

Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines

Sarah L. Eddy^{1,*} and Sara E. Brownell²

¹*The University of Texas at Austin, College of Natural Sciences, 1 University Station G2500, Austin, Texas 78712, USA*

²*School of Life Sciences, Life Sciences Building 204, Post Office Box 874501 Tempe, Arizona 85287-4501, USA*

(Received 30 January 2015; published 1 August 2016)

[This paper is part of the Focused Collection on Gender in Physics.] This focused collection explores inequalities in the experiences of women in physics. Yet, it is important for researchers to also be aware of and draw insights from common patterns in the experiences of women across science, technology, engineering and mathematics (STEM) disciplines. Here, we review studies on gender disparities across college STEM on measures that have been correlated with retention. These include disparities in academic performance, engagement, self-efficacy, belonging, and identity. We argue that observable factors such as persistence, performance, and engagement can inform researchers about what populations are disadvantaged in a STEM classroom or program, but we need to measure underlying mechanisms to understand how these inequalities arise. We present a framework that helps connect larger sociocultural factors, including stereotypes and gendered socialization, to student affect and observable behaviors in STEM contexts. We highlight four mechanisms that demonstrate how sociocultural factors could impact women in STEM classrooms and majors. We end with a set of recommendations for how we can more holistically evaluate the experiences of women in STEM to help mitigate the underlying inequities instead of applying a quick fix.

DOI: [10.1103/PhysRevPhysEducRes.12.020106](https://doi.org/10.1103/PhysRevPhysEducRes.12.020106)

I. INTRODUCTION

Throughout this Focused Collection, researchers are discussing the challenges facing women¹ in physics in the hope of generating productive conversations and interventions that change these experiences. These conversations can be strengthened by extending their scope beyond physics to include findings from other science,

technology, engineering and mathematics (STEM) disciplines. Extensive work on the experiences of women has been done across STEM disciplines and several common patterns of inequalities have emerged. In many STEM disciplines, women are underrepresented at the undergraduate level and this underrepresentation is often exacerbated at the graduate, postdoctoral, tenure-track faculty, and practicing STEM professional level [5,6]. Even if women persist to become STEM faculty and working professionals, these women tend to have lower visibility in their fields than their male colleagues: they participate less than males at academic conferences, particularly in more prestigious settings [7,8], and are less likely than males to hold the more prestigious first or last author position on manuscripts [9]. These broad findings illustrate that many issues facing women in physics transcend disciplinary boundaries, making the experiences of women across STEM fields highly relevant to the experiences of women in physics.

Increasing the retention of women in undergraduate STEM disciplines may be a first step in the complex and daunting task of improving the representation of women as STEM professionals. Currently, women switch out of undergraduate STEM majors at a higher rate than their male peers [6,10] and this may be especially true for female underrepresented minority students [11]. These women are not leaving college, but moving to other non-STEM majors, indicating that there is something about

*Corresponding author.
sarah.eddy@austin.utexas.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

¹It is important to acknowledge that although we focus on males and females in this article, gender identity exists on a spectrum and that more than two genders exist in the human experience. In addition, even individuals that identify as women do not all share a common experience. Gender is a complicated identity based on individuals' internal experiences of who they are. Thus, individuals can vary in the degree to which they identify with a particular gender, how important gender is to their identity, the gender roles associated with their gender, and how their gender identity influences their experience in different settings such as a classroom [1–3]. In addition, gender is only one of a multitude of social identities that make up who we are and how we react in certain settings. Other identities, such as race and/or ethnicity, also influence a students' experiences of their gender in the classroom [4].

college-level STEM that is driving them away. Even women who persist through their undergraduate majors enter graduate school at lower rates than their male peers in 75% of STEM disciplines [12]. If we can prevent undergraduate STEM programs from being a selective filter favoring males [13], we may be able to increase the number of women at the professional level in STEM fields.

However, we argue that only focusing on the numbers of women in STEM may be shortsighted as college retention data alone underestimates and potentially masks gender disparities that exist across STEM. These more subtle gender disparities could impact students in ways that influence the long-term retention of women in graduate school, postdoctoral training, and academic positions. Further, retention numbers alone do not provide any insights into the underlying mechanisms that contribute to this drop-off in numbers of women in STEM.

We propose that we need to look deeply into the underlying factors that influence retention across STEM disciplines in addition to physics in search of commonalities and potential solutions. Therefore, we critically reviewed the published literature to identify gender disparities in multiple factors shown to influence retention in STEM. Although our review is not comprehensive, these findings give us insights into the mechanisms underlying retention disparities. Further, we present a framework that STEM instructors and researchers could use to critically examine relationships among the different types of gender inequalities that may exist in their classrooms. Through this framework, we describe a set of possible ultimate causes of these gender inequalities that are situated in theories from social psychology. Targeting these deep underlying causes could be a more effective strategy for mitigating gender disparities in college physics classrooms.

Critical evaluation of the current literature on gender disparities in college STEM also highlights the pressing need for additional research. It is our hope that this article can provide a deeper conceptual understanding of the complexities facing females in STEM and stimulate a robust dialogue among researchers and educators across disciplines on ways to improve the experience of female students in college STEM classrooms. Furthermore, we urge discipline-based education researchers to move beyond the natural sciences to begin conversations and collaborations with social psychologists and sociologists to better study some of these underlying mechanisms. We have much to learn and share with each other.

II. GENDER INEQUALITIES IN COLLEGE STEM BEYOND RETENTION

The number of students enrolled in a college STEM major or course is only a superficial diagnosis of gender inequalities in that discipline or course. While low numerical representation is indicative of a problem, the lack of a numerical gender gap does not necessarily indicate an

absence of gender inequalities. For example, in biology, where 60% of undergraduate students on average are women, simply examining the numerical representation of female students may lead to the assumption that there are no gender inequalities. Yet, we have observed gender disparities in achievement, participation, and comfort in multiple biology classrooms (see Eddy, Brownell, and Wenderoth [14]). Additionally, the numerical underrepresentation of a group of students does not tell us anything about why there are so few students. In this article, we focus on possible underlying factors such as academic achievement, engagement, and affective measures. While we are not claiming that these are the only factors that matter, we present them as examples of growing evidence for the types of gender inequalities that extend beyond coarse-grained enrollment numbers. For each factor, we also comment on aspects of data collection, analysis, and interpretation that may impact the findings.

Here, we only included papers that explored these factors in the context of the undergraduate STEM experience to document gender gaps specific to this population. We did not include studies of gender gaps in K–12 or in other college settings (e.g., studies done with nonmajors outside of STEM classrooms). Although there have been many studies on math and science attitudes of undergraduate women not majoring in STEM, we chose not to include these studies because there is evidence that many women in STEM differ in critical ways from women not in STEM. For example, many STEM women hold different beliefs about STEM related abilities and gender than non-STEM women [15]; and there is some evidence that STEM women tend to be less gender normative in terms of gender roles [16–18]. For example, a study of computer science majors found that female majors did not differ from male majors on a sex roles scale, but did significantly differ from female nonmajors [16]. Thus, to be conservative, we only included studies done with STEM majors or in STEM classrooms.

The papers collected here constitute a representative, although not comprehensive, sample of work in the area of gender inequalities in STEM. A challenge with a synthesis of gender inequalities is that while some studies report gender gaps in undergraduate STEM as the main finding, other papers include this information as a supplemental section in a course or program evaluation paper. This makes it difficult to collate the findings on gender inequalities because many papers do not even use gender as a key word and thus are not archived by search engines in this way. We have done our best to find as many papers as possible, but want to acknowledge that we likely have not captured everything that has been published.

A. Academic performance

A common way to measure the academic climate for women in college STEM courses is to compare the academic performance of females to that of males.

Academic performance across STEM is a predictor of retention [19,20], so a consistent achievement gap between students could be an important factor in why more women than men leave STEM. However, one study comparing women who leave STEM to women who remain in STEM did not find a significant difference in academic performance between these two populations [21]. This finding suggests that although academic performance is important for the retention for all students, it may not, on its own, explain the gender differences in persistence in STEM.

In the studies that we reviewed from a range of STEM disciplines including engineering, biology, chemistry, physics and math, we do not see a consistent gender gap in performance across or within disciplines (Table I). There is also no consistent pattern comparing lower division to upper division courses. Overall there may be a trend toward men outperforming women, but these results are mixed. A closer examination of the statistical analyses and outcome variables used in these studies suggests a more nuanced pattern of gender inequalities and a need for more research in this area.

The majority of studies examining differences in academic performance are not randomized trials. Instead, they are observational studies of students who have self-selected into classrooms. In these quasirandom study systems, it is critical for researchers to control for differences between students. Lack of random assignment of groups to a treatment can produce systematic biases that confound study results [39,40]. For example, an important systematic bias exists in the population of males and females who attend college: college-bound women systematically have higher high school GPAs than men that attend college [41]. This implies that, all else being equal, females ought to do better in their college classes than males. Thus, even if males and females appear as though they are performing equally in a course, an achievement gap may actually be observed if students were matched by an indicator of ability (such as prior performance in high school or college courses).

In the majority of studies that controlled for a measure of student ability, males seem to outperform female students (Table I). The only achievement gap that favored females after controlling for student ability was found in a physics class for nonmajors [27]. In the studies that do not control for ability, there is either no achievement gap or one in favor of females (Table I). This pattern holds true even within a discipline: in biology, the one study of introductory biology classes that did not use a control for student ability did not find an achievement gap [24], but our own work in introductory biology courses that controlled for student performance in their prior college classes found an achievement gap that favored males [14].

In addition to controlling for academic differences between students, the outcome variable that is measured can influence whether or not researchers observe an achievement gap. For example, Kost-Smith, Pollock, and Finkelstein [42] observed an achievement gap in physics

classrooms when they used exam performance as their outcome variable, but not when they used overall class performance because women tended to do better than men on homework assignments. Even in studies that only use exams as their outcome variable, the characteristics of exam items used can influence the presence and size of observed gender gaps on those exams (Wright *et al.* [43], McCullough [28]). In addition, how researchers choose to measure learning can impact achievement gaps. For example, Willoughby and Metz [44] tested multiple ways of measuring learning gains and found achievement gaps between males and females when using normalized gains (post-test minus pretest divided by pretest), but not other gain calculations. Thus, researchers should carefully consider what outcome variable is most informative to their research question. For example, course grade can be useful for determining who can continue in the major if there are grade cutoffs, but exam or concept inventory performance may be a better indicator of actual student learning.

