Received: 18 February 2019 | Revised: 2 October 2019 | Accepted: 3 November 2019

DOI: 10.1002/sce.21557

GENERAL SECTION

Science Education WILEY

Locating and understanding the largest gender differences in pathways to science degrees

Eben B. Witherspoon | Christian D. Schunn

Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania

Correspondence
Eben B. Witherspoon, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara St., Pittsburgh, PA 15217.
Email: eben.witherspoon@pitt.edu

Abstract

While gender parity has been achieved in overall science degree earning, large gaps still exist within many science disciplines. Further, studies addressing gender inequity in science often ignore a large source of undergraduate science degree earners: those who enroll in science courses intending to pursue careers in health or medicine. This study examines pathways toward or away from science degrees in $N = 4{,}345$ men and women enrolled in early science courses at a large undergraduate research university. Importantly, to understand shifts in students' academic intentions and how pathways to science may be differential by gender, this study analyzed students' incoming major and career intentions, estimates of incoming academic abilities, and relative performance in science and various non-science courses. Results show that while men and women initially intending to pursue a science major graduate with science degrees in equal numbers, the plurality of science degree earners are students entering college intending health or medical careers. Further, from those subgroups, a significantly larger proportion of men end up in science, while a significantly larger proportion of women end up outside of STEM completely. Understanding disciplinary differences in gender barriers to science participation can help inform interventions that specifically target those populations.

KEYWORDS

equity, gender, medical education, retention, science education

1 | INTRODUCTION

It is well established in the literature that despite recent progress toward equity, women and other groups historically underrepresented in science are still less likely to persist in STEM (science, technology, engineering, and mathematics) careers. Many studies have used the metaphor of a “leaky STEM pipeline” to explain this phenomenon, suggesting that the key to reversing women’s attrition in STEM-degree earning lies with identifying and “plugging” key points along the pipeline that leak differentially by gender. However, recent critiques have suggested that the pipeline metaphor inadequately represents the many different pathways that lead to STEM careers (Blickenstaff, 2005; Cannady, Greenwald, & Harris, 2014). Particular early entry points identified under this paradigm have been shown to represent relatively few eventual STEM participants. For example, more than 80% of students who eventually earn a STEM degree only begin to focus on STEM after they enroll in high school (Maltese & Tai, 2011). Other evidence suggests that shifts in STEM plans may not occur until even later, through the first 2 years of college. Even common indicators of STEM participation at the end of high school, such as taking calculus, have been shown to produce relatively few actual STEM participants in college, with one study showing students *without* this indicator being 3.5 times *more* likely to participate in postsecondary STEM coursework (Cannady et al., 2014). Therefore, it is important to examine these alternative pathways to science to understand the wider variety of routes that men and women undergraduate students take in and out of STEM and how these may differ from their initial academic intentions. Relying on early indicators and ignoring gender differences that may appear later and at less established entry points to science and STEM fields can allow important barriers to true gender equity to persist undiscovered.

An additional limitation of the STEM pipeline paradigm is that it necessitates an aggregation of the many possible routes to STEM, as well as an aggregation of the differences among the disciplines that make up STEM. Most saliently for considerations of gender equity, disciplinary differences in gender stereotypes and the representation of women in a particular field can underlie the processes through which students navigate to and away from science disciplines, offering explanations for why certain sectors of the sciences remain segregated by gender (Koester, Grom, & McKay, 2016; Leslie, Cimpian, Meyer, & Freeland, 2015). Gender differences in attrition have been shown to be highly variable by discipline, both between STEM, as well as within the sciences (Cheryan, Ziegler, Montoya, & Jiang, 2017; D. I. Miller, 2018). Further, gender differences in attrition have been shown to be highly variable by discipline, both between STEM, as well as within the sciences (Cheryan et al., 2017; Koester et al., 2016). For example, while aggregate levels of STEM graduates suggest parity between women and men (e.g., 1:1 in STEM overall), much of this effect derives from high representation in some specific STEM disciplines (e.g., 2:1 in Life Sciences), which hides continued much lower participation for women in others (e.g., 1:5 in Engineering; National Science Foundation, 2015).

Further, even undergraduates taking science courses within a single discipline are also likely to vary greatly by their particular career intentions. Many studies ignore a large and growing population of those students who often also differ by gender: students who enroll in science courses intending medical and health professions (Aschbacher, Li, & Roth, 2010; J. D. Miller & Solberg, 2012; Morgan, Gelbgiser, & Weeden, 2013). Studies using nationally representative US data show that 13% of men and 34% of women leave high school with an interest in pursuing health or medicine (Sadler, Sonnert, Hazari, & Tai, 2012). Some reports suggest that students in the health sciences are more likely to switch majors (35%) than students in science (28%); however, less is known about how these various pathways may differ by gender (Chen, 2013). In sum, while the broad STEM definition has been useful in identifying a broad range of disciplines that show gender disparity at the aggregate level, attending more closely to where differences within alternative pathways like health and medicine may lie is important to lay the groundwork for potential future interventions.

2 | THEORETICAL FRAMEWORK

2.1 | Relative academic performance

Reports on the Scholastic Aptitude Test (SAT) have shown that performance differences across the test vary by gender, with a large score advantage to men in the math section, a large advantage to women in the writing section, and relatively small differences in the verbal section (Mattern, Camara, & Kobrin, 2007). Students may internalize the scores they receive on the SAT as fundamental statements about aptitude in (or out of) science (Vincent-Ruz, Binning, Schunn, & Grabowski, 2018). Alternatively, these aptitude estimates could be predictors of course performance, which then drives beliefs about disciplinary aptitude. Interestingly, the writing section has been shown to be more correlated with first-year college performance than the verbal and math sections combined (Kolbrin, Patterson, Shaw, Mattern, & Barbuti, 2008). In later more advanced coursework, writing competencies may play an even larger role (Brownell, Price, & Steinman, 2013; Yalvac, Smith, Troy, & Hirsch, 2007) and humanities and social sciences in particular can involve substantial amounts of writing.

In terms of college grades, science course performance over the first 2 years of undergraduate study has been shown to have a sustained impact on students' motivations and persistence in their STEM major and career intentions (Cromley, Perez, & Kaplan, 2016). For women with initial interest in Health and Medical careers, introductory chemistry and biology courses have been shown to be the primary drivers of changing interest in continuing on those career paths (Barr, Gonzalez, & Wanat, 2008; Barr, Matsui, Wanat, & Gonzalez, 2010). However, prior studies of those effects do not provide information on where students go if they choose to leave pre-medical study; critically, it is unknown whether there are gender differences in this "production function" toward science from the pre-medical track. For example, students may give up on medical school goals if they are struggling in Organic Chemistry or Introductory Physics (because they believe very high grades are needed when applying to medical school) but still choose to complete a biology or neuroscience degree. Alternatively, they might switch majors to the humanities (like English or History) or the social sciences (like Anthropology). Therefore, understanding the impact of relative academic performance of men and women who indicate an early preference for science, medicine, or health careers could provide insight into when these shifts in major and career decisions may occur, and if certain academic considerations are more or less important for groups of students with different career intentions.

While some still argue that gender differences in science performance can be attributed to innate biological differences or cognitive ability in the sciences, substantial evidence from cognitive psychology and the learning sciences suggests that only a few such gender differences exist, they are small, and they exhibit relative strengths in both directions (Else-Quest, Hyde, & Linn, 2010; Hyde, 2005; Spelke, 2005). Instead, there is strong evidence supporting sociocultural explanations of historical differences in gendered science performance and participation.

