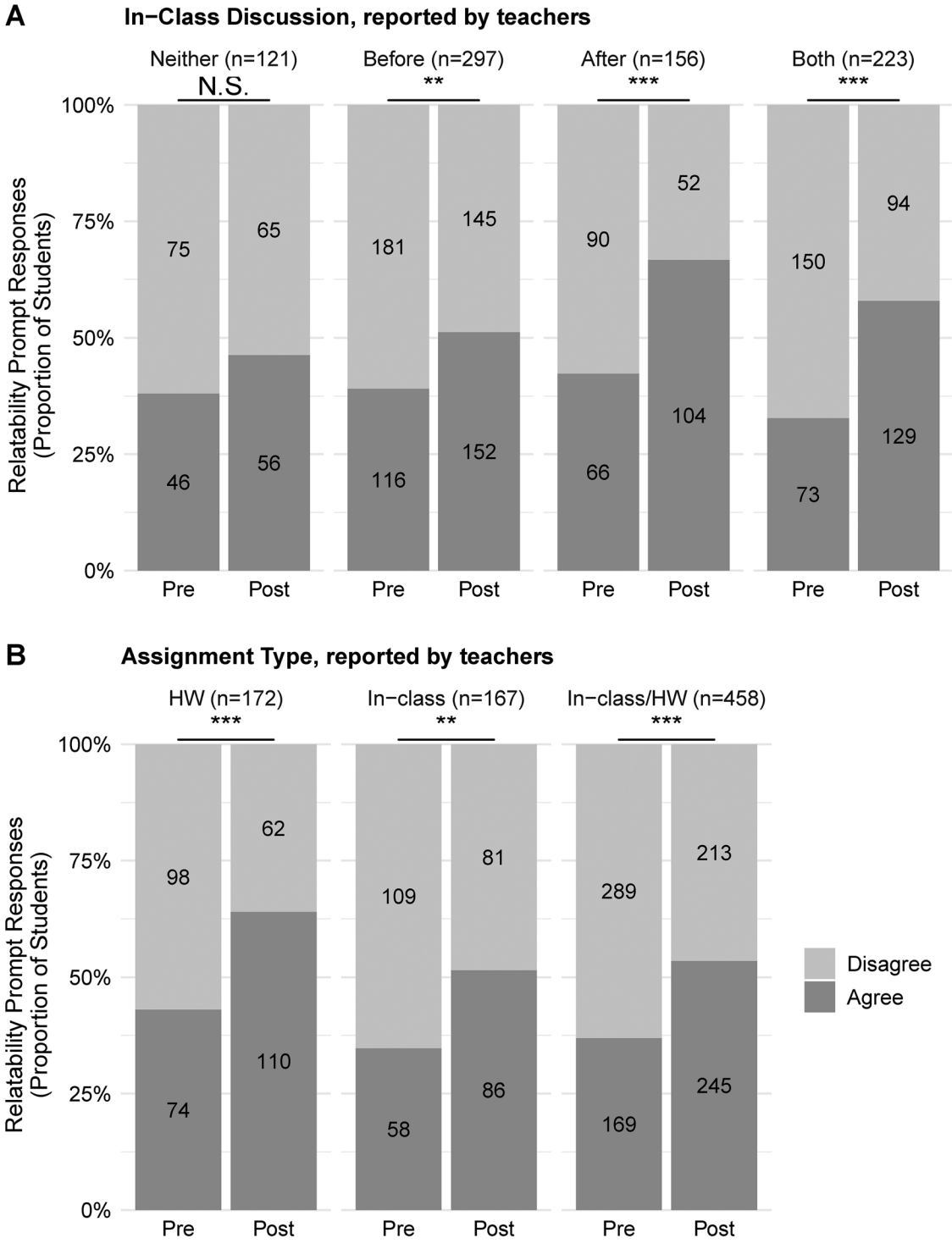

FIGURE 3. Proportion of students agreeing with Relatability prompt, disaggregated by implementation strategies as reported by teachers. The proportion of students who agreed ("strongly agree" and "somewhat agree") and disagreed ("strongly disagree" and "somewhat disagree") with the Relatability prompt, disaggregated (A) by In-Class Discussions for Scientist Spotlights, grouped by teachers who reported having In-Class Discussions with students: Neither before nor after the Scientist Spotlights intervention ($n = 121$ students), only Before ($n = 297$ students), only After ($n = 156$ students), or Both before and after ($n = 223$ students); and (B) by Assignment Type, grouped by teachers who reported students completed the assignment as homework (HW; $n = 172$ students), In-Class ($n = 167$ students), or a mix of both ($n = 458$ students). McNemar's chi-square tests show pre–post differences are significant for teacher reported In-Class Discussions Before, After, and Both, as well as all Assignment Types (N.S. = not significant, ** $p < 0.001$ and *** $p < 0.0001$, Bonferroni's adjustment for significance $p < 0.00714$).

