Localizing and Understanding Mechanisms of Gender Differences within Pathways Towards and Away from Science Degrees

by

Eben Blake Witherspoon

B.A., Colby College, 2009

Submitted to the Graduate Faculty of

the School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Eben Blake Witherspoon

It was defended on

August 21, 2019

and approved by

Kevin R. Binning, Assistant Professor, Psychology

David I. Miller, Ph. D., American Institutes of Research

Elizabeth Votruba-Drzal, Professor, Psychology

Christian D. Schunn, Professor, Learning Sciences and Policy

Localizing and Understanding Mechanisms of Gender Differences within Pathways Towards and Away from Science Degrees

Eben Blake Witherspoon, PhD

University of Pittsburgh, 2019

Despite decades-old research revealing gender differences in retention in science, technology, engineering, and mathematics (STEM), persistent gaps in women's participation remain in some undergraduate science courses and majors. Across three studies, this dissertation seeks to better identify the location and sources of persistent gender differences in pathways into and out of the Sciences, with a special focus on an understudied population that drives many of the larger trends: undergraduate pre-medical students.

In part, the studies relate the persistence of gender gaps in the sciences to an over-application of the "STEM Pipeline" metaphor, which ignores a number of other factors that play a role in undergraduates' choice of majors and careers, including historical gender differences within particular science domains. I show that pre-health and pre-medicine are particularly important pathways for understanding gendered attrition in science because they represent a large population of students enrolling in introductory science courses. In addition, these pathways produce a high proportion of eventual science degree earners. However, relative attrition by gender persists within the long sequence of required pre-med science courses, especially in later physical sciences courses (i.e., Organic Chemistry, Algebra-based Physics).

In addition to localizing these phenomena for pre-medical students at particular points in time, the studies focus on mechanisms both academic (as relative successes and failures) and social-psychological, through which students' experiences in these courses influence their beliefs

about their abilities, their performance, and eventually their decisions to persist within science courses and majors. The results from this collection of studies brings depth and specificity to a pathways model as a more accurate alternative to the problematic pipeline model of STEM participation, by identifying courses and discipline-specific psychological mechanisms along under-explored, high-bandwidth pathways to science, which could be targets for intervention. Importantly, this approach shifts the focus of sources of gender differences in the sciences away from immutable, pre-existing differences in prior experiences and performance, and towards emphasizing the agency higher education institutions have in addressing more malleable, concurrent elements of women's experiences in undergraduate physical science courses, which either work to perpetuate or mitigate earlier differences.

Table of Contents

List of Tables

List of Figures

Acknowledgements

"So many things are possible, just as long as you don't know they are impossible."

- Norton Juster, *The Phantom Tollbooth*

To my mentor and advisor Chris Schunn, I am truly thankful for the endless hours you put into supporting me in achieving this goal. You were always present and ready to meet me where I was, and I appreciate the commitment you showed in helping me build the tools I needed to find answers, rather than giving me the answers I thought I needed. Thank you for practicing what you preach in your approach to developing independent researchers, and for continuing to believe that I could become one.

A huge thank you to my committee members, Kevin Binning, David Miller, and Elizabeth Votruba-Drzal, for guiding and advancing my thinking in this work. It's been an honor getting to know each of you through this process, and I would be proud if my future work is able to continue the intellectual lineage of the important and valuable research you all do. I would also like to thank all of the LSAP faculty, and especially Mary Kay Stein, for providing consistent support and feedback around research, writing, and navigating academic life. I also want to thank my fellow LSAP and School of Education colleagues, and especially Monica Navarro, Ross Higashi, and Paulette Vincent-Ruz, for their contributions, support, advice and solidarity during all the struggles and accomplishments that come with doctoral studies.

To my entire family, and especially my parents, without whom none of this would be possible: thank you for taking the road less travelled. You continue to provide a model for a life lived with principle, and have always remind me to stay focused on the things that matter most. In a long list of unearned privileges that I have been afforded in my life, having you as my parents is by far one of the greatest.

Finally, thank you to my wife and best friend, Amy, who in this endeavor, as in all things, has been my steadfast advocate and source of support. I will forever be in awe of your resilience, equanimity, and kindness, and I feel incredibly grateful to have such an amazing partner in my corner to share this and many other adventures with. I love you.

1.0 Introduction

There has been a proliferation of research in education in the past few decades about how to best support students' persistence in science, technology, engineering, and mathematics (STEM) within K-12 (Dorph & Bathgate, 2017; Sadler, Sonnert, Hazari, & Tai, 2012), through higher education (X. Chen & Soldner, 2013; Morgan, Gelbgiser, & Weeden, 2013), and continuing into careers in those fields (Cannady et al., 2017; Kimmel, Miller, & Eccles, 2012). Prior estimates have predicted the United States would need a 34% annual increase in students who receive undergraduate STEM degrees to meet rising demand for all STEM professionals (PCAST, 2012). More recent predictions still show the demand for STEM jobs in the life and physical sciences in the U.S. increasing by 10% through 2026, faster than the average growth rates for other occupations (U.S. Bureau of Labor Statistics, 2017a). This does not account for the increasing number of STEM support careers that also require some science knowledge and training: science teachers, health technicians, support professionals, and even science journalists. Current reports show that the number of jobs that require at least some level of technical expertise in science is three times the number of science graduates now working in those jobs, with others estimating a 13:1 ratio of science jobs to qualified job seekers (National Science Board, 2018; New American Economy Research Fund, 2017).