Students' perceptions of relative academic performance provide one strong source of influence on attitudes and behaviors which can impact decisions about academic persistence. Perceptions of competence and expectations of success in a career can lead to the development of interest in, and eventually goals of pursuing, a particular career (Lent, Brown, & Hackett, 1994). Recent meta-analyses have shown that particularly regarding choice options in STEM fields as a whole, the association between outcome expectations and career goals may be higher for women than men (Lent et al., 2018). However, how this feedback is interpreted may be highly variable by discipline. For example, negative self-evaluations of these academic differences may be particularly salient for women in historically male-dominated science disciplines (Beyer & Bowden, 1997; Eccles, 1994; Kugler et al., 2017).

Further, given the many possible careers a student might pursue, *relative* strengths in performance (and corresponding *relative* expectancies for success) are as important to understanding academic choices. For example, if women who are high performing in science-relevant skills like math are also more likely than men to also have strengths in verbal ability, this may give them a wider range of viable and desirable alternatives to science careers (Wang, Eccles, & Kenny, 2013). In other words, this "relative strengths hypothesis" suggest that women may have more viable non-science options because of their verbal skills than their male peers, and therefore may be more likely to choose alternative career paths in less gender-stereotyped disciplines. However, less is known about differences in the strengths of these associations by gender within particular Science disciplines, and differential sources of outcome expectations.

Differences in disciplinary contexts and students' outcome expectations are likely to also influence the way these academic experiences are interpreted. In particular, it is unclear whether early indicators of prior performance like the SAT drive both attitudes about science, subsequent course performance and persistence, or whether course performance alone drives persistence. Longitudinal analyses testing this mechanism in engineering used a cumulative measure of college grade point average (GPA) to show that a more proximal indicator of performance is a stronger predictor of persistence than relatively distal standardized application test scores for both men and women (Lent et al., 2015, 2016). Few studies have examined how relative performance in specific disciplines may provide differential feedback to influence gendered persistence in science and health-related fields. An examination of the impact of discipline-specific academic performance on persistence could offer an important contribution to this literature.

2.2 | The current study

A number of studies have examined factors that predict students' college major selection, which in turn is an indicator for future careers. While many studies looking at gender differences conclude that disparities found are not a function of work-family goals or prior academic preparation, some suggest gender differences begin in high school because major intent upon leaving high school can be a strong predictor or initial college major selection (Morgan et al., 2013; Sadler et al., 2012). However, existing longitudinal studies have not been able to obtain both incoming major intent of college students with sufficient detail to precisely estimate students incoming academic plans/career goals, as well as rich data on students' course experiences. Our study uses fine-grained institutional data and surveys to identify students' major intentions upon entering college, indicators of prior academic preparation, average grade performance in courses for different discipline groupings (e.g., STEM and non-STEM), graduation degree data, and graduate school entrance exam-taking to assess how these initial intentions change throughout the course of the college experience.

In particular, we are interested in addressing the following three research questions. First, *what academic plans are the largest sources of science degrees*? While it is likely that a high proportion of students who graduate with science degrees enter college with the intent to pursue science, there are also a large number of students who begin on a path toward related careers (i.e., Health and Medical professions) who may then shift into the sciences. Second, *are there varying gender differences within the different pathways to science degrees*? That is, are men and women who begin with a plan to pursue Science, Health or Medicine more likely to shift into science, or into a particular non-science field. Third, *are gender differences in persistence to science degrees within different pathways mediated by relative strengths in academic performance*? For example, are distal indicators of academic preparation such as the SAT or more proximal predictors such as course GPA strong predictors of shifts in major, and are there disciplinary differences in the strength of these effect by students' initial career intentions?

3 | METHODS

3.1 | Sample

In the current study, we examined institutional data records from N = 4,345 undergraduate students in the College of Arts and Sciences and College of General Studies at a large undergraduate research university in the northeastern United States (henceforth, "the University"). We excluded students who matriculated into specific colleges (i.e., College of Engineering, College of Nursing) as these students show a high level of commitment to a particular pathway, and therefore are unlikely to exhibit similar behavior to the population of students pursing a Bachelor of Arts or Sciences degree. The University is broadly representative of similar institutions in the United States with a relatively selective admissions rate (approximately 60%): it offers over 100 undergraduate majors, the majority (60%) of students are from in-state, with a smaller number (5%) of international students, and while there is large variability in family income (*SD* = $122,000), students tend to come from higher income brackets (*Median* = $111,000).

Sampled students were those enrolled at any point in an Introductory General Chemistry course, and who had matriculated to the University between the first semester of 2009 and the final semester of 2012. This course was selected for the sample definition because it is a core, required introductory course for a variety of Science majors, and rarely taken by non-Science intending majors as a method of fulfilling their general education requirement in the natural sciences; therefore, it is likely to predominantly sample those students who enter college intending Science careers, our population of interest. Four cohorts of students were used to ensure that patterns found were not specific to particular instructors or random within-section student groupings, while allowing at least a 5-year graduation window for all students before our final data collection in the Fall of 2018; according to University reports of graduation rates, about 81% of students at the University finish their undergraduate degrees within this timeframe.

The racial and ethnic diversity of our sample roughly mirrored that of the University as a whole; students were predominantly White (74%), with Asian (12%), Black (7%), and Hispanic (2%) students making up the next largest ethnic groups (all other races and ethnicities each represented <1% of the data). The primary predictor variable, gender, was coded as 1 if the student self-identified as women (55%) and 0 if the student self-identified as men (45%). There were *N* = 17 students who had not indicated any gender on our survey; these observations were treated as missing and removed from analysis. All University data were provided for analysis with Institutional Review Board approval.

3.2 | Measures

3.2.1 | Intended major

The primary predictor of interest for this study was the incoming academic plan for students in the College of Arts and Sciences and the College of General Studies, which was information about the students' intended major or career collected prior the first year of classes, at the point of matriculation to the University. Categories were coded across 59 unique plans into five separate general categories (see Table 1): Science, Medicine, Health, Undecided, and Non-Science. Intent to pursue Science was defined as students who indicated an intent to major or pursue a career in any of the natural sciences, and was also coded separately to allow for separate follow-up analyses of degrees earned in the life sciences (i.e., Biology, Neuroscience), and physical sciences (i.e., Chemistry, Physics, Geology) because of the large gender differences across those areas. Medical intending were those students who specifically selected Pre-Medicine, while Health were all others intending health careers that did not require medical school. Finally, Undecided students were those who explicitly marked they were undecided in their major or career intent, and the Non-Science category consisted of a combination of majors, primarily in the social sciences and humanities (i.e., Anthropology, Psychology, Political Science, English, Music). Due to our specific focus on pathways toward and away from Science, a small remaining number of students (*N* = 40) indicating mathematics/technology-related disciplines (i.e., Mathematics, Statistics, Computer Science) were removed from analyses. There was also some missing information for students' intended major (*N* = 188); however, correlations show this missingness to be not systematic across gender or any other variables of interest (*r*s < .10), and so these cases were removed from analyses, leaving a final *N* = 4,140.

3.2.2 | Degrees earned

Bachelor's degrees earned by students in the sample, the primary outcome variable, were gathered from University historical data and coded across 498 unique degrees and degree combinations into seven general degree

TABLE 1 Examples of intended major and degree category codes

Category	Examples ordered from most to least
Intended major	
Science	Biological Sciences, Chemistry, Neuroscience, Physics, Environmental Studies, Geology, etc.
Medicine	Pre-Medicine
Health	Pharmacy, Pre-Physical Therapy, Pre-Dentistry, Clinical Dietetics/Nutrition, Pre-Rehabilitation Science, Pre-Athletic Training, etc.
Non-Science	Psychology, Pre-Education, Anthropology, Political Science, History, Pre-Law, English Literature, Philosophy, Spanish, Music, etc.
Undecided	Undecided
Degree	
Science	Biology, Microbiology, Chemistry, Neuroscience, Physics, Geology, etc.
Life Sciences	Biology, Microbiology, Neuroscience, etc.
Physical Sciences	Chemistry, Physics, Geology, etc.
Science + MCAT	Any of the degrees above, and also took the MCAT
Other + MCAT	Any of the degrees below, and also took the MCAT
Non-STEM	English, History, Philosophy, African Studies, Arts, Music Theater, Music, Languages, Religious Studies, Anthropology, Economics, Political Science, Psychology, Sociology, Business Accounting, Finance, etc.
Math/Engineering	Mathematics, Accounting, Applied Math, Computer Science, Statistics, Civil Engineering, Mechanical Engineering, Electrical Engineering, Chemical Engineering, etc.
Health	Nursing, Dental Hygiene, Pharmaceutical Sciences, Sports Medicine, Nutrition and Dietetics, Emergency Medicine, etc.