TABLE 8. Pre- and postassessment agreement of secondary students' Relatability to Scientists, grouped by Implementation Strategies reported by teachers[a]

			Preassessment	Postassessment	Disagree to agree	McNemar's Chi-square	
		n	Agreement % (*n*)	Agreement % (*n*)	Pre–post Shift % (*n*)	χ^2	*p* value
	Overall	797	38% (301)	55% (441)	27% (213)	68	**<2.2e-16**
In-Class Discussion	Neither	121	38% (46)	46% (56)	19% (23)	2.3	0.13
	Before	297	39% (116)	51% (152)	23% (67)	13	**4.1 e-04**
	After	156	42% (66)	67% (104)	31% (49)	23	**1.8e-06**
	Both	223	33% (73)	58% (129)	33% (74)	33	**9.8e-09**
Assignment Type	In-Class	167	35% (58)	51% (86)	25% (42)	13	**3.1e-04**
	Homework (HW)	172	43% (74)	64% (110)	30% (52)	18	**2.2e-05**
	In-Class & HW	458	37% (169)	53% (245)	26% (119)	35	**3.80e-09**

[a]Based on researcher-adjusted code of closed-ended responses and disaggregated by Implementation Strategies reported by teacher-researchers: In-Class Discussion and Assignment Type. Bold values indicate significant shifts using McNemar's chi-square tests. For seven comparisons, Bonferroni's adjustment for significance is $p < 0.00714$.

> "The boundaries of what defines a scientist are still blurry to me, but I do not feel like I know, personally, scientists that might look like me or have the same background as me."

While these students described an overlap between scientists they learned about and aspects of their identity and background, or questioned what defines being a scientist, they students still doubted whether or not shared traits imply that they, too, could be scientists.

The manner in which scientists represent and describe themselves can shape how relatable they might be for students. For example, to explore the impacts of science media for students who might be science averse, a previous study invited non-science majors to share what they thought about various scientists featured in podcast episodes of Story Collider (www.storycollider.org). The non-science majors found the most relatable scientists to be ones with whom they shared values and personal characteristics (Yonas *et al.*, 2020). Therefore, future research could explore how scientists describe themselves in media resources that get referenced in science classes and the extent to which these representations might correspond to shifts in students' relatability to scientists and possible science selves.

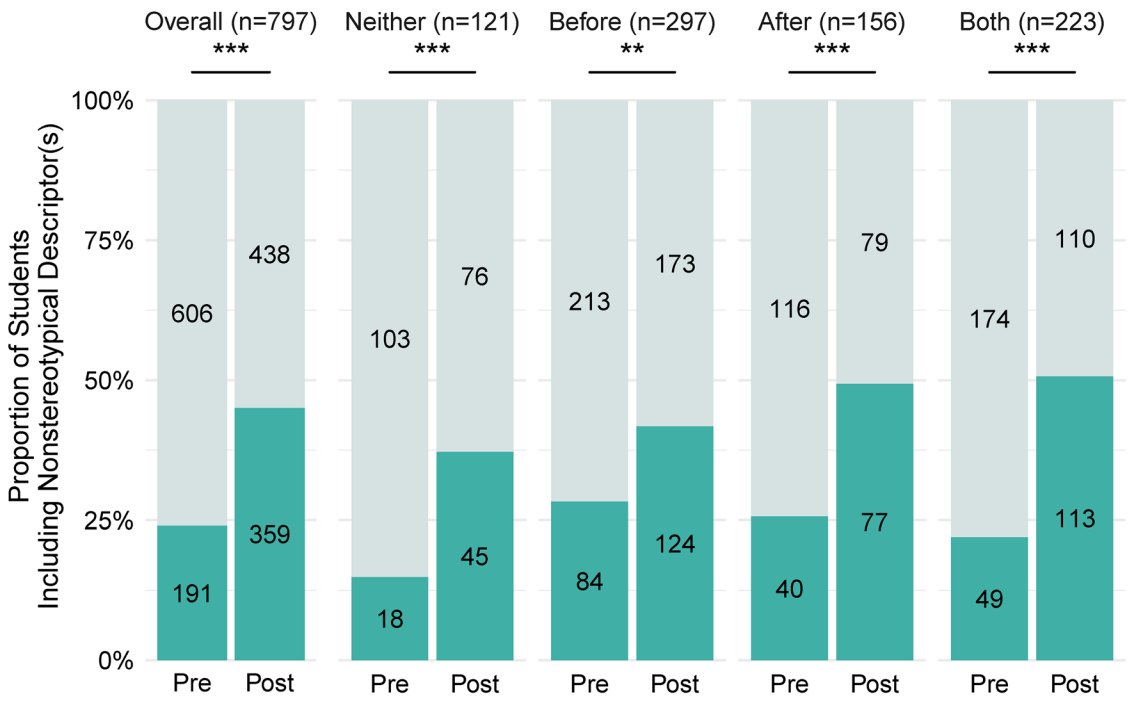

FIGURE 4. Proportion of students including Nonstereotypical Descriptors, grouped by In-Class Discussions reported by teachers. Teachers reported having In-Class Discussions with students Neither before nor after the Scientist Spotlights intervention ($n = 121$ students), only Before ($n = 297$ students), only After ($n = 156$ students), or Both before and after ($n = 223$ students). McNemar's chi-square tests show significant increase in the proportion of students including at least one Nonstereotypical descriptor in their written responses to the Stereotypes prompt ($p < 0.001$ and ***$p < 0.0001$, Bonferroni's adjustment for significance $p < 0.0125$).**

TABLE 9. PCIR Summary Statistics for Pre- and Post-Assessment, grouped by In-Class Discussions reported by teachers