In order to fulfill the high level of demand for workers in these careers, a number of reports (National Science Board, 2015; PCAST, 2012) have argued that it will be necessary to expand participation in the science workforce to a broader and more equitable sample of the population. The economic and societal importance of broader participation in the science workforce is more than simply matching supply with demand; diversity in science has been shown to be important for promoting continued innovation (Atkinson & Mayo, 2011; Page, 2007). Further, equitable representation is also critical for ensuring that research and development remains responsive to problems impacting minoritized populations historically underrepresented in science fields. For example, health disparities continue to persist for women, and particularly women of color, even when controlling for a variety of other contributing factors (Williams, 2008); some of these disparities have been linked to differences in the patterns of the medical practices of male and female physicians (Siriwardena et al., 2012; Tsugawa et al., 2017). Further, science careers are typically higher paying, generating a possible avenue for economic empowerment for students who graduate from college well-prepared to enter science careers (Beede et al., 2011; Oh & Lewis, 2011). However, large differences in earning can still appear between particular sciences disciplines; for example, there is a $9,000 median wage gap between students earning a Bachelor's degree in the physical sciences and those earning a degree in the life sciences (Carnevale, Cheah, & Hanson, 2015). While not strictly required, a majority of science careers in the future are much more likely to be given to those with a bachelor's degree and some undergraduate training in the sciences. Therefore, research in STEM persistence should also attend to broader concerns about systemic inequities in access and opportunity within higher education and within and between science careers.

1.1 From STEM Pipelines to Science Pathways

Much of the research on STEM persistence focuses on STEM as an aggregate outcome, perpetuating a problematic "pipeline" metaphor. The typical presentation of the STEM pipeline suggests a single, direct pathway to STEM careers, with specific points along the way that predict persistence, and which often differentially “leak” historically underrepresented groups (Alper, 1993; Berryman, 1983). While initially useful for identifying broad trends and a general focal area for further research, continuing this pipeline metaphor runs the risk of ignoring the variety of pathways students may take to reach degrees and careers related to science, technology, engineering, and mathematics (Blickenstaff, 2005; D. I. Miller, 2018). Research in K-12 education has shown that factors previously identified as early indicators of STEM participation—such as middle school science experiences, or taking calculus and Advanced Placement courses in high school—actually account for only a small number of eventual STEM participants (Cannady, Greenwald, & Harris, 2014; Husbands Fealing & Myers, 2012). Within higher education, research often has focused on predicting degree earning from factors such as intended field of study or initial selection of an academic major upon entering college, an approach ignoring the high degree of variation in students plans throughout the first few years of college. This pipeline approach to research in STEM persistence in higher education can therefore potentially obscure a number of different large pathways taken by undergraduate students towards and away from science degrees which may appear after they declare their initial academic intentions.

There are many possible pathways that can be followed between, towards, and away from various science and science-related degrees (see Figure 1). While some students will remain within the pathway they initially declare, many other students will “try on” a variety of pathways as they consider various career and post-graduation goals, and as their initial academic choices shift. Sometimes these pathways will lead towards a science degree, and sometimes these pathways will lead away from science. Particularly in a liberal arts undergraduate environment in the US that typically requires various general education courses, the pathway taken will be influenced by both the quality of experiences students have in the core courses for their selected majors (e.g., a bad chemistry class experience by a chemistry major; Barr, Matsui, Wanat, & Gonzalez, 2010), as well as the experiences students have in their non-major general education requirements (e.g., an excellent history class experience by a chemistry major; or an excellent chemistry class experience by a history major; Milsom & Coughlin, 2016).

Understanding the overall patterns of degree and career outcomes will require understanding which pathways are most commonly taken. The relative size of a pathway towards any given outcome will depend upon both the size of the starting population (e.g., how many students begin with a given career goal) and the relative frequency at which students reach a particular outcome from any given starting place (e.g., are students more likely to continue towards a science career goal or more likely to change to another career goal).