Abbreviations: MCAT, Medical College Admissions Test; STEM, science, technology, engineering, and mathematics.

categories: Health, Social Science, Arts & Humanities, Science, Math, Engineering, and Business (see Table 1 for a detailed coding scheme). It is important to note here that Science degree earners were separated into those students who took the Medical College Admissions Test (MCAT) and those who did not; there are a number of students who earned a Science degree and yet sat for the MCAT exam, which is a strong indication that their career intent is toward a medical career and not a career in the pure Sciences; therefore, we wanted to be able to distinguish between those two distinct groups of students. Medical school in the United States is a postbaccalaureate degree, and pre-medical students earn their undergraduate degrees in a variety of fields (e.g., in neuroscience, in chemistry, in biology, in psychology) with medicine-specific degree earned at the undergraduate level at most universities, including the one studied here. Therefore, we chose to use the entrance exam for medical school as a strong indicator for intent to pursue a medical degree and career. We also distinguished another, smaller group of students who sat for the MCAT but graduated with non-Science degrees. In our final sample, N = 3,409 (82%) of students had an earned degree recorded by the University, while N = 731 (18%) students did not, which matched reported general attrition based on University records for 5-year graduation rates.

3.2.3 | Academic performance

We used institutional warehouse data to obtain students SAT scores (Math, Verbal, and Writing) and GPAs as indicators of academic performance. The University GPA scale awards an A for 4.0, B for 3.0, C for 2.0, and D for 1.0, with a "plus" or "minus" grade awarded at 0.25 grade points above or below, respectively (e.g., C+ = 2.25, C − = 1.75); any grade below a 0.75 received an F. With these grades, we separately calculated GPAs in the first 2 years of Arts & Humanities classes, Social Sciences courses (defined using the same categorization system as for academic plan and degree as above), and Science & Math courses. Mathematics course grades were included with science grades as mathematics ability is closely related to performance in many introductory Science courses

(Sadler & Tai, 2007), and the plurality of students are required to take at least one Mathematics course as a general education requirement. See Section 3.3 for how these variables were transformed into relative performance ratios.

3.3 | Analyses

Data were analyzed using χ^2 tests to examine the level of significance for comparisons of proportions between systematically divided subgroups. We first counted the raw numbers of students in our sample who had indicated each category of intended major, by gender. Next, we determined the category of the degree earned by all students who graduated, also by gender (see Table 2 for full details). However, like most incoming undergraduate cohorts in the United States in recent decades (McFarland et al., 2019), this sample contained a higher number of women than men overall, which limits the interpretability of direct numeric comparisons. Therefore, we used these frequencies to calculate the *proportion* of men and women who graduated within each degree category, as well as the proportion who graduated with each degree who had entered from each intended academic major category.

Our initial analyses were focused on understanding the proportion of men and women who stayed with their initial intent to study Science, to identify any baseline gender differences in Science retention in our sample; that is, we examined Science degrees earned by only those students intending to major in Science. To assess this, we compared the number of men and women who had initially intended to major in Science with no indication of an intent to continue to medical school, and then had successfully earned a Science degree and not gone on to take the MCAT (i.e., "Science [No MCAT]").

The second set of analyses examined the proportion of men and women who had indicated an intent to pursue any non-Science major, but had then gone on to earn a Science degree, to identify which non-Science intended major categories were the largest producers of Science degrees, as well as determine whether there were gender differences in students who switch into Science. To do this, we again calculated the number of students who earned a "Science (No MCAT)" degree by gender, but this time as a proportion out of the total number of students who had started in each category of intended major.

The third set of analyses addressed attrition from Science and STEM, by looking at different categories of non-Science degrees that were earned by men and women, both overall and from each of the intended major categories. In other words, for each intended major, if students were *not* graduating in the Sciences, what degree had they ended up with, and are any of those attrition patterns differential by gender? To examine this, we disaggregated non-Science degree earners (i.e., "All Non-Science") into various subcategories (including separately those students who went on to take the MCAT after earning a Science or non-Science degree, as described in Section 3 above). We then again calculated the proportion of students in each intended major who earned each of those types of degrees, by gender. For all analyses, χ^2 tests were used to determine if there were significant differences ($\alpha = .01$) in the proportion of men and women who continued on to earn each degree from each intended major, and raw proportions or odds ratios are provided as proxies of effect sizes. This relatively conservative α level was selected to reduce the possibility of achieving a significant effect by chance when conducting multiple χ^2 tests.

Finally, we performed a series of mediation analyses to begin to understand potential explanatory mechanisms for any gender differences found in academic plan and eventual degree earned. Mediation analyses provide an understanding of the hypothesized processes that underlie observed relationships between a predictor and outcome variable, by including a third mediating variable which is related to both the predictor and the outcome (Mackinnon, Fairchild, & Fritz, 2007). The direct relationship between the predictor and the outcome is explained (or "mediated") by this third variable, to the extent that the strength of the direct effect is reduced when mediation variables are included. In particular, as a test of the relative strengths hypothesis, we utilized mediation analyses to first analyze whether gender differences found in any of the observed academic plan-to-degree earning pathways were mediated by relative strengths in verbal and writing performance compared with math performance, as operationalized by relative SAT scores. However, since there were only differences in math and writing but not verbal SAT (see Table 3), we compare ratios of writing to math SAT scores since they could feasibly have an impact

TABLE 2 Frequency table of students intended undergraduate major categories, and degrees earned, by gender

	Science degree (and no MCAT) (N = 956)		Non-Science degree (or MCAT) (*N* = 2,066)										Total	
			Math/Engineering		Science + MCAT		Other + MCAT		Health		Non-STEM			
Intended major	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Science	163	227	47	26	57	39	6	8	16	51	86	144	375	495
Medicine	133	135	21	10	109	58	32	31	31	61	105	183	431	478
Health	80	84	19	26	8	8	4	4	112	296	86	183	309	601
Undecided	45	56	94	34	13	6	5	2	20	55	101	116	278	269
Non-Science	15	18	9	4	2	3	1	6	3	15	33	64	63	110
Total	436	520	190	100	189	114	48	51	182	478	411	690	1,456	1,953

Abbreviations: MCAT, Medical College Admissions Test; STEM, science, technology, engineering, and mathematics.