	Both (n = 223)				After (n = 156)			
	Mean (M)		Paired t test		Mean (M)		Paired t test	
	Pre	*Post*	t(222)	p value	*Pre*	*Post*	t(155)	p value
Recognition	3.05	3.22	–3.7	**3.0e-04**	3.10	3.27	–2.9	0.0044
Interest	4.08	4.14	–1.3	0.19	4.21	4.21	–0.06	0.95
Performance/Competence	3.74	3.89	–3.8	**1.5e-04**	3.77	3.80	–0.56	0.58
	Before (n = 297)				**Neither (n = 121)**			
	Mean (M)		**Paired t test**		**Mean (M)**		**Paired t test**	
	Pre	***Post***	**t(296)**	**p value**	***Pre***	***Post***	**t(120)**	**p value**
Recognition	3.31	3.39	–2.4	0.019	2.87	2.90	–0.49	0.63
Interest	*4.25*	*4.19*	1.6	0.11	*3.90*	*3.81*	1.4	0.16
Performance/Competence	3.82	3.88	–1.7	0.094	3.60	3.74	–2.6	0.0092

Bold values indicate significant shifts using the two-tailed paired t test. Italics indicate Mean values that decreased from Pre to Post. For 12 comparisons, Bonferroni's adjustment for significance is $p < 0.00417$.

Scientist Stereotypes: More Secondary Students Describe the Types of People That Do Science with Nonstereotypes after the Scientist Spotlights Intervention Than Before. While the DAST-C is a widely used assessment tool to measure K–12 students' scientist stereotypes (Ferguson and Lezotte, 2020), we investigated whether secondary students' written responses to the previously developed Stereotypes prompt could offer additional insights. Notably, students' written responses to the Stereotypes prompt were codable with the existing rubric (developed by Schinske *et al*., 2015; adapted by Aranda *et al.*, 2021).

Because the Stereotypes prompt effectively solicited secondary students' descriptions about the types of people that do science, we anticipate that this prompt could be used in future studies to assess stereotypes about scientists. In the present study, secondary students were able to offer nonstereotypical descriptions that might not be easily depicted in a drawing, for example:

> "After learning about the different types of people that do science, I know that not all scientists come from genius families or have science that has run in their families for decades. These scientists pursue their jobs because of their interest in it and some even go against their parent's wishes of them doing a specific career to do so. I have learned that anyone, no matter gender, age, or race can become a scientist which shows the diversity in the system and how far science as a whole has come."

It is hard to imagine how a student could draw "anyone" or include a person's background, interests, and families in a drawing. The percentage of secondary students offering nonstereotypical descriptors like these in their written responses nearly doubled after the Scientist Spotlights intervention. Further, we were intrigued to find that very few students in the present study offered negative stereotypes of the types of people that do science, and significantly fewer students included negative stereotypes in the postassessment responses relative to the preassessment responses. Interestingly, there was also a significant increase in secondary students who included positive stereotypes following the Scientist Spotlights intervention. Such descriptions may stem from biographical resources typically featured in Scientists Spotlight assignments. Do scientists describe themselves using positive stereotypes in order to affiliate with values and expectations of the scientific community? Finally, a small proportion of students actually named scientists, so researchers who are interested in knowing which scientists are known to students might prompt students to list all the scientists they can name as a separate prompt.

Stereotypes oversimplify human complexity, so we considered how to code students' descriptions and scientists based on previous work while factoring in our unique study context. The present and previous studies in Scientist Spotlight assignments derived themes for scientist stereotypes from earlier research (Mead and Metraux, 1957; Dikmenli, 2010). Our coding team interrogated the nuances of stereotypes, the nature of identity, and the importance of intersectionality through our review of students' written responses. We were confronted with a range of figures and descriptions that had not yet been conceived by earlier studies in scientist stereotypes but warranted consideration, names such as "Elon Musk," "Steve Jobs," and "Mark Rober," which we considered to be stereotypical based on the existing rubric's stereotypical descriptors. We also reassessed "Bill Nye the Science Guy," who was previously coded as a Nonstereotypical scientist. Given the overlap in characteristics of Bill Nye and stereotypes in the existing coding rubric, we opted to instead code Bill Nye as a stereotypical scientist, which aligned with another study in scientist stereotypes with undergraduates (Thomas *et al.*, 2006). From examples like these, we anticipate that our understanding of stereotypes will continue to shift over time as media, textbooks, and curricular supplements like the Scientist Spotlight assignments intentionally challenge scientist stereotypes with counter-stereotypes.

A subset of science instructors, teachers, and educators with professional development in inclusive teaching use resources like Scientist Spotlight assignments in their courses to make up for the paucity of diverse representation in textbooks. We have yet to explore the representation of scientists in curricula more broadly. Future studies could explore whether and how science instructors, teachers, and educators intentionally (or unintentionally) represent scientists in their courses, whether or not students engage with metacognitive reflection about these

representations, and how this ultimately shapes students' perceptions of scientist stereotypes.

Science Identity: Secondary Students Increase Their Recognition and Performance/Competence in Science following the Scientist Spotlight Intervention. To facilitate investigations of science identity at larger scales, we explored the PCIR instrument as a possible quantitative tool in the present study. We found that the Scientist Spotlights intervention corresponded to significant increases for students overall in key predictors for student persistence in science majors and careers—Recognition and Performance/Competence. Further, when disaggregated by classes grouped by in-class discussions reported by teachers, we observed significant shifts for students in classes with discussions *both* before and after receiving the assignments. Our findings suggest promising shifts in key aspects of science identity, especially for a broad population of secondary students who might not have otherwise considered a science major or career.