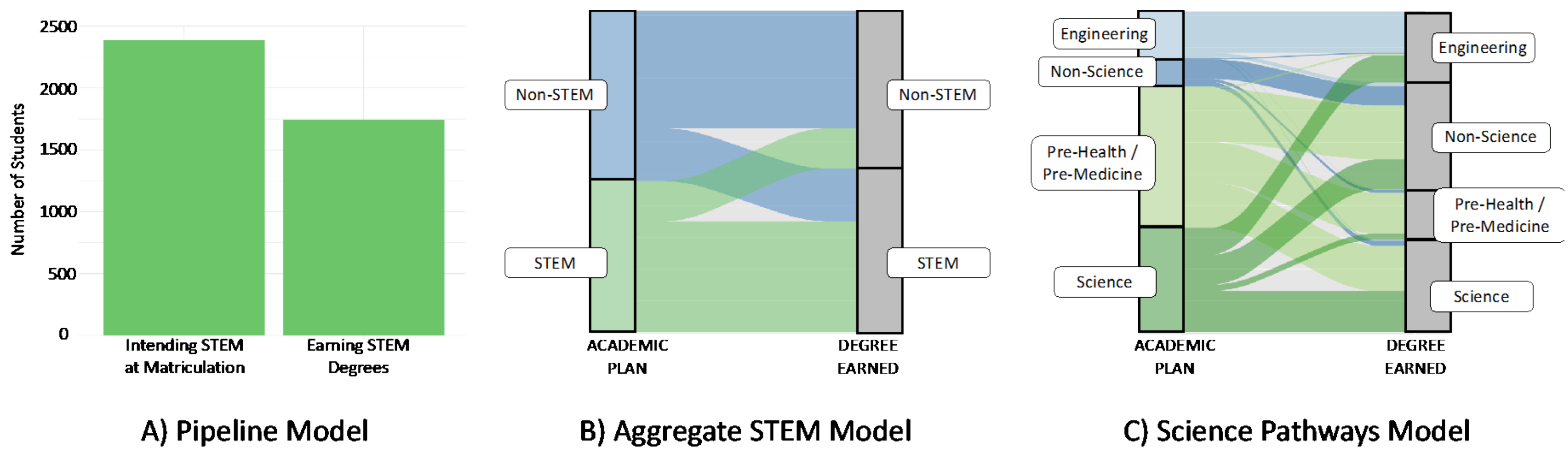


Figure 1. An example of three potential modeling paradigms for STEM attrition

A) STEM pipelines, B) aggregate STEM pathways model and C) science pathways.

1.2 Equity vs. Equality in Gender Variation by Science

The pathways framework also provides a different lens for considering gender diversity in STEM. It is well established that, on average, women have been historically underrepresented across most undergraduate STEM majors and careers (Koester, Grom, & McKay, 2016; Maltese & Tai, 2011; Seymour & Hewitt, 1996; Xie & Shauman, 2003). Recent reports suggest that women now earn more than half (58%) of the bachelor's degrees awarded in the United States, and an equal share (50%) of STEM degrees (National Center for Science and Engineering Statistics, 2017). However, within STEM, the number of women earning undergraduate degrees has surpassed men in some science fields (e.g., Biology, 58%); while gender gaps in degrees earned by women have persisted or increased in other male dominated fields (i.e. Physics, 18%; Engineering, 19%, Computer Science, 18%; see Cheryan, Ziegler, Montoya, & Jiang, 2016; National Science Foundation, 2015). Focusing on gender equality in STEM in the aggregate can ignore the extent to which in which gender differences may continue to remain siloed within particular science disciplines (see Figure 2).

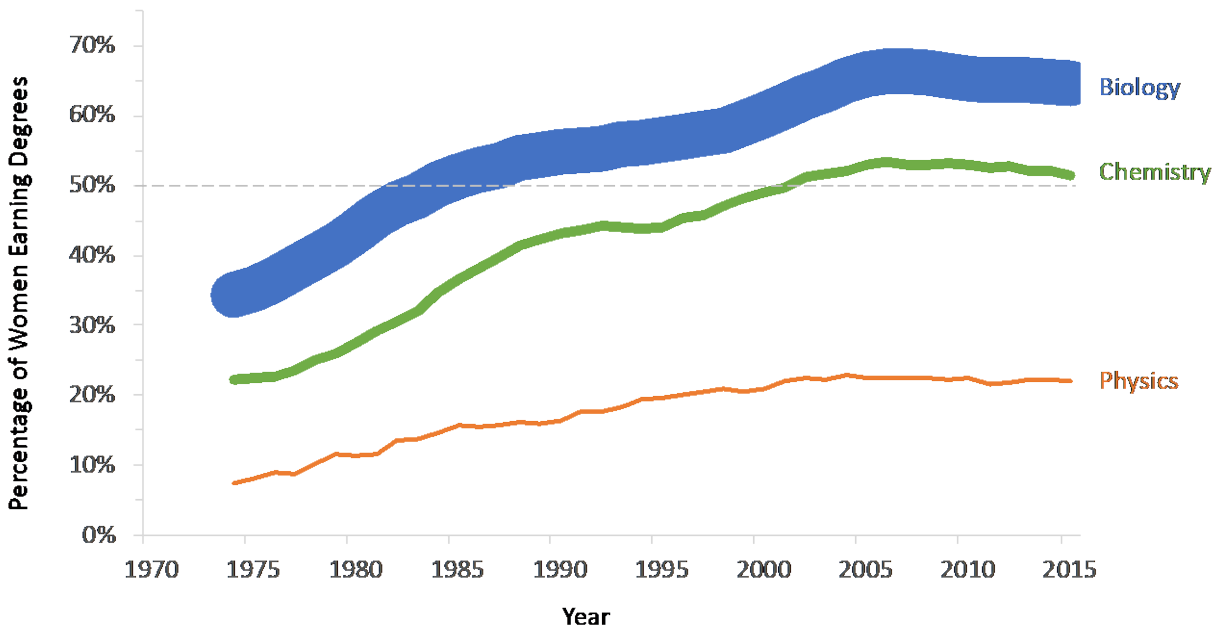


Figure 2. The proportion of women earning degrees in Biology, Chemistry, Mathematics and Physics, as a 5-year moving average from 1970 – 2015.