Finally, while academic performance gives an account of how a student is achieving, this is still a relatively coarse-grain approach that does not provide insights into why differences in performance might exist. The gap in achievement could stem from a number of different factors and achievement gaps in different settings may have non-overlapping causes. In addition, the lack of an achievement gap may lead to incorrect assumptions that gender equality has been achieved when in reality other unexamined gender disparities remain that could impact long-term retention in STEM.

B. Engagement

While less well studied than academic performance, there is also evidence for gender inequalities in engagement in college STEM (Table I). A student's engagement in STEM can include participation in the classroom (e.g., answering an instructor's question or talking in small groups), as well as participation in activities that are STEM related outside of the classroom (e.g., study groups, clubs, research). Participation and engagement within one's discipline has been shown to predict retention [45,46]. In addition, participation can be an important indicator of other affective measures such as a student's perception of and anxiety about performance in a STEM course [35,47]. In this article, we focus on participation in the classroom setting because most studies have been done in this context.

Although there is a rich literature exploring participation differences based on gender in non-STEM college classrooms [48–50], the evidence for gendered patterns of in-class participation in STEM classrooms is sparse (Table II). Studies using student self-reported participation in college STEM classrooms have shown that female students across disciplines report lower participation or less comfort with participation when compared with males [35,36]. This appears to be true even in STEM disciplines with large

TABLE I. A sample of papers reporting on academic performance and classroom engagement of males and females in two or more STEM classes or with STEM undergraduates.

| Discipline | Student level: Lower level (LL, 1st or 2nd year), Upper level (UL, 3rd + year) | Sample size: No. Students (S), classes (C), and institutions (I) | Type of students: Majors (Ma), Nonmajors (N), Mixed majors (B), Unknown (?) | Outcome measured | Direction of bias: Favors females (F) or favors males (M) or equity (=) | Reference |
|---|--|---|---|--|--|-----------|
| Across STEM (Includes Biology, Physical Sciences, and Engineering) | UL | S:? C:NA I: 197 | Ma | Cumulative GPA in Major at Graduation [®] | F | [22] |
| Across STEM (Includes Biology, Engineering, Physical sciences, Math, and Computer science) | UL | S: 1508 C: NA I: NA | Ma | Performance on GRE [®] : Verbal Quantitative Analytical Cumulative GPA in Major at Graduation [®] | M M = M or = (biology)** | [23] |
| Biology | LL | S: 269 C: 2 I: 1 | Ma | Course Performance [®] | = | [24] |
| | LL | S: ~5000 C:23 I: 1 | B | Exam Performance | M | [14] |
| | UL | S: <5728 C: 5 (sampled multiple times) I: 1 | Ma | Course Performance | M or =** | [25] |
| Chemistry or biochemistry | LL | S: 218 C: 1 I: 1 | Ma | Course Performance [®] | = or F** | [24] |
| | UL | S: <11 313 C:8 I: 1 | Ma | Course Performance | M or =** | [25] |
| | | S: 11 114 C: 2 (sampled multiple times) I: 1 | B | Course Performance | M | [26] |
| Physics | LL | S: 191 C: 4 I: 1 | Ma | Course Performance [®] | = | [24] |
| | LL | S:663 C: 8 I: 8 | N | Course Performance | F | [27] |
| | LL | S: 1265 C: 8 I: 8 | B | Course Performance | M | [27] |
| | ? | S: ~175 C: NA I:? | N | Force Concept Inventory | M | [28] |
| Engineering | LL | S: 121 C: 5 I: 1 | Ma | Course performance [®] | = or M** | [29] |
| | LL | S: 228 C: NA I: 1 | Ma | Cumulative first year GPA in major | = or M** | [30] |
| Math | LL | S: 322 C: 1 (sampled multiple times) I: 1 | ? | Course performance [®] | F | [31] |
| | LL | S: 269 C: ? I: 1 | ? | Course performance [®] | F | [32] |
| | LL | S: 326 C: 15 I: 10 | N | Course performance [®] | = | [33] |

(Table continued)

Academic performance*: Results vary

TABLE I. (Continued)

| | | | | | | | |
|--|--|-----------|---|---|--|---|-------------------------|
| Participation inequities: Women participate at lower rates in STEM classes | Across STEM (Includes Biology, Physics, Social Sciences, and Humanities) | LL and UL | S: >17 per class C: 26 I: 1 | B | Classroom observations of participation in front of whole class | M | [34] |
| | Across STEM (Includes Biology, Computer science, and Engineering) | LL and UL | S: >20 per class C: 7 I: 1 | ? | Survey: Classroom Experience Questionnaire | M | [35] |
| | Across STEM (Includes Biology, Chemistry, and Engineering) | LL | S: 205 C: NA I: 1 | ? | Survey: Intimidation in small groups | M | [36] |
| | Across STEM 'Natural sciences' | ? | S: ? C: 16 I: 1 | ? | Classroom observations of whole class participation | M | [37] |
| | Biology | LL | S: >150 per class C:3 (multiple iterations of each) I:1 | B | Classroom observations of whole class participation | M | [14] |
| | | LL | S: 624 C: 1 (multiple iterations) I: 1 | B | Survey: Intimidation in small groups and in whole class discussions whole class small groups | M | Eddy <i>et al.</i> [38] |

** If 2 patterns are listed, the first is the predominate one and the second is less common.

® These studies did not control for differences between students.

numbers of females: we found self-reported participation differences between males and females in multiple large introductory biology classes [38]. In this study, men felt more comfortable answering instructor questions in front of the whole class and were more likely to prefer to take on the role of leader in small group work compared to women. However, self-report is a measure of the subjective experiences of students, it does not necessarily always correlate with observed differences. Thus, it is important for studies to also directly measure the participation in STEM classrooms.

Fewer studies have observed participation in college STEM classrooms, although the studies that do exist are large in scope, span the range of common class sizes, and include lower and upper division courses. Sternglanz and Lyberger-Fick [37] is the earliest study that we found on student participation in STEM college classrooms. The researchers in this study observed 16 "natural science" courses that had predominately male enrollment and were taught by male instructors. They found that male undergraduates were more likely than female undergraduates to answer instructor questions and initiate interactions with the instructor in class. Society and gender roles have changed since this study was published in 1977, yet even in more modern studies we find similar patterns of gendered engagement. A more recent study explored student engagement in

classroom discussions in 26 lower and upper division humanities and science (biology and physics) courses at one institution [34]. These courses were all generally small in size relative to most science classes (the average class size for their "large" classes was 35 students) and, unfortunately, researchers did not disaggregate the science and humanities courses in their analysis. Across both classroom types, they found that gender inequalities in participation grew over time in college, with males speaking up more as they progressed from introductory to upper-division courses. To our knowledge, the most recent observational study of gender gaps was work by us in 26 introductory biology classrooms at an R1 university [14]. The class sizes in this study ranged from 159 to approximately 500 students. Although small group work occurred frequently in many of these classrooms, the study focused only on student responses to instructor posed questions to the whole class. Even though the class composition was 60% female, female voices were heard answering questions only 40% of the time. A similar pattern was seen for students asking questions of the instructor in class, although there was more variability in this among classes so the effect was not significant. Taken together these three studies suggest a gender gap in participation that consistently favored males in STEM classrooms. Further work across more classrooms

and STEM disciplines will be necessary to assess the robustness of this pattern.

A major limitation of these observational studies is that it is impossible to determine the cause of the gender difference in participation from observation alone. Class participation in these studies involves two people: the instructor and the student. Thus, gender disparities could be due to instructors not calling on females as much as males (as has been seen in some K–12 classrooms [70,71]), female students not feeling confident enough to raise their hands to answer a question [72], or a combination of both. These are two different causes of participation differences and would necessitate different interventions. More sophisticated observational studies that go beyond simple counts of who is answering questions could better disentangle these influences. For example, documenting who is raising their hands, who shouts out answers, and who is called on could distinguish between student and instructor contributions to this participation bias. Researchers could also track individual students to determine if the gendered pattern is a result of high participation by a few individuals or is more evenly distributed among the majority of students in the class. Finally, the quality of both instructor questions and student answers could be assessed. It has been demonstrated in K–12 literature that instructors hold different expectations for some populations relative to others [73] and this may impact who they call on to answer a certain type of question.

A second limitation of these studies is that they only focus on one context of classroom participation: participation in front of the whole class. As more classrooms move to student-centered instruction [74], students are increasingly working with each other in pairs or small groups. Gendered patterns of participation and the quality of this participation in these contexts need to be explored as well. Chi and Wylie's [75] Interactive, Constructive, Active, Passive (ICAP) framework is one way that researchers could conceptualize the quality of student interactions in small groups and explore whether gender influences student contributions. In this framework, interactive participation (where students take turns talking and build on each other's ideas) is hypothesized to lead to greater student learning than constructive participation (where one student in the group does all the talking). Monitoring participation in small group work is challenging for instructors, particularly in large classes where it is nearly impossible for them to engage with every group. However, these contexts need to be explored to fully characterize gender inequalities in participation that may be meaningful for retention in STEM.

C. Affective measures: self-efficacy, belonging, and science identity

Undergraduate students hold beliefs about their own abilities in STEM, their peers' abilities, and about STEM

disciplines in general, all of which can influence how students perceive their relationship to their major. These attitudes and experiences can be important predictors of retention in STEM. For example, a survey across two cohorts of STEM majors at one university found that students' perceptions of their own fit within a STEM discipline influenced the retention of women in the major, especially in male dominated fields [76]. Below we outline three affective measures that are related to STEM retention: self-efficacy, belonging, and disciplinary identity. These measures are not inclusive of all the possible affective measures impacting retention, but they are the ones that have been most studied in STEM classrooms. The amount of research that has been done on each measure in STEM settings varies greatly and conclusions that can be drawn are preliminary. However, themes are beginning to emerge (Table II).

1. Self-efficacy

Self-efficacy is a measure of the strength of one's belief that one can complete a task or goal. There are many types of self-efficacy, but the studies in this review focus on academic self-efficacy: one's confidence in one's ability to master academic subjects and coursework [77]. A meta-analysis of 38 studies relating self-efficacy to achievement and 18 studies relating self-efficacy to persistence in academic settings found that self-efficacy significantly impacted both measures [78]. This effect remained even when researchers restricted the sample to the subset of studies done in a college context. Thus, identifying differences in self-efficacy between males and females in STEM classrooms has the potential to help explain differences in retention.