As a service to the community, we also share an emergent finding of the present study—PCIR could be a viable instrument for measuring shifts in aspects of science identity following the Scientist Spotlights intervention. Because systematic qualitative analysis is both training and time intensive, we explored evidence for validity of a pre-existing, quantitative instrument to measure key aspects of science identity. Such a tool could be useful for large-scale studies, but it had yet to be determined whether there would be evidence of validity for this instrument with secondary school students. There was preliminary evidence at the community-college level that the Scientist Spotlights intervention corresponded to shifts in measures of science identity from the PCIR instrument (unpublished data). Our findings from EFA/CFA showed that the PCIR instrument performed as expected with secondary school science students.

One may wonder whether and how Scientist Spotlight assignments in secondary school settings shift students' college career intentions in relation to STEM fields. In a previous study, Dou *et al.* (2019) used PCIR to investigate how students' scores predicted their choice of a STEM major in college. A higher PCIR score corresponded to higher odds that the student pursued a STEM major. Additionally, Dou *et al.* (2019) found that participants who consumed science and science fiction media had higher STEM identity. Could Scientist Spotlight assignments, with resources from popular science media like the Story Collider podcast and TED Talks, play a key role in introducing secondary students to popular science media? Future work could consider longitudinal studies with students in secondary schools, community colleges, and university partnership programs to assess the long-term impacts of Scientist Spotlight assignments on students' major and career choices. Colleges with a higher proportion of regional students would be ideal settings.

Exploring Differences between the Present Study and Prior Work: Why Did an Even Higher Proportion of Secondary Students Not Shift in Their Relatability to Scientists?

After reviewing results of the present study, we speculated as to why the proportion of secondary students who related to scientists did not increase as much as previously observed in college-level biology. Further, the effect size was small. Even though the shifts were significant following the Scientist Spotlights intervention, the proportion of secondary school students who shifted from "disagree" to "agree" with the Relatability prompt was not as great as one might predict based on previous findings.

Let us consider the proportion of secondary students who agreed with the Relatability prompt relative to the proportion of undergraduates who agreed in Aranda *et al.* (2021). In the present study, 27% of secondary students (N = 797) shifted from "disagree" to "agree" for the Relatability prompt following at least three Scientist Spotlight assignments. However, in the previous study across participating courses in the biology department, 36% of undergraduates (n = 752) shifted from "disagree" to "agree" following at least three Scientist Spotlight assignments (Aranda *et al.*, 2021).

To investigate why there was an attenuated increase in the proportion of secondary students who agreed with the Relatability prompt relative to undergraduates in Aranda *et al.* (2021), our study explored two possible explanatory variables: 1) the differences in implementation strategies and 2) the demographics of the participants. We found that implementation strategies, but not demographics, appeared to correspond to variations in student outcomes. In the following sections, we review our findings from the present study and address what factors could be considered for future research on Scientist Spotlight assignments.

Implementation of Scientist Spotlight Assignments with In-Class Discussions Corresponded with the Proportion of Students Relating to Scientists. How an instructor introduces and implements an assignment could shape how students perceive it. Given that the instructors in Aranda *et al.* (2021) had received department-wide professional development in scientific teaching (Owens *et al.*, 2018), we wondered whether the proportion of secondary students agreeing with the Relatability prompt was lower than previously observed for undergraduates due to differences in instructor implementation. Our findings suggest that implementation—namely, engaging students with in-class discussions—could be a driving factor in differential student outcomes from the Scientist Spotlights intervention.

To explore these potential variations in implementation, we designed a retrospective reflection survey to ask teacher-researchers whether or not they engaged students in a discussion before or after implementing Scientist Spotlight assignments. Further, teacher-researchers shared whether they gave the assignments in-person, as homework, or a mix of both. It may be important to note that, for the present study, Scientist Spotlight assignments were mostly implemented in school districts that went remote for the academic year due to the coronavirus pandemic. Regardless, some teachers had large pre–post shifts in the proportion of secondary students who agreed with the Relatability prompt, suggesting that remote instruction was not an absolute impediment to student shifts for this intervention.

From this analysis, we found that in-class discussions can make a difference for students' shifts in relatability following these assignments. Teachers who reported no class discussions had the lowest proportion of students shift from "disagree" to "agree" for the Relatability prompt; however, teachers who

reported that they engaged students with in-class discussions both before and after implementing Scientist Spotlight assignments had highest proportion of students shift. There did not appear to be differences for students in classes with teachers who reportedly implemented Scientist Spotlight assignments in-class and/or as homework.

These findings suggest that direct analyses of the nature of in-class discussions, instructor language, and other measurements of classroom observation are needed. Schinske *et al*. (2016) hypothesized that noncontent instructor language, or Instructor Talk, could be a useful tool for securing student buy-in for Scientist Spotlight assignments that teach science content through counter-stereotypical scientists' stories. Future research could use in-class recordings to measure Instructor Talk (Seidel *et al.*, 2015; Harrison *et al.*, 2019) or follow the Classroom Discourse Observation Protocol (Kranzfelder *et al.*, 2019) to assess variation of implementation strategies for Scientist Spotlight assignments across instructors. Beyond observational studies, the outcomes of *intentional* implementation strategies with Scientist Spotlights could be assessed in conjunction with discussion-based activities. For example, the Ecological-Belonging Intervention fosters student reflection and discussion and has been shown to eliminate performance gaps in "threatening classroom contexts" (Hammarlund *et al.*, 2022). Findings from this work could support evidence-based professional development for instructors using Scientist Spotlight assignments.