Line thickness represents relative size of degree earning population by discipline. (DATA SOURCE: National Science Foundation WebCASPAR, 2019; https://ncsesdata.nsf.gov/webcaspar/)

Ignoring these differences can be problematic for educational equity, both from the perspective of reducing barriers to students' ability to pursue their individual interests, as well as through the lens of economic equity in pay and influence in science. For example, from an economic-opportunity perspective, some of the largest-growing and highest-paying STEM fields, such as computer science and engineering, exhibit the smallest gender gaps in pay, and yet remain highly gendered in terms of participation (Cheryan et al., 2016; Oh & Lewis, 2011). Medical and health pathways may provide another example where *equality* in women's participation in a science field may obscure a lack of *equity*. That is, despite reaching parity in women's participation at the point of medical school, failing to investigate further may obscure differential attrition within pre-medical science courses when considering the *proportion* of men and women who initially enter intending pre-med (Fiorentine, 1987).

1.3 Knowing Where vs. Knowing Why Gendered Variation Appears Along Pathways

A critical component of understanding where gender variation occurs along pathways to science is to begin to also understand the social and psychological mechanisms that produce these differences, and how they may be situated within the context of a particular discipline. An extensive foundational study of undergraduate attrition from the sciences by Seymour and Hewitt (1996) shows that a large proportion of students, and women in particular, choose to leave science despite demonstrating a high level of academic performance, suggesting factors other than ability are likely to be large contributors to these attrition decisions. More recent meta analyses have demonstrated that the scant evidence of relatively small biological sex differences in motivation or ability are insufficient for explaining large differences in participation, therefore suggesting these differences are socially constructed (Hyde, 2005). Importantly, while acknowledging the influence prior academic experiences in a particular discipline can have, this line of work suggests instead attending to malleable aspects of the concurrent context, which may also heavily contribute to the perpetuation of gendered attrition in the sciences.

Researchers have pointed to constructs within social cognitive theory (Bandura, 1986) as particularly valuable for understanding the interplay between the individual, behaviors and context (Cromley, Perez, & Kaplan, 2016; Trujillo & Tanner, 2014). Frameworks like Expectancy-Value Theory (EVT) and Social Cognitive Career Theory (SCCT) propose that academic decisions like studying for an exam, re-enrolling in courses, or persisting in a major are driven in part by students' expectations of success (i.e., self-efficacy, expectancies of success), the importance of a particular goal (i.e., interest, values), and the barriers and supports to persistence that exist in a particular context (Eccles, 1994; Lent, Brown, & Hackett, 1994, 2000; Wigfield & Eccles, 2000). Because of similarities in these two models, some scholars have recently suggested combining elements from each theory to simplify into two major constructs of achievement motivation; “Can I do this?” (i.e., self-efficacy beliefs, expectancies of success) and “Do I want to do this?” (i.e., utility value, interest; Linnenbrink-Garcia et al., 2018).

Therefore, in my approach to understanding gender equity in the sciences, it is not enough to simply look at equality in aggregate outcomes or focus on prior indicators of success. Instead, it will be necessary to understand the specific way in which men and women’s decisions to continue towards or move away from science may differ by the specific pathways the follow, the context of specific courses and disciplines, and in the concurrent social and psychological mechanisms that drive students’ decisions in those contexts.

1.4 Overview of the Literatures Reviewed

My empirical work in the dissertation takes a three-stage approach. First, at the largest grain size, it is useful to identify common pathways to science in order to reveal the relative importance of each pathway towards and away from science for men and women (e.g., from medicine goals to science degrees). Second, at an intermediate grain size, it then becomes possible to locate the particular points along pathways that produce the gendered patterns of retention or attrition (e.g., organic chemistry). The existence of course-specific points of gendered attrition begin to establish a more nuanced understanding of gendered attrition, as well as setting a priority for where to study underlying mechanisms of attrition. Third, at the most fine-grained size, it is possible to test the specific psychological mechanisms that contribute to gender differences in persistence within these specific courses (e.g., the role of chemistry self-efficacy for organic chemistry or having a high sense of belong in the physics classroom). Work at this grain size can then become the basis for eventual recommendation and design of efficient, targeted interventions that aim to address malleable aspects within the context of these observed attrition patterns.

Across this collection of studies implementing this three-staged approach, I review a range of research literatures, including 140 papers in total. Because a different set of prior work (although clearly with some overlap) is relevant to each of the three stages, the literature reviews are separated by each stage, and they are therefore presented within each chapter. Rather than redundantly reviewing all of those combined studies here, this chapter provides a brief overview of the covered literature reviews. The following paragraphs summarize the main topics, noting when one topic is relevant to two of the three separate studies (see Table 1 for a summary).

Table 1. Literatures reviewed and location of reviews in the collection of studies.

Topic	Study 1	Study 2	Study 3
Pathways to Science	Limitations of the STEM pipeline model	Pre-med and pre-health as large pathways that contribute a high number of students to introductory science courses	
Course-Specific Gender Differences		Gender differences in Physical Science courses	
		Gender differences in Life Sciences courses	
Social-Psychological Mechanisms	The role of the perceptions of relative academic performance on retention		
		The role of competency beliefs on grades and retention	
		Expectancy-Value Theory (EVT)	Contextual sources of competency beliefs

1.4.1 What are the Largest Undergraduate Pathways to Science Degrees by Gender?

While it is important to attend to differences in common science, technology, engineering, and mathematics pathways, it is also reasonable to consider alternative pathways that prepare for careers related to the sciences. In Study 1, I review prior work that has approached STEM participation from the aggregate pipeline approach and weigh the relative advantages and disadvantages of this paradigm. In particular, I demonstrate how certain early markers of STEM pipeline participation that have been identified have been recently shown to in fact be relatively poor indicators of actual persistence in STEM. I also in Study 1 point to examples that demonstrate how focusing on students taking a variety of different science courses can begin to highlight the location of disciplinary differences, that may begin to break up our conception of STEM as a monolithic domain.