TABLE 3 Overall means by gender and effect size/statistical significance of genders differences in performance, by academic plan (Health, Science, and Medicine) for SATs and cumulative GPAs in Natural Science, Social Science, and Arts & Humanities courses

	Health (*N* = 1,109)			**Science (*N* = 1,047)**			**Medicine (*N* = 1,087)**		
	Men (*N* = 400)	**Women (*N* = 709)**		**Men (*N* = 467)**	**Women (*N* = 580)**		**Men (*N* = 523)**	**Women (*N* = 564)**	
	Mean (*SE*)	**Mean (*SE*)**	***d***	**Mean (*SE*)**	**Mean (*SE*)**	***d***	**Mean (*SE*)**	**Mean (*SE*)**	***d***
Relative strengths for women									
SAT Writing Score	**595** (3.6)	**609** (2.6)	**0.20****	**614** (3.3)	**627** (3.0)	**0.19****	**617** (3.5)	**634** (3.0)	**0.22******
Arts & Humanities GPA	**3.33** (0.03)	**3.49** (0.02)	**0.29******	3.41 (0.03)	3.48 (0.02)	0.13˜	3.37 (0.02)	3.49 (0.02)	0.19˜
Social Science GPA	3.18 (0.04)	3.30 (0.02)	0.16˜	3.25 (0.04)	3.31 (0.04)	0.08	**3.21** (0.04)	**3.33** (0.03)	**0.17****
Equivalent by gender									
SAT Verbal Score	612.4 (3.68)	613.0 (2.71)	0.01	636.5 (3.40)	638.7 (3.26)	0.03	636.1 (3.45)	638.4 (3.25)	0.03
Relative strengths for men									
SAT Math Score	**660** (3.1)	**633** (2.6)	**−0.41******	**665** (3.2)	**643** (2.9)	**−0.32******	**672** (3.1)	**647** (3.0)	**−0.35******
Science GPA	2.72 (0.05)	2.68 (0.03)	−0.04	**2.77** (0.04)	**2.66** (0.04)	**−0.12****	**3285** (0.04)	**2.60** (0.04)	**−0.29******

Note: Rows are organized starting with relative strengths for women and ending with relative strengths for men. Statistically significant differences are in bold.

Abbreviations: GPA, grade point average; SAT, Scholastic Aptitude Test.

˜p < .05.

*p < .01.

**p < .001.

on relative performance by gender. By a similar logic based on observed relative strengths and weaknesses, we also tested as mediators two ratios comparing students' GPA in their Science courses to their GPA in their Arts and Humanities courses, and to their GPA in their Social Sciences courses. For each mediation model, we report standardized regression coefficients as an effect size for individual paths and the proportion of the total effect mediated, as well as conduct follow-up Sobel–Goodman tests (Mackinnon, Lockwood, Hoffman, & West, 2002; Sobel, 1982) to determine the strength and significance of any potential mediation effects, and the significance of each direct and indirect pathway.

These analyses were performed in two ways, to provide a test for robustness of our findings and of our assumptions about different timepoints where students' major and career decisions may shift. First, we modeled an early attrition function that included performance in introductory science courses; that is, only grades through the first three semesters were used to calculate these ratios. Second, we modeled GPA ratios up to and including a later and well-known gatekeeping course for many science majors: Introductory Organic Chemistry. This latter function showed similar patterns, suggesting the findings to be robust; however, because of the later attrition point, it was less representative of our original sample. Therefore, we present the early attrition results in the main text, and have included the later attrition function as Appendix A.

4 | RESULTS

4.1 | Gendered patterns toward and away from science

As would be expected, students who entered college with an intent to pursue Science majors and careers were most likely to graduate with a Science degree relative to students entering with other academic plans, producing a larger number of science degree graduates ($N = 390$). However, it is important to note that the Medicine ($N = 268$) and Health ($N = 164$) academic plans combined to contribute more Science degree earners than those initially intending Science (45% from Medicine and Health vs. 41% from Science, of all non-medical Science degree earners). By contrast, relatively few students graduated with Science degrees who entered as Undecided ($N = 101$, 11%) or from entirely non-Science plans ($N = 33$, 3%). Thus, an overall understanding of science degree production must attend to the large contribution of students (at this university and across the United States) initially intending Medicine or Health (see Figure 2, top).

When looking at gendered productivity to Science degrees in absolute numbers across all academic plans, more women graduated with a Science degree ($N = 520$) than did men ($N = 436$), further supporting recent reports of growing gender parity in science in terms of raw numbers of degrees conferred. The balance in gendered yields held across all five of the examined academic plans. However, this pattern reflected (sometimes large) differences in initial starting numbers: there were more women than men beginning the path toward science (1.2:1), toward medicine (1.1:1), and especially toward health (1.6:1) and non-Science (1.8:1, see arrow widths in Figure 1).

When focused on yield relative percentages by gender, a less rosy picture emerges: there was a marginal gender difference in the percentages of women and men entering who earn a Science degree at this general level (27% women vs. 30% men, $\chi^2(1) = 4.55$, $p = .033$). Importantly, the differential rate of producing science degrees predominantly came from two non-Science plans: Health and non-Science (see top of Figure 2). The percentage earning Science degrees who entered with an intent to major in Science was not significantly different by gender (46% women vs. 43% men, $\chi^2(1) = 0.49$, $p = .48$). By contrast, a large portion of the overall gender disparity found in earned Science degrees came from the significantly lower percentages of women compared with men who entered with a Health plan and ended up with a Science degree (14% women vs. 26% men, $\chi^2(1) = 19.60$, $p < .001$). This effect was further compounded by the fact that health was the largest academic plan for women ($N = 709$, 31% of all women, see Figure 2).

Thus, contrary to the common "leaky pipeline" metaphor, the University appears to be equally successful at retaining both women and men who initially intend to study Sciences, similar to what others have reported

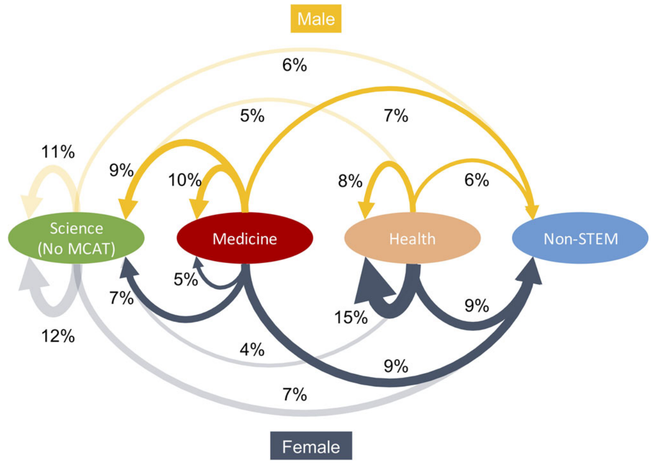

FIGURE 1 Diagram of the largest pathways into and out of various degree categories. Percentages represent how many students followed each pathway for each gender's total sample, with faded arrows representing no significant gender differences in percentages. Arrow widths are proportional to the raw number of students within each pathway. For clarity, only pathways that represent ≥5% of the data are shown [Color figure can be viewed at wileyonlinelibrary.com]

(Cheryan et al., 2017). However, this finding is qualified somewhat when categorizing the natural sciences into Physical Sciences (i.e., Chemistry, Physics, Geology) and Life Sciences (i.e., Biology, Microbiology, Neuroscience). Overall, this sample was dominated by Life Sciences degrees ($N = 777$, 82%) following national trends (National Science Foundation, 2015), and students intending Medicine or Health were even more likely to earn degrees in Life Sciences compared with students intending Science (86% from Medicine and Health vs. 79% from Science,

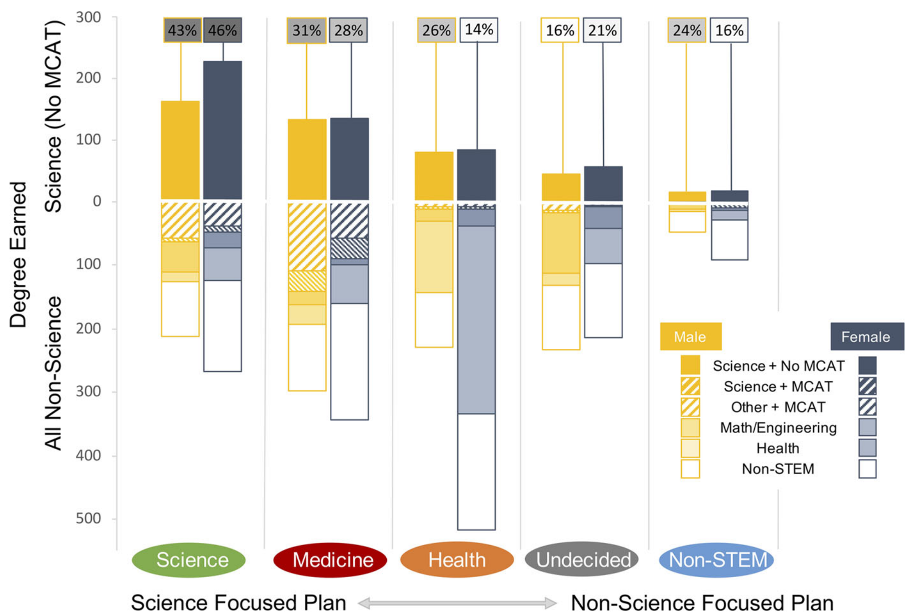

FIGURE 2 Descriptive statistics of the number of Science degrees earned from each academic plan category by gender (top), and the number of disaggregated non-Science degrees earned from each academic plan category by gender (bottom). Relative percentage yield to science for each academic plan category by gender are indicated inside the boxes above the top bars [Color figure can be viewed at wileyonlinelibrary.com]