We found eight studies that documented self-efficacy at the beginning of a student's college career. Five out of the eight studies focused on engineering and demonstrated that male first year undergraduates had higher self-efficacy than females. Besterfield-Sacre and colleagues' [57] study spanning 15 institutions was the most robust. It showed that incoming male engineering majors rated their confidence in their basic engineering skills higher than female engineering majors across all institutions (although this difference was not significant at three institutions). This pattern of higher confidence for men remained even after a year of coursework in engineering, although the pattern was weaker: gender differences in self-efficacy were only present at seven of the institutions. In studies at single institutions, Felder *et al.* [29] and Jagacinski [58] both found patterns of lower academic self-efficacy for first year female engineering majors relative to male engineering majors. Concannon and Barrow [59], on the other hand, found no difference between male and female engineering majors in some measures of self-efficacy, but did find gender differences in self-efficacy in engineering specific abilities.

TABLE II. Survey of articles reporting on gender gaps in affective measures in STEM majors or STEM courses.

| Discipline | Student level: Lower level (LL, 1st or 2nd year), Upper level (UL, 3rd+ year), Longitudinal (LO) | Sample size: No. students (S), classes (C), and institutions (I) | Type of students: Majors (Ma), Nonmajors (N), Mixed majors (B), Unknown (?) | Direction of bias: Favors females (F) or favors males (M) or equity (=) | Reference | |
|--|--|--|---|---|-----------|------|
| <i>Self-efficacy inequalities: Women report lower confidence in their STEM abilities</i> | Across STEM (Includes Biology, Chemistry, Computer science, Math, Physics, Geology, Engineering, and Technology) | UL | S: 352 C: 49 I: 1 | Ma | M | [51] |
| | Across STEM (unknown disciplines) | UL | S:169 C: NA I: 1 | Ma | M | [52] |
| | Computer science | ? | S: 56 C: NA I: 1 | Ma | M | [53] |
| | Engineering | LO | S: 1983 C: NA I: 1 | Ma | M | [54] |
| | | ? | S: 147 C: 5 I: 1 | Ma | M | [55] |
| | | LL | S:288 C: NA I: 1 | Ma | M | [56] |
| | | LL and LO | S: 6180 C: NA I:15 | Ma | M | [57] |
| | | LL | S:121 C: NA I: 1 | Ma | M | [29] |
| | | LL | S: 117 C: 1 I: 1 | Ma | M | [58] |
| | | LL and UL | S: 726 C: NA I: 1 | Ma | = or M | [59] |
| Chemistry | LL and LO | S: 84 C: NA I: 1 | Ma | M | [60] | |
| | LL and LO | S: 384 C: 1 I: 1 | B | M or = | [61] | |
| Math | LL | S: 326 C: 15 I: 10 | N | M | [33] | |
| <i>Belonging inequalities: Women less likely to feel they belong in STEM</i> | Across STEM | ? | S: 47 C: NA I: 1 | Ma | M or = | [62] |
| | Across STEM | UL (graduate students) | S: 149 C: NA I: 1 | Ma | F | [63] |
| | Across STEM | LL | S: 53 C: NA I: 1 | Ma | M | [64] |
| | Physics | LL | S: 2177 C:5 I:1 | B | M | [65] |

(Table continued)

TABLE II. (Continued)

| | | | | | | |
|--|--|----|--------------------------|----|-----------------------|------|
| Science identity: Women identify less strongly with STEM | Across STEM (Includes Biology, Chemistry, and Physics) | LL | S:7505 C:? I: 40 | Ma | F or = = M or = | [66] |
| | Across STEM (Biology, Chemistry, and Physics) | LL | S:1799 C: ? I: 1 | Ma | = or M | [67] |
| | Across STEM (Chemistry and Engineering) | UL | S: 320 C:24 I:? | Ma | M | [68] |
| | Across STEM | UL | S: 251 C: NA I: >1 | Ma | = | [69] |

Studies on academic self-efficacy with first year students in other STEM disciplines have also shown that men hold higher self-efficacy than women. In a study that tracked self-efficacy in a first year chemistry course for mixed-science majors at a single institution, researchers found that the gender gap in self-efficacy was mediated by a race by gender interaction: no gap was present between male and female students who were black or white, but there was a significant gap between Asian American and Latina(o) males and females, which decreased over the term [61]. This study is of note as it was the only study in our sample that included prior academic achievement as a control in their model when predicting self-efficacy. Thus, the gaps in this study were between males and females of matched ability in terms of their performance on the math SAT. It is also the only study in our sample that explored the interaction between gender and race or ethnicity. In contrast to this study, Dalgety and Coll [60] found that the gender gap in self-efficacy in a chemistry class for majors actually increased over the term. Finally, a study with nonmajors in a college algebra courses across 10 institutions found that males had higher self-efficacy than females [33].

Studies on advanced undergraduates (third and fourth years) across STEM reveal that the academic self-efficacy of men remains higher than that of women even beyond the first year of college. A study of third year students across STEM disciplines who were all enrolled in a supplemental program found that female STEM majors reported significantly lower academic self-efficacy compared with male STEM majors, particularly on study skills, test-taking skills, and coping with test anxiety [52]. A study across 49 different upper division STEM courses found that women reported lower confidence in their academic ability in courses in their discipline compared to men [51]. Finally, in a study comparing computer science (CS) majors to nonmajors, female CS majors reported lower self-efficacy

than male majors and had lower self-efficacy than even male non-CS majors on a measure of self-efficacy related to working with computers [53].

Another type of self-efficacy that likely impacts retention is the belief that one can excel at a job in one's chosen field. Two studies have examined this type of self-efficacy in engineering. Cech *et al.* [56] measured students' self-efficacy at four different institutions during their first year of college and then four years later to determine whether they persisted in an engineering major. They found that a student's belief that he or she could excel in an engineering career significantly predicted retention in engineering, and female engineering majors had lower confidence in this ability to excel than males. The second study, conducted at one institution, also found that male engineering majors had higher career self-efficacy than female majors [59].

Each of the studies described above only captures a snapshot of self-efficacy and not how an individual student's self-efficacy changes over the duration of their STEM major. It is plausible that the gender gaps in self-efficacy seen in advanced undergraduates are actually smaller than they were initially, but cross-sectional and longitudinal studies are needed to address this question. However, if students are not individually followed through time in longitudinal studies, it is still not possible to disentangle two explanations for the observed changes in self-efficacy: (i) a student's self-efficacy increases over time or (ii) students with lower self-efficacy dropped out of the major, so their lower self-efficacy scores are not averaged in the advanced undergraduate group. We did find one study that followed almost 2000 students longitudinally as they advanced through the college engineering sequence [54]. These researchers showed that the gap in self-efficacy between female and male engineering majors actually widened as students advance through their major [54]. However, they did not individually track students in their

analysis, which could be done by accounting for repeated measures, limiting the interpretation of these results.

In summary, these studies suggest that women in STEM report lower self-efficacy beliefs than men at the beginning of their college career, but more longitudinal studies are needed to further probe this area.

2. *Belonging*

Belonging is the experience of feeling accepted as a member of a group. Sense of belonging can be in reference to multiple different groups; one can belong to a discipline as a whole (e.g., belong in science), belong in a major (e.g., physics major), belonging in the classroom or other communities (e.g., Physics 101 class), or even belong to a small working group in the context of the larger class (e.g., one's lab group).

The connection between belonging and retention has been demonstrated in two studies using very different methods. In one study, female undergraduate students in a college level calculus class who reported a higher sense of belonging to math were more likely to express an intent to pursue further math courses beyond the calculus series [79]. The second study used an observational approach to measure belonging [80]. The researchers postulated that the amount of time that faculty in STEM departments spent socializing with colleagues was a proxy for their sense of belonging and that this time would predict how engaged a faculty member felt in his or her job (the measure of intent to persist in this study). Forty-five faculty members wore recording devices and researchers found that women who engaged in fewer social conversations with their male colleagues felt more disengaged from their job [80]. Fewer social conversations seemed to have no impact on men. Thus, there is some evidence that sense of belonging has the possibility to influence persistence in STEM.

There have been few studies on belonging in a discipline done with students in STEM classrooms and the results of these studies are mixed (Table II). Stout *et al.* [65] documented belonging in five calculus-based physics classrooms at one institution and found that male students had a stronger sense of belonging to physics than female students. A second study of college level calculus classes at another institution found that a sense of belonging to math for all students declined over the semester and the belonging of female undergraduates was significantly lower than male undergraduates in the middle of the course [79].

We found three additional studies that examined sense of belonging of STEM students in contexts other than their discipline. One study experimentally manipulated the gender ratio of participants at a hypothetical STEM conference and found that female STEM majors felt a lower sense of belonging at that conference when the gender ratio favored men [62]. This study implies that women may feel a lower sense of belonging in STEM environments where they perceive that they are in the

minority. A second study focused on the sense of belonging at an institution for STEM transfer students [64]. Female STEM transfer students experienced a greater difference between their sense of belonging and their desired level of belonging at this institution than did male STEM transfer students. Another study of almost 150 first year graduate students across STEM disciplines found that at the start of their first year, female students reported a higher sense of belonging in their program than their male peers [63]. Although at first glance this may seem like a positive, it actually suggests that only women with a heightened sense of belonging consider graduate school whereas men with both higher and lower senses of belonging pursue graduate school. Thus, belonging may be a limiting factor for women in pursuing STEM careers beyond their bachelor's degree.

3. *Disciplinary identity*

Disciplinary identity is an indicator of the extent to which students perceive their own identity to be aligned with the identity of practitioners in their discipline. The relationship between this disciplinary identity and retention has been implied by one study in STEM that spanned three populations: advanced undergraduates, graduate students, and post-doctoral scholars [81]. Within each of these populations, individuals with greater science identity also expressed greater commitment to becoming a research scientist.

Only a few studies of disciplinary identity have been done with STEM students. In general, these studies indicate that there is a relationship between gender and the strength of one's STEM identity (Table II). In a sample of incoming first year students across 40 institutions, students on average reported neutrally to three questions asking 'do you consider yourself a [biology, chemistry, or physics] person', but this response was moderated by gender and race for students intending to pursue a STEM career after college [66]. For these STEM-oriented students, white women identified more with biology than white men, while men and women of color did not differ from each other in biology identity; men and women of all races and ethnicities identified equally with chemistry; and white men had a stronger identification with physics than white women, but men and women of color did not differ from one another in physics. One point of caution in the interpretation of this study's findings is that the same students answered each of the three questions, yet the questions were analyzed as if they were functioning independently. It is possible that a student's response on one question might have influenced his or her response on another (i.e., students may have thought they could only identify with one of the disciplines represented or the order in which the questions were presented could influence their responses on subsequent questions).