As part of Studies 2 and 3, I review more specific evidence demonstrating how students intending to study the medical and health professions provide an example of one large and gendered degree pathway that is often overlooked, and yet may act as an important on-ramp to science. I acknowledge that some research considers students who switch into health and medicine as examples of a loss to STEM (Lindemann, Britton, & Zundl, 2016), despite evidence to suggest that almost half of students in the US enrolled in introductory undergraduate science courses do so with the intent to pursue medicine. However, I instead build on research suggesting that interest in medicine acts as an early catalyst towards science participation, which may later shift towards interest and retention in specific science disciplines (Crowley, Barron, Knutson, & Martin, 2014), leading us to adopt the expanded acronym "STEMM" (science, technology, engineering, mathematics and medicine) that includes this large population of students (J. D. Miller & Kimmel, 2010).

1.4.2 Where are the Specific Courses that Show Patterns of Attrition by Gender?

Gender differences during the undergraduate years can occur at the level of academic major (e.g., who majors in physics or engineering vs. biology or neuroscience). However, because all science and science-related majors in the US require students to take courses from various disciplines, it is also important to consider that gender differences can appear from experiences in various courses across the different disciplines which are required within a major. Similarly, there can be additional broad coursework requirements based on later career or graduate school intentions that may provide experiences that vary by gender. Therefore, in addition to focusing on how the disaggregation of STEMM can illuminate diverse pathways to science degrees, attending to differences between specific courses within different science disciplines may help pinpoint where gender differences in those pathways appear.

As part of Studies 2 and 3, I review the evidence in the literature that identifies key differences by gender in courses both between the broader categories of the life sciences (i.e., Biology and Neuroscience) and the physical sciences (i.e., Chemistry and Physics), as well as a body of work in pre-medicine that locates retention issues within specific physical science courses often required along the pre-medical pathway: Organic Chemistry and Introductory Physics. In addition to the existence of large differences in women's representation in the physical sciences, particularly in pre-medicine these physical science courses have been shown to be large contributors to a loss of interest and decision to switch out of pre-medicine (Barr, 2010). In Study 3, I even more narrowly focus on one required pre-medical course in the physical sciences primarily enrolling non-majors and identified as a potentially problematic domain in terms of gender differences in grades, attitudes and attrition: Algebra-based Introductory Physics.

1.4.3 Which Academic and Attitudinal Mechanisms Predict Gendered Persistence?

In parallel with our review of various pathways to science and specific courses that may show large differences in grades, attitudes and persistence by gender, throughout the three studies I also conduct an examination of the educational psychology literature regarding plausible academic and attitudinal mechanisms that could produce these observed gender differences.

For Study 1, I primarily situate the review within a hypothesis about the impact of domain-relative academic performance on grades, persistence, and beliefs about abilities (Breda & Napp, 2019; Marsh, 1986; M. Te Wang, Eccles, & Kenny, 2013). This line of work suggests that in addition to using academic performance feedback to compare their own abilities to the abilities of others, students also compare academic performance feedback within one particular domain to their own relative performance in other domains, in order to make persistence decisions and draw inferences about abilities. For example, women may move towards social sciences and humanities because of relative strengths and higher performance in writing and verbal performance.

For Study 2, I again revisit the literature of relative academic performance, but expand the review of academic mechanisms to include potential attitudinal mechanisms that are shown to also exert a strong influence on students' grades and retention. Specifically, I identify specific constructs related to a broader motivational framework Expectancy-Value Theory (EVT; Eccles, 1994) which have shown particularly large differences for men and women in science courses: interest (i.e., students' intrinsic interest or enjoyment); identity (i.e., the relationship between the task and the students' sense of self) and especially competency beliefs (i.e., whether or not students expect to be successful in the course based on beliefs about their abilities in the subject).

Finally, in Study 3, which focuses on physical sciences, I focus on the competency beliefs literature because it is a particularly powerful source of grades and attrition that often show gender differences in the physical sciences. I begin with a review of sources of competency beliefs that have been identified previously in the literature, in order to direct attention towards unpacking how gender differences in competency beliefs develop throughout a students' experience with an undergraduate physical science course. Then, I propose connections between the extant literature on different categories of self-efficacy sources, and the literature from two constructs used to measure students' perception of their science learning environments; sense of belonging, and implicit theories of intelligence. Importantly, I also demonstrate evidence that these constructs may be highly malleable, providing an avenue for intervention.

1.5 Summary of the Literature Review

To summarize, a review of the literature in each article leads to the following primary open questions with each of the three layers. First, while there has been a recent call to move away from the STEM(M) pipeline metaphor and towards an understanding of a variety of pathways to science-related careers, there is little research that has investigated the relative size and gendered nature of the some of the most common undergraduate pathways to science degrees. Pathways to the medical and health professions represent one relatively understudied sequence of science courses, which may also be a large contributor to students who eventually earn a science degree.