Demographics of Secondary Students Did Not Predict Shifts in Students' Relatability to Scientists following the Scientist Spotlights Intervention. We considered that another possible explanation for the attenuated pre–post shifts in secondary students' agreement with the Relatability prompt was differences in the demographics of secondary students. However, we observed significant pre–post increases in the proportion of SOC and White students, as well as PEER and non-PEER, who reported knowing an important scientist to whom they could personally relate. Previous work on Scientist Spotlight assignments assessed their impact on the outcomes of college biology students who are mostly from racial and ethnic backgrounds that have been excluded from the sciences (e.g., Schinske *et al.*, 2016; Aranda *et al.*, 2021). While the secondary school setting is a key difference in the present study, we also noticed that the majority of secondary student participants identified as White (~20%) and/or Asian (~55%). Because disaggregated student responses still showed significant pre–post shifts in agreement across racial and ethnic demographics, we doubted that the demographic composition of students was the driving factor for attenuated shifts in the present study relative to previous studies. Importantly, we must acknowledge that like any social construct, there is heterogeneity hidden by these racial and ethnic labels including but not limited to the country of origin, generation of immigration, linguistic capital, and familial access to educational opportunities.

Like race and ethnicity, binary gender demographic groups—women and men—also had attenuated but significant pre–post shifts in relatability. It is important to note that the majority of students in the present study self-identified in binary gender categories. Given the disproportionate attrition and minimal attention of trans* and gender-nonconforming individuals from STEM in college (Casper *et al*., 2022a,b; Maloy *et al*., 2022), additional work is needed to amplify perspectives of trans* and gender-nonconforming students with regard to assignments like Scientist Spotlights.

These findings across demographic groups suggest that Scientist Spotlight assignments may be a worthwhile intervention not only for expanding possible science selves for underrepresented groups but also for challenging implicit assumptions held by majority groups about scientist stereotypes. One qualitative study showed secondary students' perceptions of scientists in the media corresponded with how they explained their perceptions of their peers (and themselves) in relation to science (Braden, 2020). When it comes to challenging implicit associations based on prejudice and stereotypes, a meta-analysis suggested that interventions focusing on counter-stereotypical exemplars are the most promising, but more research is needed (FitzGerald *et al.*, 2019). One may wonder how challenging implicit assumptions with counter-stereotypical exemplars might also challenge ongoing discrimination in STEM fields, as discussed in the *Introduction*. Future research might consider a deeper exploration of student perspectives through interviews or focus groups to uncover how these implicit assumptions are reinforced or challenged when learning about counter-stereotypical scientists in STEM curricula and in popular science media, as featured in Scientist Spotlight written assignments.

Other Possible Explanations for the Attenuated Shift in Secondary Students' Relatability to Scientists: The Selection of Scientist Spotlight Assignments and Student Authorship. In addition to possible explanations for differential student outcomes already explored, there are also a couple of key differences between the present study in secondary schools and prior work in undergraduate biology. Here we address the manner in which Scientist Spotlight assignments were selected and the potential role of student authorship to garner student buy-in for assignments like these.

Which Scientist Spotlight assignments are chosen to be included in a course might affect student buy-in for these assignments. In the study of Aranda *et al*. (2021), undergraduates in a service-learning course collaborated with instructors across the biology department to develop Scientist Spotlight assignments that the instructors used in their courses. In the present study, secondary school teacher-researchers chose from an online database (www.scientistspotlights.org), which included the Scientist Spotlight assignments developed from the student–instructor collaborations of the previous study. The former method of Scientist Spotlight selection might garner student buy-in because the assignments were developed by near-peers and customized for that particular college course. Since our study was completed, teacher-researchers reviewed, and our team updated, the database of Scientist Spotlight assignments to be even more accessible for the high school level. Future studies could develop Scientist Spotlight assignments in partnership with secondary school teachers and students to ensure the assignments are appropriate for the grade level and desired learning outcomes.

Student authorship is another notable phenomenon that differentiates the previous study from the present one. In the previous study, upper-division biology students who authored their

own Scientist Spotlight assignments had significant shifts in their relatability to and descriptions about scientists (Aranda *et al.*, 2021). These shifts were much larger than what we observed in the present study. Therefore, one might also consider the role of student authorship in fostering students' relatability to scientists. When students are tasked with finding scientists that they would like to feature in Scientist Spotlight assignments, they have the opportunity to identify a scientist to whom they *already* relate. Therefore, future work could investigate whether there are greater shifts in secondary students' relatability to scientists after authoring their own Scientist Spotlight assignments.