A second large study of students in enrolled in introductory courses for majors in biology, chemistry, and physics at one institution revealed that, on average, science identity between male and female students is equal, but this effect is moderated by how strongly one associates with one's gender [67]. They found that female STEM students who perceive gender as more important to their personal identity had weaker science identity. The salience of gender did not impact the science identity of male STEM students.

The results for advanced undergraduates are also mixed. In a study with primarily third year students in engineering and chemistry classes at one institution, male students had a stronger STEM identity than female students [68]. However, a second study using a national sample did not find a significant difference between the STEM identities of male and female students [69]. The students in this second study were recruited from the listserv of a national research meeting, though, so they may already have gone through a selective filter before the study because undergraduates with low science identity are unlikely to attend such a meeting.

4. Limitations of affective measures

One of the largest limitations of these studies is that the majority of them only sampled students once. Without following students through time, it is impossible to determine the impact of a degree program or course. It is interesting that the one study in the sample that did follow students longitudinally through their major revealed that the gap in the affective measure between male and female students grew over time [54]. This implies that there was something about the college experience that is disadvantaging female students and could potentially lead to lower retention.

In addition, it is important to think carefully about how one measures these affective factors. While Likert-scale surveys were used to assess self-efficacy, belonging, and science identity in the studies described above, it is important for researchers to be critical of the validity of surveys prior to administering them. While well-designed published surveys are validated in a particular context, questions are only truly valid for the study population that they were tested on, and it is important for researchers interested in administering a previously designed survey to determine if the questions are valid in their specific context (e.g., questions designed for graduate students may be misinterpreted by undergraduates; see Benson [82] and Netemeyer, Bearden, and Sharma [83]). Another way of investigating these affective measures is through interviews or focus groups, but these qualitative methods often take more time to both collect and analyze data and the smaller sample sizes may introduce more biases in interpretations. As such, affective measures are sensitive to the specific questions and context in which the data was collected and studies should be evaluated with these criteria in mind.

D. Conclusion: Gender gaps in STEM

Overall, there is evidence for gender gaps in multiple factors across a diverse range of college STEM disciplines, classrooms, and programs. Although the gender inequalities presented above are not meant to be comprehensive and clearly more work needs to be done before we can draw any definitive conclusions, preliminary patterns are emerging. Although there is not a strong pattern for academic performance and too few studies on participations to make any broad conclusions, there do seem to be consistent gendered patterns in affective measures. In the studies collected here, women enter college STEM expressing lower self-efficacy and science identity than men and as they advance in their college career, the gaps do not close. The consistencies in affective measures as compared to performance may be due to the fact that academic performance is measured in many diverse ways (e.g., post-test, exam performance, concept inventories) and is not as well defined a construct as some of the affective measures. Another possibility is that students are so motivated to achieve that they are able to overcome differences in affective measures to perform equally [84].

Our second conclusion is a call for a shift towards more systematic data collection. Currently data collection in STEM is done in a somewhat piecemeal way and many of the studies in this review report on only one type of data. More studies that collect multiple measures, longitudinally on the same sample of students, would provide us a more nuanced picture of student experiences in our majors and classrooms.

III. UNDERLYING SOURCES OF GENDER INEQUALITIES IN COLLEGE STEM

Studies that have collected multiple measures on the same students reveal that affective measures such as self-efficacy, science identity, and belonging tend to be correlated with each other and change over time, even within one semester (Good *et al.* [79]). Similarly in multiple studies, these measures have been shown to be correlated with achievement and with other affective measures not mentioned above (e.g., values orientation, goal orientation, and many more Hernandez *et al.* [85] and Perez, Cromley, and Kaplan [86]). These correlations imply that there are possibly underlying processes that influence all of these measures together. In the next section, we adapt a motivational model for career interest developed by Wang and Degol [87] to help illustrate these relationships and the proposed underlying factors that could unite them.

A. Adapted framework that incorporates observed factors, psychological factors, and underlying sociocultural factors

Wang and Degol's framework explores how motivation influences career choices. It draws on studies of student

experiences in K–12 and connects three levels of influence on student career choices (Fig. 1). This framework seems to be supported by the college level literature reported here, although the evidence from college STEM is more tentative as there have been fewer studies. Students in most STEM classrooms have already made their initial career choice by choosing their major. Thus, in our adaption of this model we replace career choice with the decision to persist in STEM.

In Wang and Degol’s framework, the *observable factors* related to career choice are student behaviors such as engagement, course enrollment, and achievement in the domain of one’s intended career. These factors align with our factors of performance and engagement in STEM classrooms.

The next level of Wang and Degol’s framework includes prior demonstrated ability and psychological factors that can affect persistence directly and/or indirectly through the observable factors. These *psychological factors* include the invisible factors of self-efficacy, belonging, and science identity that we described earlier, as well as many others. The idea behind the connection between these psychological factors and the observable factors is that students will be more motivated to achieve, engage, and persist if they feel as though they can be successful and if the value they see in their major is worth the cost of remaining in it (Wang *et al.* [87]). Self-efficacy is one measure of their belief that they can be successful in STEM; science identity and belonging

are two measures of the value they see in their STEM major.

Finally, Wang and Degol propose that underlying *sociocultural factors* influence these psychological factors. These sociocultural factors are larger cultural norms and values that students and instructors bring into the classroom. They influence students in many ways including their ability to see themselves as scientists (science identity), how other students treat them (which can influence belonging), and their beliefs that they can be successful in STEM (self-efficacy). These sociocultural factors may be the underlying sources of many of the gender inequalities documented in college STEM.

We make two additional modifications to Wang and Degol’s model to make it better fit the experiences of women as they progress through their STEM majors. First, Wang and Degol focus explicitly on forces that motivate students, but it is also important for us to consider that there are events outside of a student’s control that can impact the observable factors. An example of such an event might be if a female student raises her hand in class, but the instructor does not call on her and calls on a male instead. In this case, the student has not achieved the outcome of talking in class because an outside factor (the instructor) has obstructed her. To account for these external factors, we added an arrow from the underlying sociocultural factors to the observable outcomes. In addition, the feedback that students receive on their

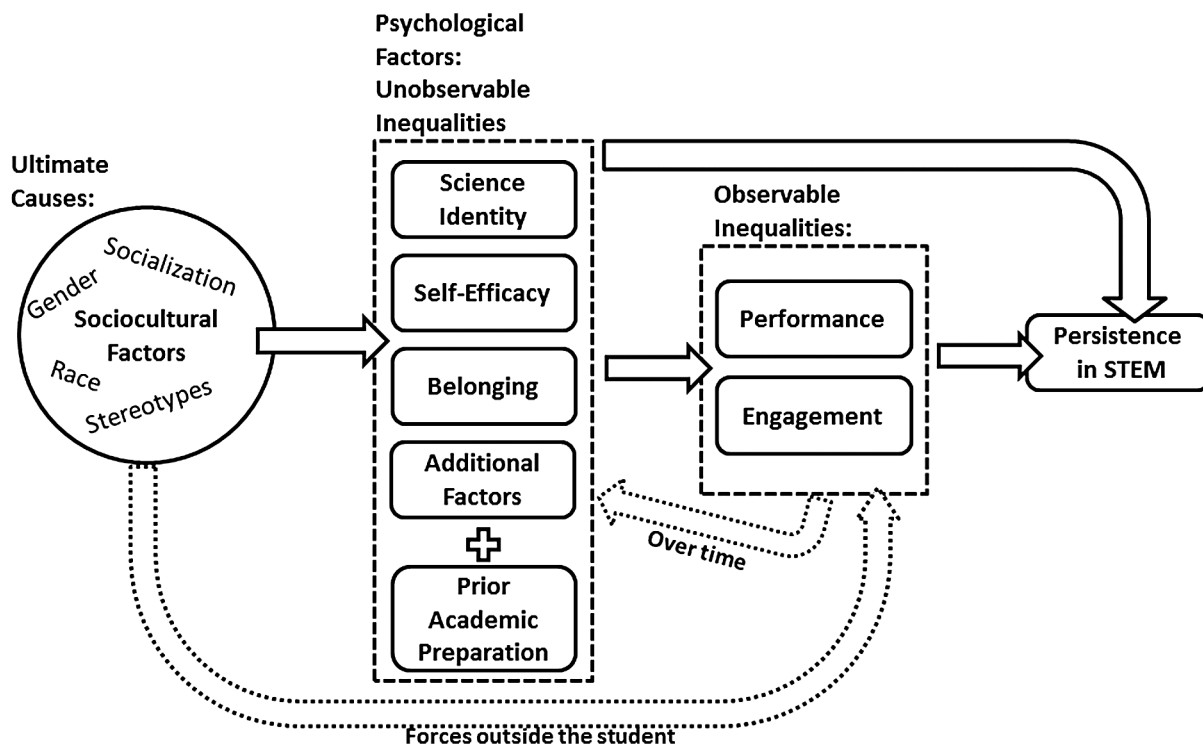


FIG. 1. Model describing how gender gaps in affective and observational measures can impact persistence in STEM and the underlying sociocultural factors influencing all of them. Modified from Wang and Degol’s [87] model for understanding career choice.

performance or engagement as they progress through an undergraduate degree will influence their subsequent perceptions of self-efficacy, belonging, and identity. To illustrate this longitudinal feedback loop, we added an arrow back from the observable factors towards the underlying factors. We propose this model as an approximation of the interactions among factors that may influence a college STEM student's retention.

We hope instructors and researchers can use this framework to begin to conceptualize gender gaps in their classrooms. While the observable factors of persistence, performance, and engagement can only provide information about who is affected, the psychological factors and sociocultural factors can both begin to explain *how* some of these observed gender inequalities may have come to exist. This also means that interventions to ameliorate gender gaps in persistence, performance, and/or engagement may only be a metaphorical band-aid to a much deeper problem. If significant changes to the experiences of women in STEM are to be made, this framework posits that the mechanisms surrounding sociocultural factors in STEM classrooms will need to be targeted.