Second, despite research showing that large differences in persistence by gender remain between the different science disciplines that make up STEMM (i.e., Biology vs. Physics), less is known about the location and character of specific science courses that are required along these large pathways to science, which may show differential attrition by gender. Importantly, more research is needed to identify whether the academic and motivational mechanisms for these course-level gender differences that may cumulate to produce large gender differences in attrition at the degree level are discipline-specific, and how these are related to perceptions of relative performance in those disciplines.

Finally, while some studies continue to note differences in academic performance, retention, and beliefs about ability between men and women in undergraduate physical science courses, there is less certainty in the literature of the sources of individual ability beliefs, and how perceptions of the specific disciplinary context may mediate the development of ability beliefs in ways that are differential by gender. Therefore, while grades may continue to be the most reliable direct indicator of retention to subsequent courses, it is also important to begin understanding the mechanisms through which relative influences of prior performance and the concurrent development of concurrent attitudes within a disciplinary context influence these more distal outcomes. This can provide important insight into how these mechanisms may operate differently for men and women in physical science courses and identify fruitful points for intervention.

1.6 Overview of the Methodological Approach

1.6.1 Study Designs that Balance Internal and External Validity

The literature on gender gaps in science retention reviewed above can be divided roughly into two categories of methodological designs. The first set are descriptive and correlational studies, which aim to observe patterns in large datasets that are drawn from a representative sample of the population. The size, scope, and increased statistical power of these studies can afford them with a high level of *external validity*, improving the claims that findings can be generalized to a broad population. The second set are experimental or quasi-experimental studies, which aim to test the validity of an underlying theoretical mechanism. These studies are therefore typically conducted within a tightly controlled laboratory environment, or purposefully draw more localized inferences from smaller samples using detailed surveys or qualitative interviews and observations. These design features afford the researcher with a high level of *internal validity*; that is, the inferences drawn in any given study are unlikely to be confounded by potential exogeneity

(Shadish, Cook, & Campbell, 2002). However, these same features limit the external validity of these studies; it is difficult to know whether or not findings from a highly controlled experimental context will generalize to a broader implementation in context, where a large number of confounding factors can potentially be introduced.

Practices informed by implementation science, developed in fields like public health, provides a blueprint for an alternative methodological approach which helps bridge tensions between external and internal validity raised above (Kilbourne, Bauer, Damschroder, Hagedorn, & Smith, 2015). This method suggests initially conducting exploratory studies using large longitudinal and cross-sectional datasets to generate hypotheses, which are then refined through theory and motivate more targeted small scale investigations that can get at potential explanatory mechanisms (Penuel & Fishman, 2012; Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007; Spillane, Reiser, & Reimer, 2007). An important step in this approach is identifying the hypothesized *unit of change* (e.g., a particular science course within pre-med) where an intervention could be effective, while remaining attuned to how differences in context may introduce barriers or supports to interventions (Confrey, Castro-Filho, & Wilhelm, 2000). In this way, non-experimental methods can be employed as a starting point for the development of experiments that are both based in theory and grounded in observations of "real world" phenomena (Archer et al., 2012). Important in the case of institutional research of science persistence in higher education, beginning with an large-scale assessment of a cross-disciplinary pathways like pre-medicine can facilitate the breaking down of "data silos", where student information becomes isolated within specific science departments, further reducing the external validity of findings discovered within each (Atkinson & Mayo, 2011). Further, this approach facilitates a more comprehensive assessment of psychological frameworks that brings together prior and current academic performance indicators, as well as attitudinal factors, to predict women's persistence within specific pathways to science.

1.6.2 Using Logistic Regression to Analyze Retention

Analyses of outcomes such as retention or enrollment often use logistic regression, a special case of regression which provides a better fit to the residuals when the dependent variable is binary or categorical. A review of studies in higher education using this method show that over half of the studies reviewed were analyses of retention or enrollment (Peng, So, Stage, & St. John, 2002). Another probabilistic method commonly used in studies of retention that are particularly interested in the timing of attrition is survival analysis, a method developed primarily in the medical fields for analyses of mortality rates (Collett, 1994). Survival analysis uses the observations remaining in the study at each time point to determine the probability of persistence to the next time point. However, the advantages of this method in medicine make them less beneficial for analyses of STEMM pathways, which is specifically interested in how students exit but also re-enter degrees at various points.

An advantage of logistic regression over other non-linear tests is ease of interpretability; results can be reported as an odds ratio, or a ratio of the proportions, which provides an effect size that is more intuitive than other non-linear methods, particularly when comparing the probabilities of retention between two groups like in the case of analyses by gender. Further, logistic regression can provide more accurate estimates for models using a large number of predictors, even with a relatively small sample size; studies suggest only a minimum observations-to-predictor ratio of 10:1 is needed, rather than 40:1 for OLS regression (Long, 1997). However, logistic regression becomes more cumbersome when there are multiple outcomes and mediational variables. Further, like any kind of regression, it is strictly correlational; therefore, it provides limited evidence for causal hypotheses or the directionality of effects.