LIMITATIONS AND FUTURE DIRECTIONS

There are some limitations of present study that could be addressed in future research. First, we noticed for the Stereotypes Prompt that the vast majority of students did not name specific scientists, despite being asked by the prompt to do so "if possible." To better understand how secondary science students conceptualize and remember scientists, future work might pose a separate question to invite students to list the names of all the scientists who come to mind. Second, there was no comparison group in the present study, and other variables could have affected student shifts in science identity. Another limitation of the present study was the examples in our coding rubric for the Stereotypes prompt. The categories and corresponding examples were informed by previous work on scientist stereotypes; however, stereotypes are culturally dependent and ever evolving. Future scholarship in scientist stereotypes might explore which long-standing stereotypes exist for a given group of participants, with special consideration to multicultural student populations. Additionally, the present study includes students from a convenient sample of teacher-researchers who graciously agreed to participate during the pandemic, so there is an opportunity to explore Scientist Spotlight implementation across a broader range of geographies and demographics, as well as in traditional classroom environments. Finally, students were grouped for In-Class Discussions based on teachers' responses on a retrospective reflection survey, which was administered following the execution of our study. Therefore, we cannot empirically confirm the consistency in how teacher-researchers managed in-class discussions. Future work might include recordings of the in-class discussions to better understand the nature of these discussions and the Instructor Talk that primes students.

CONCLUSION

Our study aimed to respond to recommendations from countless colleagues to implement Scientist Spotlight assignments in secondary school settings. Across a dozen school districts, we evaluated whether the impact of these written assignments on multiple measures of science identity work for secondary science students, as previously observed for college-level biology students. Indeed, findings from our work suggest that Scientist Spotlight assignments correspond to increases in secondary students' relatability to scientists. After receiving at least three Scientist Spotlight assignments, students overall used more nonstereotypical descriptors to describe the types of people that do science. Nonstereotypical descriptors are more inclusive, so when secondary science students write, "Anyone can be a scientist," they seem to be expanding the definitional boundaries of who can be a scientist to include *anyone*—even, quite possibly, themselves. Additionally, we found evidence of validity for a quantitative instrument of science identity (PCIR), which appears to be a viable instrument for large-scale studies in Scientist Spotlight assignments. Finally, initial evidence suggests that in-class discussions can amplify the impact of assignments like Scientist Spotlights on multiple measures of science identity. Our results suggest there would be great value in exploring how science instructors introduce Scientist Spotlight assignments and how instructors facilitate conversations about these assignments with their students. With thoughtful implementation, Scientist Spotlight assignments can innovate science curricula by increasing the representation of scientists to reflect our students and our society and by advancing efforts to promote diversity, equity, and inclusion in the sciences.

ACKNOWLEDGMENTS

This work would not have been possible without the generous support of our funder. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award no. R25GM129823. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We express our sincerest gratitude to the teacher-researchers who signed up to contribute to this initiative in the midst of ongoing racial injustices, extreme weather events like wildfires, and the global pandemic. Being a teacher in today's world is challenging enough, and this scholarship is indebted to your interest and enthusiasm to contribute. We also sincerely appreciate the students for participating, families for consenting, and principals for offering letters of support for teachers to collaborate. The growing Scientist Spotlights database is made possible through the contributions of student-authors/emerging scientists at Foothill Community College, San Francisco State University, and many other institutions. Also, we thank the San Francisco State University Science Education Partnership and Assessment Laboratory (SEPAL) colleagues, both past and present, who offered key insights throughout the development of this project—including Patricia Devera, Lucia Hau, Elleanor Pangilinan, Leticia Rivero, Francisco Sanchez-Vasquez, Katherine White, Maria José Pastor-Infantas, Jarred dela Cruz, Dianna Cazarez, and Kate Gelinas, as well as Dr. Andrea Burton and professors Maurina Aranda, Mallory Rice, and Brie Tripp, who offered constructive comments on the paper. We also acknowledge Dr. Heather Perkins for guidance on using the PCIR instrument in this novel context and Dr. Jennifer Kurushima for invaluable feedback on the article.

REFERENCES

Appiah, K. A. (2020, June 18). The case for capitalizing the "B" in Black. *The Atlantic*. Retrieved July 15, 2022, from www.theatlantic.com/ideas/archive/2020/06/time-to-capitalize-blackand-white/613159/

Aranda, M. L., Diaz, M., Mena, L. G., Ortiz, J. I., Rivera-Nolan, C., Sanchez, D. C., ... & Boorstin, S. N. (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.

Asai, D. J. (2020). Race matters. *Cell*, *181*(4), 754–757.

Bandalos, D. L., & Finney, S. J. (2010). Exploratory and confirmatory factor analysis. In Hancock, G. R., & Mueller, R. O. (Eds.), *Quantitative Methods in the Social and Behavioral Sciences: A Guide for Researchers and Reviewers* (pp. 98–122). New York, NY: Routledge.

Barton, A. C., Carlone, H., Cook, M., Wong, J., & Brickhouse, W. S (2008). Seeing and Supporting Identity Development in Science Education. In Kanselaar, G, Jonker, V., Kirschner, P. A., & Prins, F. J. (Eds.), *International Perspectives in the Learning Sciences: Cre8ing a learning world. Proceedings of the Eighth International Conference for the Learning Sciences – ICLS 2008* Vol. 3 (pp. 214–220). Utrecht, The Netherlands: International Society of the Learning Sciences.

Boone, H. N., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension*, *50*(2), 1–5.

Braden, S. K. (2020). "Scientists can't really talk to people": Unpacking students' metacommentary on the racialized and gendered science nerd trope. *International Journal of Multicultural Education*, *22*(2), 87–108. https://doi.org/10.18251/ijme.v22i2.2439

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.