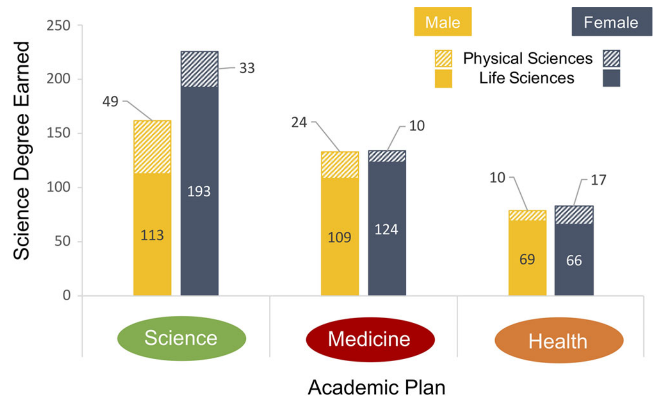

FIGURE 3 Number of students earning degrees in either the Life Sciences (solid) or Physical Sciences (lined), who initially declared either Science, Medicine, or Health academic plans, by gender [Color figure can be viewed at wileyonlinelibrary.com]

$\chi^2(1) = 7.21$, $p < .01$; see Figure 3). By gender, more women than men earned a degree in the Life Sciences compared with the Physical Sciences (86% women vs. 77% men, $\chi^2(1) = 13.94$, $p < .001$). However, from students intending Medicine and Health, there were no significant differences in the yield to Life Sciences between women and men (88% women vs. 84% men, $\chi^2(1) = 1.18$, $p = .278$). Instead, significant differences in the yield to the Life Sciences between women and men were located within the group of students initially intending Science (85% women vs. 70% men, $\chi^2(1) = 13.13$, $p < .001$; see Figure 3), likely reflecting initial differences in discipline interests. Most importantly, students intending Medicine and Health are an important source of both Physical and Life science degrees for both men (41% of Physical Sciences/61% of Life Sciences) and women (45% of Physical Sciences/50% of Life Sciences; Figure 3).

Another worrisome pattern from a science policy perspective emerges when examining gendered differences in what degrees were obtained when not earning non-medical Science degrees (see lower half of Figure 2). These other outcomes varied from being closely connected to science (e.g., pursuing medicine, math, or engineering) to those more distant from science, especially non-STEM degrees. Most concerning for the goal of gender equity in STEM, women from all academic plan categories were more likely than men to move away from STEM entirely, graduating without any type of STEM degree (overall 35% women vs. 28% men, $\chi^2(1) = 19.24$, $p < .001$; see Figure 2, bottom). It is important to note that even with our sample of students with no initial intent to pursue Mathematics and Engineering, men who initially intending Science were more likely to end up graduating with these degrees than women. This suggest that even men who move away from Science are more likely to stay within STEM broadly, providing another example of how the aggregation of STEM can lead to a misrepresentation of the character of continued gender gaps. The largest gendered rate differences for earning non-STEM degrees were those students who entered the University intending to pursue Medicine (38% women vs. 24% men, $\chi^2(1) = 20.30$, $p < .001$), with a difference also observed in those intending to pursue Science (29% women vs. 23% men, $\chi^2(1) = 4.16$, $p = .041$; see Figure 2, bottom). That the effect occurred in those two academic plans is somewhat surprising since these plans were most closely aligned with an intent to enter a science field.

4.2 | Mediators of pathway gender differences

The next set of analyses used mediation to test the relative strengths hypothesis in the two pathways that produced the plurality of science degree earners and that also showed large gender differences: (a) Medical academic plans leading to more women than men graduating with entirely non-STEM degrees and (b) Health academic plans leading to fewer women than men graduating with Science degrees (without taking the MCAT).

Specifically, mediation models for these two groups included ratios of relative STEM and non-STEM academic performance in SAT scores, and ratios of GPA in Science courses compared to either GPA in Arts and Humanities or GPA in Social Science courses, as possible mediators between entering with a Health academic plan and leaving with a non-medical Science degree, and between entering with a Medical academic plan and leaving with a non-STEM degree.

Results of the mediation analyses for the Medical academic plan to non-STEM pathway showed that this differential gender effect was mediated by relative STEM and non-STEM academic performance for students (56% of total effect for Medicine plans). Focusing on the specific mediating variables, women performed significantly higher on Writing than Math SAT scores, and had higher GPAs in both their Arts and Humanities and Social Sciences courses than in their Science courses. However, only those relative performances in GPA were shown to be significantly associated with graduating with a non-STEM degree (see Figure 4a). That is, the indirect effects of having a higher Arts and Humanities GPA and having a higher Social Science GPA relative to their Science courses, were significant mediators ($\beta = .07$, $p < .001$) of women's higher likelihood of graduating with a non-STEM degree. It is important to note that in a model including the SAT ratio, these indirect GPA effects are significant whereas the SAT ratio is not a significant mediator in that model. This suggests that more proximal academic experiences of relative performance in Arts and Humanities and Social Sciences classes are larger drivers of these students' decisions to shift to a non-STEM major than fixed ability differences or long-standing attitudes shaped by prior academic performance on the SAT.

Mediation analysis of the gendered Health to non-medical Science pathway showed that women with this initial academic plan were also slightly more likely to have a higher Writing to Math SAT ratio, and that this variable was also (not significantly) associated with pursuing a non-medical Science degree. However, different from the Medicine analysis, women with a Health subplan did not show any significant differences in either their Arts and Humanities or Social Sciences grades relative to their Science grades, and further none of these factors were associated with earning a non-medical Science degree (see Figure 4b). In the Health subplan, gender also contributed to variation in Science degrees indirectly through relative GPA and SAT, but these indirect effects were marginally significant ($\beta = -.02$, $p = .04$), only explained a small proportion (20%) of the total effect, and the direct path between women and Science degrees remained significant even after including these mediating variables.

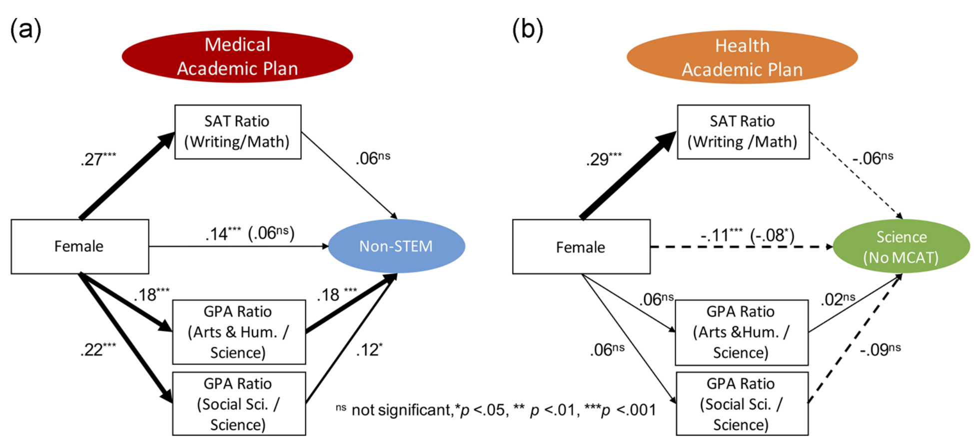

FIGURE 4 Mediation analyses of the effect of relative Scholastic Aptitude Test (SAT) scores and relative grade point average (GPA) on gender differences in students with (a) initial academic plans in Medicine, who graduate with a non-STEM degree and (b) initial academic plans in Health, who graduate with a Science degree and do not take the Medical College Admissions Test (MCAT). (Thickness of lines indicates relative strength of associations; dashed lines indicate a negative association. Covariation between all academic variables included in model but not shown for clarity.) [Color figure can be viewed at wileyonlinelibrary.com]

Therefore, overall the lower proportion of women moving from Health to Science was not well-explained by these academic performance variables, suggesting a different mechanism is likely responsible for this particular trajectory; we discuss potential alternative explanations for this finding in Section 5 below.