1.6.3 Using SEM to Test Mechanisms

In order to provide a holistic assessment of the combination of prior and current academic factors, and contextualized motivational factors that may influence science persistence, in these studies I will utilize structural equation modeling (SEM) that help to address the possibility of bidirectional effects in a single model. For example, SEM can test whether a model that has direct effects between self-efficacy and performance is the best fit, or if fit is improved when self-efficacy instead indirectly effects performance through interest (Ainley & Ainley, 2019). Rather than accounting for the relationships between the predictors by holding them constant as in multiple regression, SEM combines both a structural path model which considers the relationships between predictors as well as with outcomes, as well as a latent measurement model inferred from observed variables. SEM modeling uses a maximum likelihood estimator to attempt to fit the hypothesized model to the data, and provides information about the adequacy of the model fit to the data (Hu & Bentler, 1999). While SEM can provide a test of the causal inferences assumed by the design of the structural model, however, it does not allow for strict causal inference when used with correlational data (Bollen & Pearl, 2013).

In addition to using SEM, the collection of longitudinal measures can contribute to strengthening the case for causal inference about the relationships between various predictors of gendered performance and persistence. Importantly for motivational research, longitudinal SEM offers a more rigorous test of mediation, and provides a stronger assumption of temporal precedence.

1.7 The Current Set of Studies

Building on the prior literature and consideration of methodological approaches, across the three studies presented in this proposal I take an approach informed by implementation science that attempts to localize the gendered nature of pathways to science participation. Using extensive data from a single large research university in the northeastern United States as a model for the larger phenomenon, I begin with a descriptive analysis using logistic regression to understand the variety of pathways that students take to and away from science degrees, and the extent to which these pathways are differential by gender. Next, I again use logistic regression to analyze the same institutional dataset, this time following multiple cohorts of students within medicine, one of the largest pathways to science, to isolate the particular courses along the medical pathway where gender differences occur. Then, I use a small survey sample and structural equation modeling to begin to test hypotheses about the potential academic and psychological mediators of differential attrition in those courses and understand if and when gender differences appeared throughout a series of science courses during the undergraduate years. Finally, I use a more comprehensive sample of longitudinal surveys to conduct a deep dive into a single course showing the largest gender differences in retention, to further investigate the mechanisms through which these differences appear using both institutional data on prior academic preparation and current performance, as well as longitudinally test the relationships between psychological factors that have been related to gender differences in science persistence.

1.8 Organization of the Dissertation

Chapters 2, 3, and 4 each present one of the studies described above. These studies were all supported by research grant Division of Undergraduate Education (DUE) #1524575 from the National Science Foundation. Each chapter is a fully contained article, following a three-article dissertation structure, complete with literature review, methods, results, and discussion. Chapter 5 presents reflection looking across the three empirical studies, discussing common themes and layouts out ideas for next steps in this research area.

2.0 Study 1: Locating and Understanding the Largest Gender Differences in Pathways to Science

(Manuscript under review at Science Education, 2019)

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 towards 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.

2.1 Introduction

It is well established in the literature that despite recent progress towards 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 et al., 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 two 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 post-secondary 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 et al., 2016; Leslie, Cimpian, Meyer, & Freeland, 2015). Gender differences in attrition have been shown to be highly variable by discipline, both between science, technology, engineering and math, as well as within the sciences (Cheryan et al., 2016; D. I. Miller, 2018). Further, gender differences in attrition have been shown to be highly variable by discipline, both between science, technology, engineering and math, as well as within the sciences (Cheryan et al., 2016; 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 (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 et al., 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

(X. Chen & Soldner, 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.

2.2 Theoretical Framework

2.2.1 Relative Academic Performance

Reports on the Scholastic Aptitude Test (SAT) have shown that performance differences across the test vary by gender, with a large point 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 two 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 et al., 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 et al., 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” towards 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 socio-cultural 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 et al., 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 (Marsh, 1986; M. Te Wang et al., 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

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.

2.2.2 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 towards 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?

2.3 Methods

2.3.1 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 (*Mdn* = $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 prior to 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 Black or Hispanic (14%) and Asian (12%) students making up the next two largest ethnic groups. 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 was provided for analysis with Institutional Review Board approval.

2.3.2 Measures

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 5 separate general categories (see Table 2): 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 towards 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 (rs $< .10$), and so these cases were removed from analyses, leaving a final N=4,140.

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 categories: Health, Social Science, Arts & Humanities, Science, Math, Engineering, and Business (see Table 2 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 towards 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 post-baccalaureate 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 no 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.

Academic performance. We used institutional warehouse data to obtain students SAT scores (Math, Verbal, and Writing) and grades as indicators of academic performance. For grades, we separately calculated GPAs in the first two 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 analysis section for how these variables were transformed into relative performance ratios.

Table 2. 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.*

2.3.3 Analyses

Data were analyzed using chi-square 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 3 for full details). However, like most incoming undergraduate cohorts in the U.S. in recent decades (Institute of Education Sciences, 2018), 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, in order 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 sub-categories (including separately those students who went on to take the MCAT after earning a Science or non-Science degree, as described in the Methods section 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, chi-square tests were used to determine if there were significant differences (α=.05) in the proportion of men and women who continued on to earn each degree from each intended major.

Table 3. Frequency table of students intended undergraduate major categories and degrees earned, by gender.