B. Sociocultural factors: Stereotypes and gendered socialization

In this section, we explore four examples of how sociocultural factors could be impacting the persistence of female undergraduates in STEM (Table III). We focus on the impact of stereotypes and gendered socialization and the mechanisms by which they may act in the classroom. We then highlight the evidence that these mechanisms could explain why we see gender gaps in psychological and observable factors.

Although we initially introduce and provide evidence of each sociocultural factor broadly, we only feature articles that assess the impact of these factors on the three populations that are most likely to be relevant to women in STEM: (a) Female undergraduates enrolled in a STEM major; (b) female undergraduates in a STEM course who may or may not be majors; (c) females who identify

positively with math. As most STEM disciplines require at least some degree of comfort with math, we felt that this third population was appropriate to include.

1. Stereotypes about women's ability in STEM can lead to stereotype threat

Few people in society today explicitly state that they believe that men are better at science or math than women, but more subtle versions of this message persist. For example, in television programs popularly watched by middle school aged students, male scientists appear 1.6 times as frequently as female scientists. This disparity in representation is even greater in shows not funded by the National Science Foundation: less than half the scientists portrayed are women [88]. The lack of female scientists portrayed in the media may send the message that science and math are activities for males. Messages like these have impact: in a worldwide comparative study, countries where people hold weaker associations between science and maleness have reduced achievement gaps between men and women on national math and science assessment exams, whereas countries with stronger associations between science and maleness have larger gender gaps [89]. The United States has an average position on the scale of the association between science and maleness, indicating that these stereotypes exist in our culture.

Even undergraduate women who enroll in STEM disciplines continue to be impacted by the stereotypes about women's ability in STEM: a sample of over 300 female STEM majors found these women still associated STEM with maleness [68] and in a smaller study approximately 25% of female STEM majors endorsed the stereotype that men are better at math [1]. The negative impact of these associations for STEM women can be seen in work done by Nosek and Smyth [15]: college women, even those majoring in STEM courses, show a strong association between math and maleness and the strength of this association predicts their performance on the quantitative section of the SAT [11]. Interestingly, college men enrolled in STEM courses show the strongest endorsement of the "men are better at math stereotype." Overall, it seems that the

TABLE III. Evidence for the influence of the underlying processes linking sociocultural factors with the psychological and observable factors for women in STEM. \checkmark indicates that there is evidence in the literature for a particular process to impact that factor. (the number in parentheses) indicates the number of peer reviewed articles supporting this connection. $\checkmark?$ indicates the link has not been explicitly tested but is implied by published studies. Please refer to text for actual citations.

| Underlying process | Psychological factors | | | Observable factors | |
|---|-----------------------|-----------------|------------------|--------------------|-----------------|
| | Belonging | Self-Efficacy | Science Identity | Performance | Persistence |
| Stereotype threat | $\checkmark(3)$ | $\checkmark(5)$ | $\checkmark(4)$ | $\checkmark(13)$ | $\checkmark(7)$ |
| Biases held by others | $\checkmark?$ | $\checkmark?$ | | | |
| Alignment of personal goals with perceived science goals | | | $\checkmark(1)$ | | $\checkmark(2)$ |
| Implicit theories of intelligence | $\checkmark(1)$ | | | $\checkmark(1)$ | $\checkmark(2)$ |

stereotypes associating males with science and math influence the experience of women in college STEM.

Stereotype threat is often used to explain the negative influence that male-centric STEM stereotypes have on women in STEM. Stereotype threat is defined as situational pressure posed by the possibility that poor performance will be judged through the lens of a negative group-relevant stereotype [90]. Some of the original studies characterizing stereotype threat were done with women who resemble STEM women: undergraduate women who liked math and for whom math performance was important to them [91]. In one of these studies, female undergraduates and a matched set of male undergraduates were given an exam with Graduate Research Exam (GRE)-like math questions. Researchers manipulated how women might feel while taking these exams by using one of two verbal prompts: either “Males are known to outperform women on this exam” (stereotype threat treatment) or “Unlike other math exams, this exam does not show gender differences” (no stereotype threat treatment). Women in the no stereotype threat treatment performed almost three times better than the women in the stereotype threat treatment. Thus, women in STEM may be experiencing stereotype threat if they are concerned that their performance will be evaluated based on stereotypes about a woman’s ability in these fields.

There is a strong body of literature suggesting that stereotype threat can impact both the retention and performance of women in STEM. Manipulations of the extent to which women experience stereotype threat have repeatedly been shown to modify women’s self-reported interest in continuing in STEM [1,68,92–95], although long-term studies of persistence are generally lacking (but see Beasley and Fischer [11]). Another robust body of literature demonstrates that stereotype threat decreases STEM women’s performance on math-based exams in artificial experimental settings [91,95–102]. However, we should be somewhat cautious of these findings because only a few investigations have explored the impact of stereotype threat on women’s performance on exams in their disciplines or in college STEM classrooms [30,103,104].

Stereotype threat impacts not only the performance and persistence of women in STEM, but also multiple psychological factors, although the number of studies that explore each factor is small. Many of these studies also demonstrate how decreases in these psychological factors are the proximate mechanisms by which stereotype threat decreases performance or retention. Studies have found that stereotype threat reduces science identity for women in a range of disciplines including biology, chemistry, physics, and engineering and some of these studies also show that this reduction in science identity mediates some of the impact of stereotype threat on persistence [68,93,95,105]. Women under stereotype threat also seem to express reduced self-efficacy across diverse STEM disciplines and this impact on self-efficacy can also mediate some of the

effect of stereotype threat on persistence [1,30,93,95,101]. Finally, sense of belonging in STEM for undergraduate women, but not graduate women, seems to be impacted by stereotype threat and mediates some of the impact of stereotype threat on performance [63,65,79,105]. This disparity between graduates and undergraduates could be because only women who have overcome stereotype threat persist to graduate school. Alternatively, stereotype threat may manifest in psychological factors other than belonging for graduate students. Thus, the mechanism of stereotype threat, which is derived from gendered stereotypes about who is good at math and science, could explain why we see gender gaps in multiple psychological factors and in observable factors in STEM majors.

2. Stereotypes about women’s ability in STEM can lead to biased peers, mentors, or instructors

A second way that math or science-gender stereotypes can impact undergraduate women in STEM is through women’s interactions with those around them who believe these stereotypes. A recent series of high profile papers reminds us that these biases appear to be fairly universal and held by people in STEM at all levels from peers to potential employers. In a study with millennial students who role played a hiring interview environment, people were more likely to hire men for a math-related job than women even when math performance was identical between the “candidates” [106]. Additionally, a study of STEM departments at elite institutions demonstrated that male faculty members employed fewer female than male graduate students and postdoctoral researchers [107]. Finally, a study across STEM indicated that faculty members held subtle gender biases when selecting a research assistant [108]. When faculty members were asked to rate identical CVs assigned either a male or female name, both male and female faculty members ranked the male candidate as more hireable and competent and reported they would pay and mentor the male candidate more than the female candidate. Thus, biases held by others may limit the opportunities offered to women in STEM.

The actions of others can also lower a student’s self-efficacy. One of the major sources of women’s self-efficacy is recognition from others [109]. There is evidence that suggests that women in STEM may not be gaining as much recognition as males. First in a study of three introductory biology courses, men were more likely to name men as knowledgeable about the class material, even after researchers controlled for students’ actual performance in the class (Grunspan *et al.* [110]). In a second study with male engineers, men were more likely to discount ideas that were presented in a female voice compared to ideas presented in a male voice [111]. These seemingly small events may also cumulatively lead to students feeling like they do not belong in an environment [112]. For example, feeling supported by one’s peers (an aspect of belonging)

was the strongest contributor to student persistence in a study of CS majors [113]. This impaired sense of belonging in STEM for women can predict their interest in continuing to pursue STEM careers [76,79].

Examples such as these demonstrate how psychological factors could be impacted by the biases people have about women in STEM. Promoting awareness of these subtle biases, especially among individual interacting with women in STEM, is another possible avenue for increasing the retention of women in STEM fields.

3. *Conflict between personal goals and stereotypes about the practices of science*

Just as students hold gender-STEM stereotypes, they also hold stereotypes about what it means to be a STEM practitioner. One of these messages is that STEM practitioners are socially isolated individuals working on esoteric problems. In two studies with a combined sample of almost 800 undergraduates from one institution, researchers found that students believe that individuals in STEM professions such as engineering, computer science, and environmental sciences are much less likely to work with others or contribute to society than individuals in non-STEM careers such as architecture or business [114,115]. One potential problem with this stereotype is that women, on average, value working with others and contributing to society (from here on called communal goals) more than men value these goals [116,117]. This difference in values is not innate in women and men, but is hypothesized, using social role theory, to be due to the historical emphasis on women being in more caregiving roles. This has led to the female gender role emphasizing more communal traits (reviewed in Diekmann *et al.* [118]).

Thus, the stereotype of the socially isolated everyday activities and purpose of STEM is in conflict with the communal values that, on average, women hold more strongly than men. Evidence of this conflict comes from multiple studies. In a series of interviews with 10 graduate students in atmospheric science, women emphasized communal career goals of service and social impact significantly more than men [119]. Additionally, in a national sample of over 9000 college students, women's interest in helping others was negatively correlated with their persistence in STEM [120]. Finally, Smith *et al.* [93] found that women in STEM who perceived that science could not address communal goals had weaker science identities and this lower science identity predicted their attrition from STEM. Based on these studies, one possible way to address the gender gap in persistence in STEM may be to increase the perception that science can support communal goals. Evidence that this type of intervention could help female students comes from a study by Brown *et al.* [121] that used an intervention to experimentally manipulate female STEM majors into believing STEM careers could fulfill communal

goals and found that this increased their interest in pursuing such a career.

4. *Implicit theories of intelligence and the stereotype of brilliant scientists*

Another example of how gendered socialization could impact the persistence of women in STEM relates to intelligence. Women, on average, are more likely to hold an entity view of intelligence: that intelligence is fixed and at some point one reaches a threshold that cannot be exceeded. Men, on average, are more likely to hold a growth view of intelligence: that as long as you work hard, you can get better [122]. This difference is hypothesized to be derived from differential treatment during childhood: girls, on average, are complemented based on innate ability rather than effort, whereas boys are complemented for the effort they put into tasks [122]. The dangers of holding an entity view include giving up in response to failure and avoiding situations where one is at risk of revealing that one has reached one's limit [123,124]. If science is a field that is seen to require high intelligence, than it may seem like a risky field for individuals with an entity mindset and they may choose to avoid it. Unfortunately, it does seem that many fields in science have the stereotype of requiring high innate ability [125].