5 | DISCUSSION

In this study, we examined the incoming academic intentions and degree outcomes for a large sample of undergraduate students to determine if there were common alternative pathways to reaching a science degree, and whether or not these pathways provided differential access to science degrees by gender. Our analyses identify two major incoming academic plans that become large contributors of students who eventually earn science degrees: students who enter with the intention to study in the Health professions or pursue a Medical career. This finding also provides support for the growing body of work suggesting that instead of a STEM pipeline, a multipathways model for STEM retention will be needed to provide a more accurate depiction of the various ways that students actually enter and leave science as they navigate their college courses, allowing for the consideration of large numbers of students who may switch into both Life and Physical science fields at later points (Cannady et al., 2014; Mervis, 2012). In fact, our analyses show that these two pathways produce *more* science degree earners than does the group of students who initially intend to study science. As others have suggested (Kimmel, Miller, & Eccles, 2012; J. D. Miller & Solberg, 2012), future studies of STEM retention that do not acknowledge these large groups of medical and health intending students are at risk of ignoring a significant and often gendered source of entrance into degrees and careers in science.

Further, and most importantly for the goals of increasing women's participation in science and STEM, these large health and medical entry points also show sizable gender differences in their propensity to be on-ramps to science or to be off-ramps from STEM entirely. Specifically, our result show that women who enter with an intention to pursue Health are *less* likely than men of that group to pursue Science, while women with an initial intention to pursue Medicine are *more* likely than men in that group to leave to a field completely outside of STEM. This finding suggest that not only should health and medical fields be considered in analyses of gender differences in STEM retention, but that within these two fields there may be specific differences in the educational experiences for men and women that impact their decisions to persist in science, or leave STEM completely, and at higher rates relative to men than from other pathways to science.

Mediation analyses show that academic factors such as SAT and GPA, which can provide a strong signal for students regarding their expectations of success and beliefs about their abilities in science, explain much of the gendered differences for the group intending Medicine. Specifically, women's relative strength in Social Science courses and especially Arts and Humanities courses compared to their Science courses may offer them more non-STEM academic options relative to their male peers. This supports the hypothesis that relative academic strengths in non-STEM fields may pull some high-performing women who initially enter intending medicine away from science, and contributes to this body of work by showing that more proximal measures of performance may be better indicators of this effect than early indicators of prior achievement like the SAT (see Thoman, Arizaga, Smith, Story, & Soncuya, 2014; Wang et al., 2013).

However, relative academic performance was not shown to be a strong mediator of the lower proportion of women who initially intend to pursue Health and end up earning non-STEM degrees. Understanding how the educational contexts and early motivations for Health intending and Medicine intending students differ could offer a number of alternative explanations for disciplinary differences found in these phenomena. Pre-medical study in the United States typically requires a long sequence of rigorous science courses such as Organic Chemistry and Physics that are often majority male, and are well-established "gatekeeper" courses which often attrit a disproportionate number of students from groups historically underrepresented in science (Barr, 2010; Barr et al., 2010). Women in majority-male disciplines are more likely to experience a "chilly climate" of social marginalization, sexism, and stereotype threats, which can reduce their performance and influence their choice to leave those fields (Beasley & Fischer, 2012; Logel et al., 2009; Shapiro & Williams, 2012). These stereotype threats may have a particularly strong negative impact on grades for women entering STEM courses from other disciplines, leading to even larger barriers for women who do not begin college with an initial intent to pursue science (Smeding, 2012; Thoman et al., 2014). While small performance differences in both directions existed for women and men in the Science intending group, our findings show that women with Health and Medicine academic plans had significantly higher grades in Arts and Humanities, Social Sciences, and SAT Writing than men in those groups. Further, women intending Health and especially Medicine had higher GPAs in non-STEM areas than women who entered college intending science. This may suggest that the explanation of greater alternative options as a source of gendered attrition may be particularly relevant to the group of women entering with Medical academic plans, as their overall academic strengths in both STEM and non-STEM areas provide them with a range of other fields to leave to which may offer a less “chilly climate.”

These findings suggest that in addition to disaggregating students’ initial academic plans, using measures of relative academic strengths can provide insight into the role relative performance plays in men and women’s decisions to switch majors or careers within pathways like Health and Medicine. Meta-analyses addressing social cognitive career theory (SCCT) have shown that prior GPA primarily influences persistence through motivational factors like self-efficacy, suggesting it is perception of ability rather than preparation that students weigh in their decisions to persist (Brown et al., 2008). Our findings argue for updates to SCCT to focus on relative self-efficacy across domains rather than absolute self-efficacy within one domain; work by Marsh (Marsh, 1986; Marsh et al., 2018) and other more recent studies also suggest that *relative* perceptions of ability may be more strongly related to self-efficacy, and therefore academic persistence decisions (Wang et al., 2013).

The literature also suggests that pre-medical academic environments are likely to be much more competitive about grades, while coursework on the pathway to Health careers may be less focused on academic competition and place a larger emphasis on authentic experiences that help develop the required skillset for those professions (Horowitz, 2009; Lempp & Seale, 2004). Particularly in introductory science courses, these different instructional environments may elicit different perceptions of the size and importance of relative academic ability, even when these ability differences are in fact relatively small. Competition has been shown to have mixed effects on future performance through simultaneously orientating toward performance goals (i.e., a focus on demonstrating relative ability) and mastery goals (i.e., a focus on achieving conceptual understanding), a dichotomy that may be particularly salient for medical students (Horowitz, 2009; Murayama & Elliot, 2012). While both men and women have been shown to demonstrate performance benefits from single-gender competition, only women experience a negative performance effect from mixed-gender competition; in part, this differential gender effect is explained through men’s higher competency beliefs, even when there are no actual performance differences (Niederle & Vesterlund, 2010). Our sample showed that compared to the number of students intending Medicine, the number of students intending Health has a much higher proportion of women. It could be that in addition to a reduced emphasis on academic competition, as women progress through Health courses they are more likely to encounter majority female classes, rather than the majority male classes they encounter in many Science and Medicine introductory courses. This experience as a majority gender in the Health pathway may mitigate some of the negative effects of stereotype threat, and in fact have a positive effect on performance for women who are competing in more homogenous courses.

5.1 | Limitations and future directions

Some limitations to this study should be considered when interpreting the findings. First, this study took place at a single institution, and therefore the patterns found here may be unique to this population. In addition, it is important to limit the interpretation and generalizability of these data to Colleges of Arts and Sciences and Colleges of General Studies; findings regarding retention to intended major are likely to show less variation in more directed programs of study like Colleges of Nursing or Engineering. While the University courses do represent a common sequence and structure for their pre-medical and health courses that is relatively typical of similar large research institutions, various demographic and regional factors could influence the learning environments and the particular way in which students perceive and navigate different majors and career pathways. Future research should attempt to apply a similar approach in a multi-institutional sample, to better understand which factors are consistently important, and whether institutional culture and interventions might be moderators of these effects. However, where national data exist, there is a match to the data in the current study. For example, the differential persistence by gender on the medical pathway matches the clear gender shift in the contrast of national data about high school career plans (Sadler et al., 2012) and national data about medical school enrollment (Association of American Medical Colleges, 2017).