Brickhouse, N. W., Lowery, P., & Schultz, K. (2000). What kind of a girl does science? The construction of school science identities. *Journal of Research in Science Teaching*, *37*(5), 441–458.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In Hoyle, R. H. (Ed.), *Handbook of Structural Equation Modeling* (pp. 261–276). New York, NY: Guilford Publications.

Carlone, H. B. (2004). The cultural production of science in reform-based physics: Girls' access, participation, and resistance. *Journal of Research in Science Teaching*, *41*(4), 392–414.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science*, *44*(8), 1187–1218.

Casper, A. M. A., Atadero, R. A., & Fuselier, L. C. (2022b). Revealing the queer-spectrum in STEM through robust demographic data collection in undergraduate engineering and computer science courses at four institutions. *PLoS ONE*, *17*(3), e0264267. https://doi.org/10.1371/journal.pone.0264267

Casper, A. M. A., Rebolledo, N., Lane, A. K., Jude, L., & Eddy, S. L. (2022a). "It's completely erasure": A qualitative exploration of experiences of transgender, nonbinary, gender nonconforming, and questioning students in biology courses. *CBE—Life Sciences Education*, *21*(4), ar69.

Chacón-Díaz, L. B. (2022). A textbook analysis to uncover the hidden contributors of science and mathematics. *Science & Education*, *31*(1), 193–211. https://doi.org/10.1007/s11191-021-00246-7

Chambers, D. W. (1983). Stereotypic images of the scientist: The Draw-A-Scientist test. *Science Education*, *67*(2), 255–265.

Dikmenli, M. (2010). Undergraduate biology students' representations of science and the scientist. *College Student Journal*, *44*, 579–588.

Dou, R., Hazari, Z., Dabney, K., Sonnert, G., & Sadler, P. (2019). Early informal STEM experiences and STEM identity: The importance of talking science. *Science Education*, *103*(3), 623–637. https://doi.org/10.1002/sce.21499

Dunkel, C., & Kerpelman, J. (Eds.) (2006). *Possible selves: Theory, research and applications*. Hauppauge, NY: Nova Publishers.

Fagerland, M. W. (2012). t-tests, non-parametric tests, and large studies—A paradox of statistical practice? *BMC Medical Research Methodology*, *12*(1), 78. https://doi.org/10.1186/1471-2288-12-78

Fagerland, M. W., & Sandvik, L. (2009). Performance of five two-sample location tests for skewed distributions with unequal variances. *Contemporary Clinical Trials*, *30*(5), 490–496. https://doi.org/10.1016/j.cct.2009.06.007

Ferguson, S. L., & Lezotte, S. M. (2020). Exploring the state of science stereotypes: Systematic review and meta-analysis of the Draw-A-Scientist Checklist. *School Science and Mathematics*, *120*(1), 55–65. https://doi.org/10.1111/ssm.12382

FitzGerald, C., Martin, A., Berner, D., & Hurst, S. (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: A systematic review. *BMC Psychology*, *7*(1), 29. https://doi.org/10.1186/s40359-019-0299-7

Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2016). Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice. *Journal of Engineering Education*, *105*(2), 312–340. https://doi.org/10.1002/jee.20118

Grimalt-Álvaro, C., Couso, D., Boixadera-Planas, E., & Godec, S. (2022). "I see myself as a STEM person": Exploring high school students' self-identification with STEM. *Journal of Research in Science Teaching*, *59*(5), 720–745. https://doi.org/10.1002/tea.21742

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

Harrison, C. D., Nguyen, T. A., Seidel, S. B., Escobedo, A. M., Hartman, C., Lam, K., … & Tanner, K. D. (2019). Investigating Instructor Talk in novel contexts: Widespread use, unexpected categories, and an emergent sampling strategy. *CBE—Life Sciences Education*, *18*(3), ar47. https://doi.org/10.1187/cbe.18-10-0215

Hill, P.W., McQuillan, J., Talbert, E., Spiegel, A., Gauthier, G. R., & Diamond, J. (2017). Science possible selves and the desire to be a scientist: Mindsets, gender bias, and confidence during early adolescence. *Social Sciences*, *6*(2), article 2. https://doi.org/10.3390/socsci6020055

Hock, M. F., Deshler, D. D., & Schumaker, J. B. (2006). Enhancing student motivation through the pursuit of possible selves. In Dunkel, C., & Kerpelman, J. (Eds.), *Possible selves: Theory, research and applications* (pp. 205–221). Hauppauge, NY: Nova Publishers.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Knekta, E., Runyon, C., & Eddy, S. (2019). One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research. *CBE—Life Sciences Education*, *18*(1), rm1. https://doi.org/10.1187/cbe.18-04-0064

Korkmaz, S., Göksülük, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *R Journal*, *6*(2).

Kranzfelder, P., Bankers-Fulbright, J. L., García-Ojeda, M. E., Melloy, M., Mohammed, S., & Warfa, A.-R. M. (2019). The Classroom Discourse Observation Protocol (CDOP): A quantitative method for characterizing teacher discourse moves in undergraduate STEM learning environments. *PLoS ONE*, *14*(7), e0219019.