Few studies in STEM have used implicit theories of intelligence to explore the experience of women. One study in college level math courses found that students had a more fixed view of math intelligence than general intelligence, but there were no gender differences in this perception [126]. However, the students in this study were in a remedial algebra class and may not reflect students in most undergraduate STEM majors. A second study in a college calculus class found that women who held a fixed view of math ability also had a lower sense of belonging, decreased math performance, and decreased intent to persist in math [79]. Finally, a study across STEM disciplines found that as a field is perceived to require more innate ability rather than hard work, the percent of female Ph.D.s is lower [125]. These studies imply that changing an individual's view of intelligence and its relationship to STEM fields may be a fruitful avenue for increasing the retention of women in STEM.

Conclusion: Underlying mechanisms.—In summary, it is important that researchers and practitioners diagnose and then address the underlying mechanisms that could be influencing the experience of women in STEM. There are many more possible mechanisms than have been laid out in this review and multiple may be occurring in the same classroom. We present these mechanisms because we believe interventions will be most effective if they address the root causes of gender inequalities rather than the superficial issues. By targeting interventions at these underlying mechanisms, we have the potential to impact all the downstream factors such as self-efficacy, belonging,

identity, performance, engagement, and ultimately persistence.

V. RESEARCH RECOMMENDATIONS

This review of research findings on gender inequalities in undergraduate STEM is intended as a starting point for instructors and researchers interested in promoting gender equity in STEM. We offer this framework as a rudder to steer the national conversation away from counting the numbers of women in STEM and toward documenting the more subtle gender gaps that are at the root of those numeric inequalities. We also propose that this framework could help instructors and researchers explore the connections and synergies between the factors impacting the experience of women in undergraduate STEM classrooms.

This framework and the accompanying literature review represent only an initial effort to better define why there are not more women in STEM, and much work remains. We call on instructors and researchers to more systematically catalog and evaluate the experiences of women in college STEM classrooms and programs so that we can better understand the extent to which gender inequalities exist. With a more extensive literature base, we can begin to identify differences and communalities between STEM fields in terms of these inequalities. Currently, STEM fields appear fairly similar in terms of the experiences of women, but this may be due to lack of data. In addition, rather than collecting gender gap data in a piecemeal fashion, we argue that it would be most informative to collect multiple measures on the same pool of students and to do this longitudinally. This type of data collection would enable researchers to identify (a) the factors that change as a student progresses through a class or program of study as well as the direction of that change, and (b) the relative contributions of different factors to performance and/or persistence in STEM.

In addition, we recommend that future research on the sociocultural and psychological factors that may impact women in STEM takes place in actual STEM classrooms. Currently, few studies explore these factors in the specific context of undergraduate STEM and results may be very different in these settings where performance measures (e.g., grades) have real impacts and students are highly motivated to achieve [84]. As STEM practitioners and discipline-based education researchers are not generally trained in psychology or sociology, it may be a more effective and efficient approach to collaborate with sociologists and social psychologists. Sociocultural factors are complicated and nuanced and expertise in the underlying theories is essential to interpreting how these factors are influencing the experience of women in STEM and for generating effective interventions. Conversely, social psychologists argue the mechanisms they have identified, like stereotype threat, are important factors in STEM classrooms, yet the majority of their studies are conducted

using psychology students or in psychology labs—a much different context from large introductory science courses. Research partnerships that span the social sciences and natural sciences are likely to be fruitful for both parties, but more importantly are likely to result in a greater impact on the next generation of female STEM scholars.

Our final suggestion is a greater emphasis on intersectionality in studies of women in STEM. Gender is only one of a myriad of identities that a student brings into the classroom and all of these identities impact the experience of students. A cis-gendered black woman from an upper-middle class background in a STEM classroom may have a very different experience than a trans-gendered Asian American student from a working class background. These intersections of identity have not been deeply explored in the discipline-based education literature. It is important for researchers to be aware that most of the studies presented here were done at primarily white institutions and, thus, the results are likely driven by the experiences of white women. Yet, the experiences of white women may not capture the experiences of women of other races and ethnicities. For example, in a study comparing math-gender stereotypes held by white and black women, white women held a much stronger belief that males are better at math than black women [127]. This means a black woman's susceptibility to stereotype threat based on gender may be different in the classroom than a white woman's. A second example is the experience of double jeopardy: women of color confront stereotypes about STEM in two social identities: gender and race. Thus, women of color may be more likely to encounter situations where they are made to feel like they do not belong in STEM. In a recent study of 60 female scientists of color, 100% said they had encountered gender bias [128] while another study on black female scientists found 80% reported encountering racial bias [129]. It is essential that research efforts begin to account for both gender and race or ethnicity, in addition to other social identities. In this review there were a few studies that acknowledged and explored the intersection of gender and race, but much more work needs to be done.

VI. CONCLUSION

Through this article we hope to emphasize that gender inequalities are occurring beyond retention and can be found in achievement, participation, and affective measures. Further, all of these measures could be impacted by underlying sociocultural factors that students and instructors bring into the classroom. To effectively ameliorate gender disparities in college STEM, we propose that we need to address the root causes of these disparities. To do so, we need to engage in dialogue with other disciplines spanning the natural and social sciences to probe common factors that influence gender disparities at the collegiate level. Working together, rather than as single disciplines, we hope to move one step closer to gender equality.

ACKNOWLEDGMENTS

The authors thank A. Huang for help with the literature review. We also thank M. P. Wenderoth and two

anonymous reviewers for their influential comments on earlier versions of this manuscript.

S. E. and S. B. have contributed equally to this work.

-
- [1] T. Schmader, M. Johns, and M. Barquissau, The Costs of Accepting Gender Differences: The Role of Stereotype Endorsement in Women's Experience in the Math Domain, *Sex Roles* **50**, 835 (2004).
- [2] B. A. Nosek, M. Banaji, and A. G. Greenwald, Harvesting implicit group attitudes and beliefs from a demonstration web site, *Gr. Dyn. Theory, Res. Pract.* **6**, 101 (2002).
- [3] K. A. Lane, J. X. Goh, and E. Driver-Linn, Implicit science stereotypes mediate the relationship between gender and academic participation, *Sex Roles* **66**, 220 (2012).
- [4] M. Ong, C. Wright, L. L. Espinosa, and G. Orfield, Inside the double bind: a synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics, *Harv. Educ. Rev.* **81**, 172 (2011).
- [5] C. Hill, C. Corbett, and A. St. Rose, Why So Few? Women in Science, Technology, Engineering, and Mathematics, *Why So Few? Women in Science, Technology, Engineering, and Mathematics* (American Association of University Women, Washington, DC, 2010).
- [6] National Science Foundation, National Center for Science and Engineering Statistics, Special Report (5-31) (2009).
- [7] T. M. Jones, K. V. Fanson, R. Lanfear, M. R. E. Symonds, and M. Higgie, Gender differences in conference presentations: a consequence of self-selection?, *PeerJ* **2**, e627 (2014).
- [8] L. A. Isbell, T. P. Young, and A. H. Harcourt, Stag parties linger: Continued gender bias in a female-rich scientific discipline, *PLoS One* **7**, e49682 (2012).
- [9] V. Lariviere, C. Ni, Y. Gingras, B. Cronin, and C. Sugimoto, Global gender disparities in science, *Nature (London)* **504**, 211 (2013).
- [10] L. Dickson, Race and gender differences in college major choice, *Ann. Am. Acad. Pol. Soc. Sci.* **627**, 108 (2010).
- [11] M. A. Beasley and M. J. Fischer, Why they leave: the impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors, *Soc. Psychol. Educ.* **15**, 427 (2012).
- [12] A. Gillen and C. Tanenbaum, Exploring Gender Imbalance Among STEM Doctoral Degree Recipients Overview (2014).
- [13] J. C. Blickenstaff, Women and science careers: Leaky pipeline or gender filter?, *Gender Educ.* **17**, 369 (2005).
- [14] S. L. Eddy, S. E. Brownell, and M. P. Wenderoth, Gender gaps in achievement and participation in multiple introductory biology classrooms, *CBE Life Sci. Educ.* **13**, 478 (2014).
- [15] B. A. Nosek and F. L. Smyth, Implicit social cognitions predict sex differences in math engagement and achievement, *Am. Educ. Res. J.* **48**, 1125 (2011).
- [16] S. Byer and S. Haller, Gender differences and intragender differences in computer science students: Are female CS majors more similar to male CS majors or female non-majors?, *J. Women Minorities Sci. Eng.* **12**, 337 (2006).
- [17] E. M. Trauth, Odd girl out: An individual differences perspective on women in the IT profession, *Inf. Technol. People* **15**, 98 (2002).
- [18] P. L. Casserly and D. Rock, Factors related to young women's persistence and achievement in advanced placement mathematics, in *Women and Mathematics: Balancing the equation*, edited by S. F. Chipman, L. R. Brush, and D. M. Wilson (Routledge, New York, 1985), pp. 225–248.
- [19] J. Price, The effect of instructor race and gender on student persistence in STEM fields, *Econ. Educ. Rev.* **29**, 901 (2010).
- [20] I. Blackman and I. G. N. Darmawan, Graduate-entry medical student variables that predict academic and clinical achievement, *Int. Educ. J.* **4**, 30 (2004).
- [21] E. Seymour, The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account, *Sci. Educ.* **79**, 437 (1995).
- [22] G. Sonnert and M. F. Fox, Women, men, and academic performance in science and engineering: The gender difference in undergraduate grade point averages, *J. Higher Educ.* **83**, 73 (2012).
- [23] J. Grandy, Gender and Ethnic Differences among Science and Engineering Majors: Experiences, Achievements, and Expectations, *Gender and Ethnic Differences among Science and Engineering Majors: Experiences, Achievements, and Expectations* (Princeton, New Jersey, 1994).
- [24] S. Lauer, J. Momsen, E. Offerdahl, M. Kryjevskaja, W. Christensen, and L. Montplaisir, Stereotyped: Investigating gender in introductory science courses, *CBE Life Sci. Educ.* **12**, 30 (2013).
- [25] L. R. Creech and R. D. Sweeder, Analysis of student performance in large-enrollment life science courses, *CBE Life Sci. Educ.* **11**, 386 (2012).
- [26] M. M. Rauschenberger and R. D. Sweeder, Gender performance differences in biochemistry, *Biochem. Mol. Biol. Educ.* **38**, 380 (2010).
- [27] R. H. Tai and P. M. Sadler, Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA, *Int. J. Sci. Educ.* **23**, 1017 (2001).
- [28] L. McCullough, Gender, context, and physics assessment, *J. Int. Womens. Stud.* **5**, 20 (2004).