Second, our current data were limited by including only academic performance as predictors of students' decision-making. Continuing to develop discipline-specific explanatory mechanisms for attrition phenomena is an important direction for future work, which could be directly tested with future studies that include measures of mediating variables from expectancy-value theory and SCCT such as interest, value, and self-efficacy. While grades can be a strong proxy for self-perceptions of ability and identify formation, and exist in both theoretical frameworks as predictors of academic choices, there are likely a number of omitted psychological variables that mediate the relationship between grades and academic decisions. For example, the relationship between objective ability and academic persistence has been shown to be mediated by self-efficacy and goals (Brown et al., 2008). Future research in this area would benefit from gathering and including more direct measures of nonacademic factors such as competency beliefs in science and science identity to see if these can explain the differential attrition to STEM from this large pool of women who may be shifting away from an initial interest in science-relevant careers in Health, and toward non-STEM career pathways.

Third, the patterns observed in this study may not generalize across countries for a number of reasons. Most saliently, the medicine and many health profession degrees are graduate degrees in the United States, whereas many other countries allow students to pursue such degrees directly from high school (Riska, 2010). Requiring a full undergraduate degree first (instead of a shortened course of foundational science courses) creates opportunities for students to change career plans. On the positive side, this Bachelor's degree requirement may produce many more students who pursue science as a career. On the negative side, this requirement may produce more gender differences in STEM. Such effects may partially explain why the United States has one of the lowest proportions of women in medicine among OECD countries (Organization for Economic Co-operation and Development, 2018).

6 | CONCLUSION

While many studies identify gender inequity in STEM fields, it is increasingly important to become clearer in identifying how disciplinary differences may contribute to or mitigate these gaps. Understanding particular fields that are both pathways into, and pathways away from science, as well as beginning to define the driving mechanisms of those transitions, will help to focus interventions on those particular fields that continue to lag behind in equitable participation by gender. In this study, we identify two large incoming academic intentions that contribute the plurality of students who eventually earn science degrees: Health and Medicine. Further, we identify that while gender gaps exist in the production of science and STEM degrees from these two groups, the character and explanatory mechanisms for these two pathways as sources of STEM and science retention or attrition may differ. Future studies in gender differences in STEM attrition should attend to these often-understudied populations and build on prior qualitative and survey-based motivational research conducted in both K-12 (Jones, Howe, & Rua, 2000; P. H. Miller, Slawinski Blessing, & Schwartz, 2006) and undergraduate (Robnett, Chemers, & Zurbriggen, 2015; Seymour, 1995; Witherspoon, Vincent-Ruz, & Schunn, 2019) environments, to directly measure nonacademic factors that could provide additional explanatory power for these functions. Importantly, identifying these mechanisms can help to target interventions specifically toward these large science-related groups of students which could continue to help remove barriers to gender equity in STEM.

ACKNOWLEDGMENT

This study was supported by research grant Division of Undergraduate Education (DUE) 1524575 from the National Science Foundation.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ORCID

Eben B. Witherspoon http://orcid.org/0000-0003-0987-3679
Christian D. Schunn http://orcid.org/0000-0003-3589-297X

REFERENCES

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching*, *47*(5), 564–582. https://doi.org/10.1002/tea.20353

Association of American Medical Colleges. (2017). *AAMC FACTS Table (No. Tab. A-1)* . Retrieved from https://www.aamc.org/data/facts/applicantmatriculant/

Barr, D. A. (2010). *Questioning the premedical paradigm: Enhancing diversity in the medical profession a century after the flexner report*. Baltimore, MD: Johns Hopkins University Press.

Barr, D. A., Gonzalez, M. E., & Wanat, S. F. (2008). The leaky pipeline: Factors associated with early decline in interest in premedical studies among underrepresented minority undergraduate students. *Academic Medicine*, *83*(5), 503–511. https://doi.org/10.1097/ACM.0b013e31816bda16

Barr, D. A., Matsui, J., Wanat, S. F., & Gonzalez, M. E. (2010). Chemistry courses as the turning point for premedical students. *Advances in Health Sciences Education*, *15*(1), 45–54. https://doi.org/10.1007/s10459-009-9165-3

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

Beyer, S., & Bowden, E. M. (1997). Gender differences in self-perceptions: Convergent evidence from three measures of accuracy and bias. *Personality and Social Psychology Bulletin*, *23*(2), 157–172. https://doi.org/10.1177/0146167297232005

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, *17*(4), 369–386. https://doi.org/10.1080/09540250500145072

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior*, *72*(3), 298–308. https://doi.org/10.1016/j.jvb.2007.09.003

Brownell, S. E., Price, J. V., & Steinman, L. (2013). A writing-intensive course improves biology undergraduates' perception and confidence of their abilities to read scientific literature and communicate science. *Advances in Physiology Education*, *37*(1), 70–79. https://doi.org/10.1152/advan.00138.2012

Cannady, M. A., Greenwald, E., & Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: Is the STEM pipeline metaphor serving our students and the STEM workforce? *Science Education*, *98*(3), 443–460. https://doi.org/10.1002/sce.21108

Chen, X. (2013). *STEM attrition: College students' path into and out of STEM fields* (NCES 2014-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. https://doi.org/nces.ed.gov/pubs2014/2014001rev.pdf

Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin*, *143*(206), 1–35. https://doi.org/10.1037/bul0000052

Cromley, J. G., Perez, T., & Kaplan, A. (2016). Undergraduate STEM achievement and retention: Cognitive, motivational, and institutional factors and solutions. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 4–11. https://doi.org/10.1177/2372732215622648

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly*, *18*, 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Horowitz, G. (2009). It's not always just about the grade: Exploring the achievement goal orientations of pre-med students. *The Journal of Experimental Education*, *78*, 215–245. https://doi.org/10.1080/00220970903352746

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, *84*, 180–192. https://doi.org/10.1002/(SICI)1098-237X(200003)84:2%3c180::AID-SCE3%3e3.0.CO;2-X

Kimmel, L. G., Miller, J. D., & Eccles, J. S. (2012). Do the paths to STEMM professions differ by gender? *Peabody Journal of Education*, *87*(1), 92–113. https://doi.org/10.1080/0161956X.2012.642276

Koester, B. P., Grom, G., & McKay, T. A. (2016). *Patterns of gendered performance difference in introductory STEM courses*. Retrieved from http://arxiv.org/abs/1608.07565

Kolbrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average* (Report No. 08-5). New York, NY: The College Board.

Kugler, A. D., Tinsley, C. H., Akerlof, G., Autor, D., Butcher, K., Figlio, D., & Lavy, V. (2017). *Choice of majors: Are women really different from men?* (Working Paper No. 23735). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w23735

Lempp, H., & Seale, C. (2004). The hidden curriculum in undergraduate medical education: Qualitative study of medical students' perceptions of teaching. *British Medical Journal*, *329*, 770–773. https://doi.org/10.1136/bmj.38202.667130.55

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*, 79–122. https://doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Miller, M. J., Smith, P. E., Watford, B. A., Hui, K., & Lim, R. H. (2015). Social cognitive model of adjustment to engineering majors: Longitudinal test across gender and race/ethnicity. *Journal of Vocational Behavior*, *86*, 77–85. https://doi.org/10.1016/j.jvb.2014.11.004

Lent, R. W., Miller, M. J., Smith, P. E., Watford, B. A., Lim, R. H., & Hui, K. (2016). Social cognitive predictors of academic persistence and performance in engineering: Applicability across gender and race/ethnicity. *Journal of Vocational Behavior*, *94*(0827470), 79–88. https://doi.org/10.1016/j.jvb.2016.02.012

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., & Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: A meta-analytic path analysis of the social–cognitive choice model by gender and race/ethnicity. *Journal of Counseling Psychology*, *65*(1), 17–35. https://doi.org/10.1037/cou0000243

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 23–34. https://doi.org/10.1081/E-EWS

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., & Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *Journal of Personality and Social Psychology*, *96*(6), 1089–1103. https://doi.org/10.1037/a0015703

Mackinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Mackinnon, D. P., Lockwood, C. M., Hoffman, J. M., & West, S. G. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*(1), 83–104.