Macdonald, R. H., Beane, R. J., Baer, E. M. D., Eddy, P. L., Emerson, N. R., Hodder, J., … & Ormand, C. J. (2019). Accelerating change: The power of faculty change agents to promote diversity and inclusive teaching practices. *Journal of Geoscience Education*, *67*(4), 330–339. https://doi.org/10.1080/10899995.2019.1624679

Maloy, J., Kwapisz, M. B., & Hughes, B. E. (2022). Factors influencing retention of transgender and gender nonconforming students in undergraduate STEM majors. *CBE—Life Sciences Education*, *21*(1), ar13.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*(9), 954.

Mead, M., & Metraux, R. (1957). Image of the Scientist among High-School Students. *Science*, *126*(3270), 384–390.

Mills, L. A. (2014). *Possible science selves: Informal learning and the career interest development process*. *Paper presented at the 11 th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA) from October 25–27, 2014 at Porto, Portugal*.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2023). *Advancing Antiracism, Diversity, Equity, and Inclusion in STEMM Organizations: Beyond Broadening Participation*. Washington, DC: The National Academies Press.

Ormand, C. J., Heather Macdonald, R., Hodder, J., Bragg, D. D., Baer, E. M. D., & Eddy, P. (2021). Making departments diverse, equitable, and inclusive: Engaging colleagues in departmental transformation through discussion groups committed to action. *Journal of Geoscience Education*, *70*(3), 1–12. https://doi.org/10.1080/10899995.2021.1989980

Owens, M. T., Trujillo, G., Seidel, S. B., Harrison, C. D., Farrar, K. M., Benton, H. P., … & Tanner, K. D. (2018). Collectively improving our teaching: attempting biology department–wide professional development in scientific teaching. *CBE—Life Sciences Education*, *17*(1), ar2. https://doi.org/10.1187/cbe.17-06-0106

Oyserman, D., Bybee, D., Terry, K., & Hart-Johnson, T. (2004). Possible selves as roadmaps. *Journal of Research in Personality*, *38*(2), 130–149. https://doi.org/10.1016/S0092-6566(03)00057-6

Pringle, R. M., & McLaughlin, C. A. (2014). Preparing science teachers for diversity: Integrating the contributions of scientists from underrepresented groups in the middle school science curriculum. In Atwater, M. M., Russell, M., & Butler, M. B. (Eds.), *Multicultural science*

education: Preparing teachers for equity and social justice (pp. 193–208). Netherlands: Springer. https://doi.org/10.1007/978-94-007-7651-7_12

Provenzo, E. F., Jr., Shaver, A. N., & Bello, M. (Eds.) (2010). *The textbook as discourse: Sociocultural dimensions of American schoolbooks*. New York, NY: Routledge. Retrieved April 29, 2022, from http://ebookcentral.proquest.com/lib/sfsu/detail.action?docID=667927

Raiche, G., Magis, D., & Raiche, M. G. (2020). *Package 'nFactors'. R package version 2.4.1.1*. Retrieved April 4, 2022, from https://cran.r-project.org/web/packages/nFactors/nFactors.pdf

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved April 4, 2022, from www.r-project.org

Revelle, W. R. (2017). *Package 'psych'. R package version 2.2.9*. Retrieved March 8, 2022, from https://cran.r-project.org/web/packages/psych/psych.pdf

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., … & Barendse, M. (2023). Package 'lavaan'. *R package version 0.6–14*. Retrieved March 8, 2023, from https://cran.r-project.org/web/packages/lavaan/lavaan.pdf

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

Seidel, S. B., Reggi, A. L., Schinske, J. N., Burrus, L. W., & Tanner, K. D. (2015). Beyond the biology: A systematic investigation of noncontent instructor talk in an introductory biology course. *CBE—Life Sciences Education*, *14*(4), ar43. https://doi.org/10.1187/cbe.15-03-0049

Shah, M., Foster, A., Talafian, H., Barany, A., & Petrovich, M. E. (2021). Facilitating and interpreting high school students' identity exploration trajectories in STEM. *Journal of Experimental Education*, *89*(3), 541–559. https://doi.org/10.1080/00220973.2020.1808941

Simpson, A., & Bouhafa, Y. (2020). Youths' and adults' identity in STEM: A systematic literature review. *Journal for STEM Education Research*, *3*(2), 167–194. https://doi.org/10.1007/s41979-020-00034-y

Slemp, K. (2020). *Latino, Latina, Latin@, Latine, and Latinx: Gender inclusive oral expression in Spanish* (Master's thesis). *Electronic Thesis and Dissertation Repository, University of Western Ontario, Canada*. Retrieved April 22, 2022, from https://ir.lib.uwo.ca/etd/7297

Sumrall, W. J. (1995). Reasons for the perceived images of scientists by race and gender of students in grades 1–7. *School Science and Mathematics*, *95*(2), 83–90.

Symington, D., & Spurling, H. (1990). The "Draw a Scientist Test": Interpreting the data. *Research in Science & Technological Education*, *8*(1), 75–77. https://doi.org/10.1080/0263514900080107

Thomas, M. D., Henley, T. B., & Snell, C. M. (2006). The Draw a Scientist test: A different population and a somewhat different story. *College Student Journal*, *40*(1), 140–148.

van Tuijl, C., & van der Molen, J. H. W. (2016). Study choice and career development in STEM fields: An overview and integration of the research. *International Journal of Technology and Design Education*, *26*(2), 159–183. https://doi.org/10.1007/s10798-015-9308-1

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), 25. https://doi.org/10.1128/jmbe.v21i1.2013