- [29] R. M. Felder, G. N. Felder, M. Mauney, C. E. Hamrin, Jr, and E. J. Dietz, A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes, *J. Eng. Educ.* **84**, 151 (1995).
- [30] G. M. Walton, C. Logel, J. M. Peach, S. J. Spencer, and P. Mark, Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering, *J. Educ. Psychol.* **107**, 468 (2014).
- [31] C. I. Brooks, Superiority of women in statistics achievement, *Teach. Psychol.* **14**, 45 (1987).
- [32] C. I. Brooks and J. E. Mercincavage, Grades for men and women in college courses taught by women, *Teach. Psychol.* **18**, 47 (1991).
- [33] M. L. Peters, Examining the relationships among classroom climate, self-efficacy, and achievement in undergraduate mathematics: A multi-level analysis, *Int. J. Sci. Math. Educ.* **11**, 459 (2013).
- [34] L. M. Fritschner, Inside the undergraduate college classroom: Faculty and students differ on the meaning of student participation, *J. Higher Educ.* **71**, 342 (2000).
- [35] G. Crombie, S. W. Pyke, N. Silverthorn, A. Jones, and S. Piccinin, Students’ perceptions of their classroom participation and instructor as a function of gender and context, *J. Higher Educ.* **74**, 51 (2003).
- [36] M. Micari and D. Drane, Intimidation in small learning groups: The roles of social-comparison concern, comfort, and individual characteristics in student academic outcomes, *Act. Learn. High. Educ.* **12**, 175 (2011).
- [37] S. H. Sternglanz and S. Lyberger-ficek, Sex differences in student-teacher interactions in the college classroom, *Sex Roles* **3**, 345 (1977).
- [38] S. L. Eddy, S. E. Brownell, P. Thummaphan, M.-C. Lan, and M. P. Wenderoth, Caution, student experience may vary: social identities impact a student’s experience in peer discussions, *CBE Life Sci. Educ.* **14**, 1 (2015).
- [39] R. Theobald and S. Freeman, Is it the intervention or the students? using linear regression to control for student characteristics in undergraduate STEM education research., *CBE Life Sci. Educ.* **13**, 41 (2014).
- [40] S. E. Brownell, M. J. Kloser, T. Fukami, and R. J. Shavelson, Context matters: Volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses, *J. Microbiol. Biol. Educ.* **14**, 176 (2013).
- [41] D. Conger and M. C. Long, Why are men falling behind? Gender gaps in college performance and persistence, *Ann. Am. Acad. Pol. Soc. Sci.* **627**, 184 (2010).
- [42] L. E. Kost-Smith, S. J. Pollock, and N. D. Finkelstein, Gender disparities in second-semester college physics: The incremental effects of a “smog of bias”, *Phys. Rev. ST Phys. Educ. Res.* **6**, 020112 (2010).
- [43] C. D. Wright, S. L. Eddy, M. P. Wenderoth, E. Abshire, M. Blankenbiller, S. E. Brownell, Cognitive difficulty and format of exams predicts gender and socioeconomic gaps in exam performance of students in introductory biology courses, *CBE Life Sci. Educ.* **15**, 1 (2016).
- [44] S. D. Willoughby and A. Metz, Exploring gender differences with different gain calculations in astronomy and biology, *Am. J. Phys.* **77**, 651 (2009).
- [45] B. J. Guzzetti and W. O. Williams, Gender, text, and discussion: Examining intellectual safety in the science classroom, *J. Res. Sci. Teach.* **33**, 5 (1996).
- [46] J. Holmes, Women’s talk in public contexts, *Discourse Soc.* **3**, 131 (1992).
- [47] P. Fassinger, How classes influence students’ participation in college classrooms, *J. Classroom Interact.* **35**, 38 (2000).
- [48] H. E. Tatum, B. M. Schwartz, and P. A. Schimmoeller, Classroom participation and student- faculty interactions: Does gender matter?, *J. Higher Educ.* **84**, 745 (2013).
- [49] J. Howard, A. Zoeller, and Y. Pratt, Students ‘ race and participation in sociology classroom discussion: A preliminary investigation, *J. Scholar. Teach. Learn.* **5**, 14 (2006).
- [50] K. L. Brady and R. M. Eisler, Sex and gender in the college classroom: A quantitative analysis of faculty-student interactions and perceptions, *J. Educ. Psychol.* **91**, 127 (1999).
- [51] W. J. Hughes, Perceived gender interaction and course confidence among undergraduate science, mathematics and technology majors, *J. Women Minorities Sci. Eng.* **6**, 155 (2000).
- [52] D. MacPhee, S. Farro, and S. S. Canetto, Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns, *Anal. Soc. Issues Public Policy* **13**, 347 (2013).
- [53] S. Beyer, K. Rynes, J. Perrault, K. Hay, and S. Haller, Gender differences in computer science students, *SIGCSE Bull.* **35**, 49 (2003).
- [54] H. Hartman and M. Hartman, Do gender differences in undergraduate engineering orientations persist when major is controlled?, *Int. J. Gender, Sci. Technol.* **1**, 61 (2009).
- [55] M. Micari, P. Pazos, and M. J. Z. Hartmann, A matter of confidence: Gender differences in attitudes towards engaging in lab and course work in undergraduate engineering, *J. Women Minorities Sci. Eng.* **13**, 279 (2007).
- [56] E. Cech, B. Rubineau, S. Silbey, and C. Seron, Professional role confidence and gendered persistence in engineering, *Am. Sociol. Rev.* **76**, 641 (2011).
- [57] M. Besterfield-Sacre, M. Moreno, L. Shuman, and C. Atman, Gender and ethnicity differences in freshmen engineering student attitudes: A cross-institutional study, *J. Eng. Educ.* **90**, 477 (2001).
- [58] C. M. Jagacinski, Women engineering students: Competence perceptions and achievement goals in the freshman engineering course, *Sex Roles* **69**, 644 (2013).
- [59] J. P. Concannon and L. H. Barrow, A reanalysis of engineering majors’ self-efficacy beliefs, *J. Sci. Educ. Technol.* **21**, 742 (2012).
- [60] J. Dalgety and R. K. Coll, Exploring first-year science students’ chemistry self-efficacy, *Int. J. Sci. Math. Educ.* **4**, 97 (2006).
- [61] S. M. Villafañe, C. A. Garcia, and J. E. Lewis, Exploring diverse students’ trends in chemistry self-efficacy throughout a semester of college-level preparatory chemistry, *Chem. Educ. Res. Pract.* **15**, 114 (2014).
- [62] M. C. Murphy, C. M. Steele, and J. J. Gross, Signaling threat: How situational cues affect women in math,

- science, and engineering settings, *Psychol. Sci.* **18**, 879 (2007).
- [63] J. L. Smith, K. L. Lewis, L. Hawthorne, and S. D. Hodges, When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns, *Pers. Soc. Psychol. Bull.* **39**, 131 (2013).
- [64] G. Townley, J. Katz, A. Wandersman, B. Skiles, M. J. Schillaci, E. Briana, and T. A. Mousseau, Exploring the role of sense of community in the undergraduate transfer student experience, *J. Community Psychol.* **41**, 277 (2013).
- [65] J. G. Stout, T. A. Ito, N. D. Finkelstein, and S. J. Pollock, How a gender gap in belonging contributes to the gender gap in physics participation, *AIP Conf. Proc.* **1513**, 402 (2013).
- [66] B. Z. Hazari, P. M. Sadler, and G. Sonnert, The science identity of college students: Exploring the intersection of gender, race, and ethnicity, *J. Coll. Sci. Teach.* **42**, 82 (2013).
- [67] J. L. Cundiff, T. K. Vescio, E. Loken, and L. Lo, Do gender-science stereotypes predict science identification and science career aspirations among undergraduate science majors?, *Soc. Psychol. Educ.* **16**, 541 (2013).
- [68] D. M. Young, L. A. Rudman, H. M. Buettner, and M. C. McLean, The influence of female role models on women's implicit science cognitions, *Psychol. Women Q.* **37**, 283 (2013).
- [69] R. D. Robnett, M. M. Chemers, and E. L. Zurbriggen, Longitudinal associations among undergraduates' research experience, self-efficacy, and identity, *J. Res. Sci. Teach.* **52**, 847 (2015).
- [70] E. R. Altermatt, J. Jovanovic, and M. Perry, Bias or responsiveness? Sex and achievement-level effects on teachers' classroom questioning practices, *J. Educ. Psychol.* **90**, 516 (1998).
- [71] C. Burns and D. Myhill, Interactive or inactive? A consideration of the nature of interaction in whole class teaching, *Cambridge J. Educ.* **34**, 35 (2004).
- [72] K. A. Rocca, Student participation in the college classroom: An extended multidisciplinary literature review, *Commun. Educ.* **59**, 185 (2010).
- [73] L. van den Bergh, E. Denessen, L. Hornstra, M. Voeten, and R. W. Holland, The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap, *Am. Educ. Res. J.* **47**, 497 (2010).
- [74] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8410 (2014).
- [75] M. T. H. Chi and R. Wylie, The icap framework: Linking cognitive engagement to active learning outcomes, *Educ. Psychol.* **49**, 219 (2014).
- [76] B. London, L. Rosenthal, S. R. Levy, and M. Lobel, The influences of perceived identity compatibility and social support on women in nontraditional fields during the college transition, *Basic Appl. Soc. Psych.* **33**, 304 (2011).
- [77] E. L. Usher and F. Pajares, Sources of self-efficacy in school: Critical review of the literature and future directions, *Rev. Educ. Res.* **78**, 751 (2008).
- [78] K. D. Multon, S. D. Brown, and R. W. Lent, Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation, *J. Counsel. Psychol.* **38**, 30 (1991).
- [79] C. Good, A. Rattan, and C. S. Dweck, Why do women opt out? Sense of belonging and women's representation in mathematics, *J. Pers. Soc. Psychol.* **102**, 700 (2012).
- [80] S. E. Holleran, J. Whitehead, T. Schmader, and M. R. Mehl, Talking shop and shooting the breeze: A study of workplace conversation and job disengagement among stem faculty, *Soc. Psychol. Personal. Sci.* **2**, 65 (2011).
- [81] M. M. Chemers, E. L. Zurbriggen, M. Syed, B. K. Goza, and S. Bearman, The role of efficacy and identity in science career commitment among underrepresented minority students, *J. Soc. Issues* **67**, 469 (2011).
- [82] J. Benson, Developing a strong program of construct validation: A test anxiety example, *Educ. Meas.* **17**, 10 (1998).
- [83] R. G. Netemeyer, W. O. Bearden, and S. Sharma, Scaling procedures: issues and applications, *Scaling Procedures: Issues and Applications* (Sage Publications, Thousand Oaks, CA, 2003).
- [84] M. J. Cullen, S. D. Waters, and P. R. Sackett, Testing stereotype threat theory predictions for math-identified and non-math-identified students by gender, *Hum. Perform.* **19**, 421 (2006).
- [85] P. R. Hernandez, P. W. Schultz, M. Estrada, A. Woodcock, and R. C. Chance, Sustaining Optimal Motivation: A Longitudinal Analysis of Interventions to Broaden Participation of Underrepresented Students in STEM, *J. Educ. Psychol.* **105**, 89 (2013).
- [86] T. Perez, J. G. Cromley, and A. Kaplan, The role of identity development, values, and costs in college STEM retention, *J. Educ. Psychol.* **106**, 315 (2014).
- [87] M.-T. Wang and J. Degol, Motivational pathways to stem career choices: Using expectancy-value perspective to understand individual and gender differences in stem fields, *Dev. Pol. Rev.* **33**, 304 (2013).
- [88] M. Long, J. Steinke, B. Applegate, M. Knight Lapinski, M. J. Johnson, and S. Ghosh, Portrayals of male and female scientists in television programs popular among middle school-age children, *Sci. Comm.* **32**, 356 (2010).
- [89] B. A. Nosek, F. L. Smyth, N. Sriram, N. M. Lindner, T. Devos, A. Ayala, Y. Bar-anan, E. Olli, J. Park, K. Schnabel, K. Shiomura, B. T. Tulbure, R. W. Wiers, N. Akrami, B. Ekehammar, M. Vianello, M. R. Banaji, and A. G. Greenwald, National differences in gender—science stereotypes predict national sex differences in science and math achievement, *Proc. Natl. Acad. Sci. U.S.A.* **106** (2009).
- [90] C. M. Steele, A threat in the air: How stereotypes shape intellectual identity and performance, *Am. Psychol.* **52**, 613 (1997).
- [91] S. J. Spencer, C. M. Steele, and D. M. Quinn, Stereotype threat and women's math performance, *J. Exp. Soc. Psychol.* **35**, 4 (1999).
- [92] A. K. Kiefer and D. Sekaquaptewa, Implicit stereotypes, gender identification and math-related outcomes: A prospective study of female college students, *Psychol. Sci.* **18**, 13 (2007).