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, *23*(7), 129–149. https://doi.org/10.3102/00028312023001129

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology*, *54*(2), 263–280. https://doi.org/10.1037/dev0000393

Mattern, K., Camara, W., & Kobrin, J. L. (2007). *SAT® writing: An overview of research and psychometrics to date* (College Board Research Report No. RN, 32). Retrieved from https://doi.org/10.2307/257084

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., … Barmer, A. (2019). The Condition of Education 2019 (NCES 2019-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144

Mervis, J. (2012). Training and workforce. What if the science pipeline isn't really leaking? *Science*, *337*(6092), 280. https://doi.org/10.1126/science.337.6092.280

Miller, D. I. (2018). *Characterizing pathways for joining STEM in college and beyond* (Doctoral dissertation). Evanston, IL: Northwestern University.

Miller, J. D., & Solberg, V. S. (2012). The Composition of the STEMM workforce: Rationale for differentiating STEMM professional and STEMM support careers. *Peabody Journal of Education*, *87*(1), 6–15. https://doi.org/10.1080/0161956X.2012.642232

Miller, P. H., Slawinski Blessing, J., & Schwartz, S. (2006). Gender differences in high-school students' views about science. *International Journal of Science Education*, *28*(4), 363–381. https://doi.org/10.1080/09500690500277664

Morgan, S. L., Gelbgiser, D., & Weeden, K. A. (2013). Feeding the pipeline: Gender, occupational plans, and college major selection. *Social Science Research*, *42*(4), 989–1005. https://doi.org/10.1016/j.ssresearch.2013.03.008

Murayama, K., & Elliot, A. J. (2012). The competition-performance relation: A meta-analytic review and test of the opposing processes model of competition and performance. *Psychological Bulletin*, *138*(6), 1035–1070. https://doi.org/10.1037/a0028324

National Science Foundation. (2015). *Integrated postsecondary education data system, 2015, completions survey*. Retrieved from https://webcaspar.nsf.gov

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, *24*(2), 129–144. https://doi.org/10.1257/jep.24.2.129

Organization for Economic Co-operation and Development (OECD). (2018). *Health care resources*. Retrieved from https://stats.oecd.org/viewhtml.aspx?datasetcode=HEALTH_REAC&lang=en#

Riska, E. (2010). Women in the medical profession. In E. Kuhlmann & E. Annandale (Eds.), *Palgrave handbook of gender and healthcare* (pp. 389–404). Retrieved from http://ebookcentral.proquest.com/lib/pitt-ebooks/detail.action?docID=1033794

Robnett, R. D., Chemers, M. M., & Zurbriggen, E. L. (2015). Longitudinal associations among undergraduates' research experience, self-efficacy, and identity. *Journal of Research in Science Teaching*, *52*(6), 847–867. https://doi.org/10.1002/tea.21221

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, *96*(3), 411–427. https://doi.org/10.1002/sce.21007

Sadler, P. M., & Tai, R. (2007). Advanced placement exam scores as a predictor of performance in introductory college biology, chemistry and physics courses. *Science Educator*, *16*(1), 1–19. Retrieved from http://eric.ed.gov/?id=EJ783418

Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education*, *79*(4), 437–473. https://doi.org/10.1002/sce.3730790406

Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles*, *66*(3–4), 175–183. https://doi.org/10.1007/s11199-011-0051-0

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles*, *67*(11–12), 617–629. https://doi.org/10.1007/s11199-012-0209-4

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*(1982), 290. https://doi.org/10.2307/270723

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American Psychologist*, *60*(9), 950–958. https://doi.org/10.1037/0003-066X.60.9.950

Thoman, D. B., Arizaga, J. A., Smith, J. L., Story, T. S., & Soncuya, G. (2014). The grass is greener in non-science, technology, engineering, and math classes: Examining the role of competing belonging to undergraduate women's vulnerability to being pulled away from science. *Psychology of Women Quarterly*, *38*(2), 246–258. https://doi.org/10.1177/0361684313499899

Vincent-Ruz, P., Binning, K., Schunn, C. D., & Grabowski, J. (2018). The effect of math SAT on women's chemistry competency beliefs. *Chemistry Education Research and Practice*, *19*, 342–351. https://doi.org/10.1039/C7RP00137A

Wang, M. Te, Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science*, *24*(5), 770–775. https://doi.org/10.1177/0956797612458937

Witherspoon, E. B., Vincent-Ruz, P., & Schunn, C. D. (2019). When making the grade isn't enough: The gendered nature of premed science course attrition. *Educational Researcher*, *48*(4), 193–204. https://doi.org/10.3102/0013189X19840331

Yalvac, B., Smith, H. D., Troy, J. B., & Hirsch, P. (2007). Promoting advanced writing skills in an upper-level engineering class. *Journal of Engineering Education*, *96*(2), 117–128. https://doi.org/10.1002/j.2168-9830.2007.tb00922.x

How to cite this article: Witherspoon EB, Schunn CD. Locating and understanding the largest gender differences in pathways to science degrees. *Science Education*. 2020;104:144–163. https://doi.org/10.1002/sce.21557

APPENDIX A

Mediation models with GPA calculated through to a later course on the pre-medical and pre-health pathway (Organic Chemistry 1).

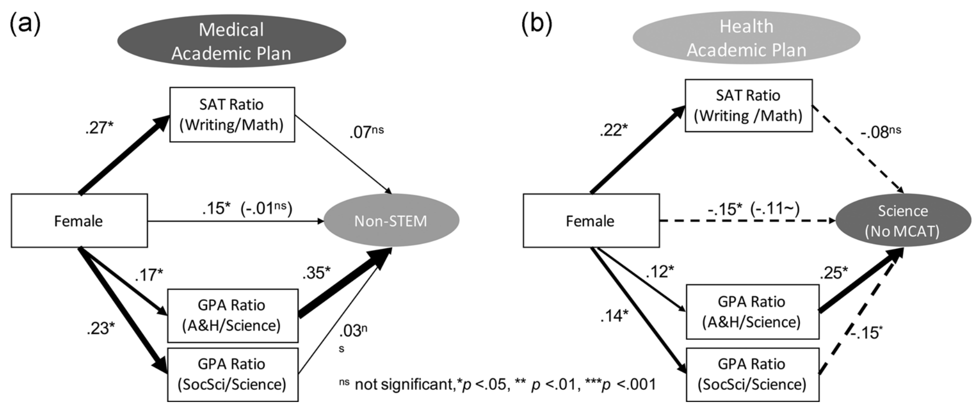

FIGURE A1 Mediation models of the (a) Medical and (b) Health pathways, using GPA ratios up to and including Organic Chemistry 1. The indirect effects of the GPA mediators for students intending Medicine explain 95% of the direct effect between gender and graduating with a non-STEM degree, and the indirect effects of the GPA mediators for students intending Health explain 6% of the direct effect between gender and earning a Science degree (without taking the MCAT). (Thickness of lines indicates relative strength of associations; dashed lines indicate a negative association. Covariation between all academic variables included in model but not shown for clarity.) GPA, grade point average; MCAT, Medical College Admissions Test; SAT, Scholastic Aptitude Test; STEM, science, technology, engineering, and mathematics