- [93] J. L. Smith, E. R. Brown, D. B. Thoman, and E. D. Deemer, Losing its expected communal value: How stereotype threat undermines women's identity as research scientists, *Soc. Psychol. Educ.* **18**, 443 (2015).
- [94] J. Steele, J. B. James, and R. C. Barnett, Learning in a man's world: examining the perceptions of undergraduate women in male-dominated academic areas, *Psychol. Women Q.* **26**, 46 (2002).
- [95] J. G. Stout, N. Dasgupta, M. Hunsinger, and M. A. McManus, STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM), *J. Pers. Soc. Psychol.* **100**, 255 (2011).
- [96] J. R. Shapiro, A. M. Williams, and M. Hambarchyan, Are all interventions created equal? A multi-threat approach to tailoring stereotype threat interventions, *J. Pers. Soc. Psychol.* **104**, 277 (2013).
- [97] R. J. Crisp, L. M. Bache, and A. T. Maitner, Dynamics of social comparison in counter-stereotypic domains: Stereotype boost, not stereotype threat, for women in engineering majors, *Soc. Influ.* **4**, 171 (2009).
- [98] M. Johns, M. Inzlicht, and T. Schmader, Stereotype threat and executive resource depletion: Examining the influence of emotional regulation, *J. Exp. Psychol. Gen.* **137**, 691 (2008).
- [99] S. L. Beilock, R. J. Rydell, and A. R. McConnell, Stereotype threat and working memory: Mechanisms, alleviation, and spillover, *J. Exp. Psychol. Gen.* **136**, 256 (2007).
- [100] A. B. Brodish and P. G. Devine, The role of performance-avoidance goals and worry in mediating the relationship between stereotype threat and performance, *J. Exp. Soc. Psychol.* **45**, 180 (2009).
- [101] D. M. Marx and J. S. Roman, Female role models: Protecting women's math test performance, *Pers. Soc. Psychol. Bull.* **28**, 1183 (2002).
- [102] P. G. Davies, S. J. Spencer, D. M. Quinn, and R. Gerhardstein, Consuming images: how television commercials that elicit stereotype threat can restrain women academically and professionally, *Pers. Soc. Psychol. Bull.* **28**, 1615 (2002).
- [103] A. Miyake, L. E. Kost-Smith, N. D. Finkelstein, S. J. Pollock, G. L. Cohen, and T. a Ito, Reducing the gender achievement gap in college science: a classroom study of values affirmation, *Science* **330**, 1234 (2010).
- [104] J. G. Cromley, T. Perez, T. W. Wills, J. C. Tanaka, E. M. Horvat, and E. T.-B. Agbenyega, Changes in race and sex stereotype threat among diverse STEM students: Relation to grades and retention in the majors, *Contemp. Educ. Psychol.* **38**, 247 (2013).
- [105] L. R. Ramsey, D. E. Betz, and D. Sekaquaptewa, The effects of an academic environment intervention on science identification among women in STEM, *Soc. Psychol. Educ.* **16**, 377 (2013).
- [106] E. Reuben, P. Sapienza, and L. Zingales, How stereotypes impair women's careers in science, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4403 (2014).
- [107] J. M. Sheltzer and J. C. Smith, Elite male faculty in the life sciences employ fewer women, *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10107 (2014).
- [108] C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, and J. Handelsman, Science faculty's subtle gender biases favor male students, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16474 (2012).
- [109] A. L. Zeldin and F. Pajares, Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers, *Am. Educ. Res. J.* **37**, 215 (2000).
- [110] D. Z. Grunspan, S. L. Eddy, S. E. Brownell, B. L. Wiggins, A. J. Crowe, and S. M. Goodreau, Males under-estimate academic performance of their female peers in undergraduate biology classrooms, *PLoS One* **11**, e0148405 (2016).
- [111] J. Wolfe and E. Powell, Biases in interpersonal communication: How engineering students perceive gender typical speech acts in teamwork, *J. Eng. Educ.* **98**, 5 (2009).
- [112] D. W. Sue, *Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation* (Wiley, New York, 2010).
- [113] L. J. Barker, C. McDowell, and K. Kalahar, Exploring factors that influence computer science introductory course students to persist in the major, *ACM SIGCSE Bull.* **41**, 153 (2009).
- [114] A. B. Diekman, E. R. Brown, A. M. Johnston, and E. K. Clark, Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers, *Psychol. Sci.* **21**, 1051 (2010).
- [115] A. B. Diekman, E. K. Clark, A. M. Johnston, E. R. Brown, and M. Steinberg, Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective, *J. Pers. Soc. Psychol.* **101**, 902 (2011).
- [116] P. T. Costa, A. Terracciano, and R. R. McCrae, Gender differences in personality traits across cultures: robust and surprising findings, *J. Pers. Soc. Psychol.* **81**, 322 (2001).
- [117] A. M. Konrad, J. E. J. Ritchie, P. Lieb, and E. Corrigan, Sex differences and similarities in job attribute preferences: a meta-analysis, *Psychol. Bull.* **126**, 593 (2000).
- [118] A. B. Diekman, E. S. Weisgram, and A. L. Belanger, New routes to recruiting and retaining women in stem: Policy implications of a communal goal congruity perspective, *Social Issues and Policy Review* **9**, 52 (2015).
- [119] S. S. Canetto, C. D. Trott, J. J. Thomas, and C. A. Wynstra, Making sense of the atmospheric science gender gap: Do female and male graduate students have different career motives, goals, and challenges?, *J. Geosci. Educ.* **60**, 408 (2012).
- [120] L. J. Sax, Retaining tomorrow's scientists: exploring the factors that keep male and female college students interested in science careers, *J. Women Minorities Sci. Eng.* **1**, 45 (1994).
- [121] E. R. Brown, D. B. Thoman, J. L. Smith, and A. B. Diekman, Closing the communal gap: The importance of communal affordances in science career motivation, *J. Appl. Soc. Psychol.* **45**, 662 (2015).
- [122] C. S. Dweck, Is math a gift? Beliefs that put females at risk, in *Why aren't More Women in Science? Top*

- Researchers Debate the Evidence*, edited by S. J. Ceci and W. M. Williams (American Psychological Association, Washington, DC, 2007).
- [123] C. S. Dweck, Motivational processes affecting learning, *Am. Psychol.* **41**, 1040 (1986).
- [124] J. S. Eccles, T. F. Adler, R. Futterman, S. B. Goff, C. M. Kaczala, and J. L. Meece, Expectancy values and academic behaviors, in *Achievement and Achievement Motives: Psychological and Sociological Approaches* (Freeman, San Francisco, CA, 1983), pp. 75–146.
- [125] S.-J. Leslie, A. Cimpian, M. Meyer, and E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines, *Science* **347**, 262 (2015).
- [126] R. L. Shively and C. S. Ryan, Longitudinal changes in college math students' implicit theories of intelligence, *Soc. Psychol. Educ.* **16**, 241 (2013).
- [127] L. T. O'Brien, A. Blodorn, G. Adams, D. M. Garcia, and E. Hammer, Ethnic variation in gender-STEM stereotypes and STEM participation: An intersectional approach, *Cult. Divers. Ethn. Minor. Psychol.* (to be published).
- [128] J. C. Williams, K. W. Phillips, and E. V. Hall, Double jeopardy? Gender bias against women of color in science, <http://www.uchastings.edu/news/articles/2015/01/double-jeopardy-report.pdf>.
- [129] C. Jones and K. Shorter-Gooden, *Shifting: The double lives of black women in America*, *Shifting: The Double Lives of Black Women in America* (Harper, New York, 2003).