Intended Major	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			
	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*
Med	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*	*1456*	*1953*

Finally, we performed multiple mediation analyses to begin to understand potential explanatory mechanisms for any gender differences found in academic plan and eventual degree earned. 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 to math performance, as operationalized by relative SAT scores. However, since there were only differences in math and writing but not verbal SAT (see Table 4), we compare ratios of writing to math SAT scores since they could feasibly have an impact 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 conducted follow-up Sorbel-Goodman tests to determine the strength and significance of any potential mediation effects and the significance of each direct and indirect pathway, as well as reported the proportion of the total effect that was mediated.

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 early attrition results in the main text, and have included the later attrition function as Appendix A, Figure 13.

Table 4. 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.

Rows are organized starting with relative strengths for women and ending with relative strengths for men.

Statistically significant differences are in bold.

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

Note. * p <.05; **p <.01; ***p <.001

2.4 Results

2.4.1 Gendered Patterns Towards 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 non-medical Science degree earners than those initially intending Science (45% vs. 41% 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 US) initially intending Medicine or Health.

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 towards science (1.2:1), towards medicine (1.1:1), and especially towards health (1.6:1) and non-Science (1.8:1, see arrow widths in Figure 3).

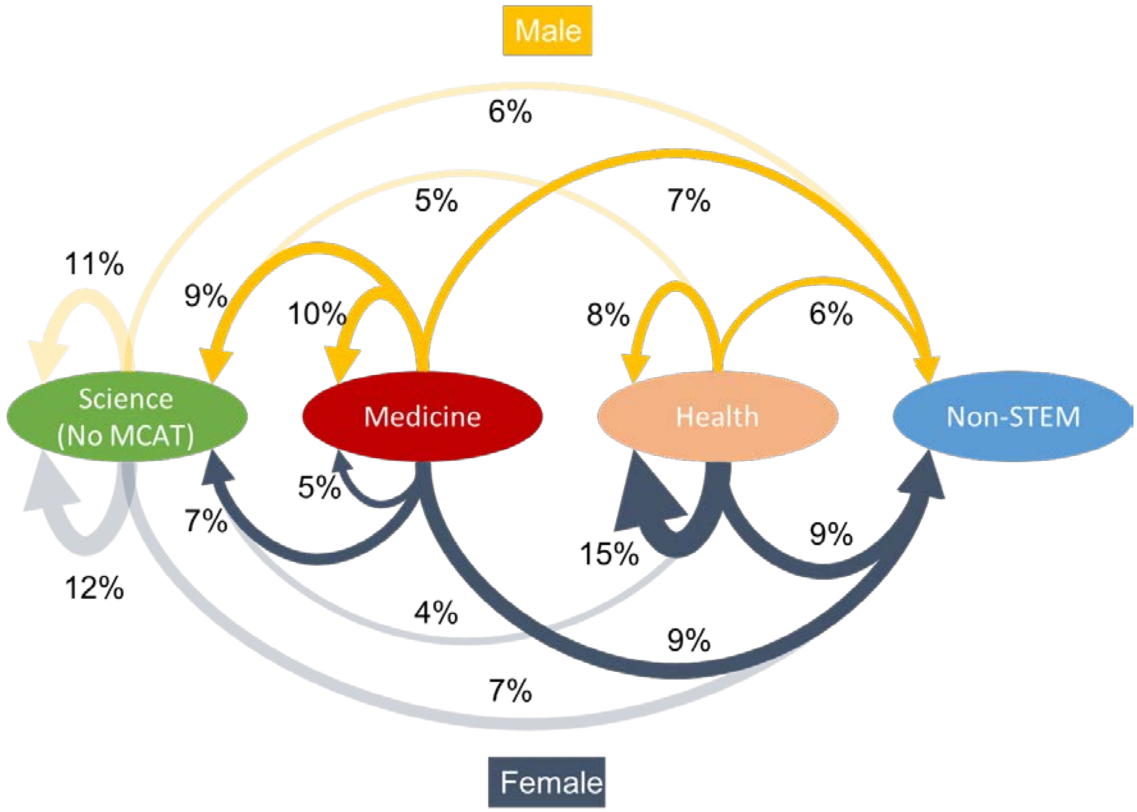


Figure 3. Diagram of the largest pathways into and out of various degree categories. Percentages represent how many students followed each pathway for each gender's total sample, with faded arrows representing no significant gender differences in percentages. Arrow widths are proportional to the raw number of students within each pathway. For clarity, only pathways that represent >= 5% of the data are shown.

When focused on yield relative percentages by gender, a less rosy picture emerges: there was a significant gender difference in the percentages of women and men entering who earn a Science degree at this general level (27% vs. 30%, $\chi^2(1)=4.55$, $p<.05$). Importantly, the differential rate of producing science degrees predominantly came from two non-Science plans: Health and non-Science (see top of Figure 4). The percentage earning Science degrees who entered with an intent to major in Science was not significantly different by gender (46% vs. 43%, $\chi^2(1)=0.49$, $p=.48$). By contrast, a large portion of the overall gender disparity found in earned Science degrees came from the significantly lower percentages of women compared to men who entered with a Health plan and ended up with a Science degree (14% vs. 26%, $\chi^2(1)=19.60$, $p<.001$). This effect was further compounded by the fact that health was the largest academic plan for women (n=709, 31% of all women).

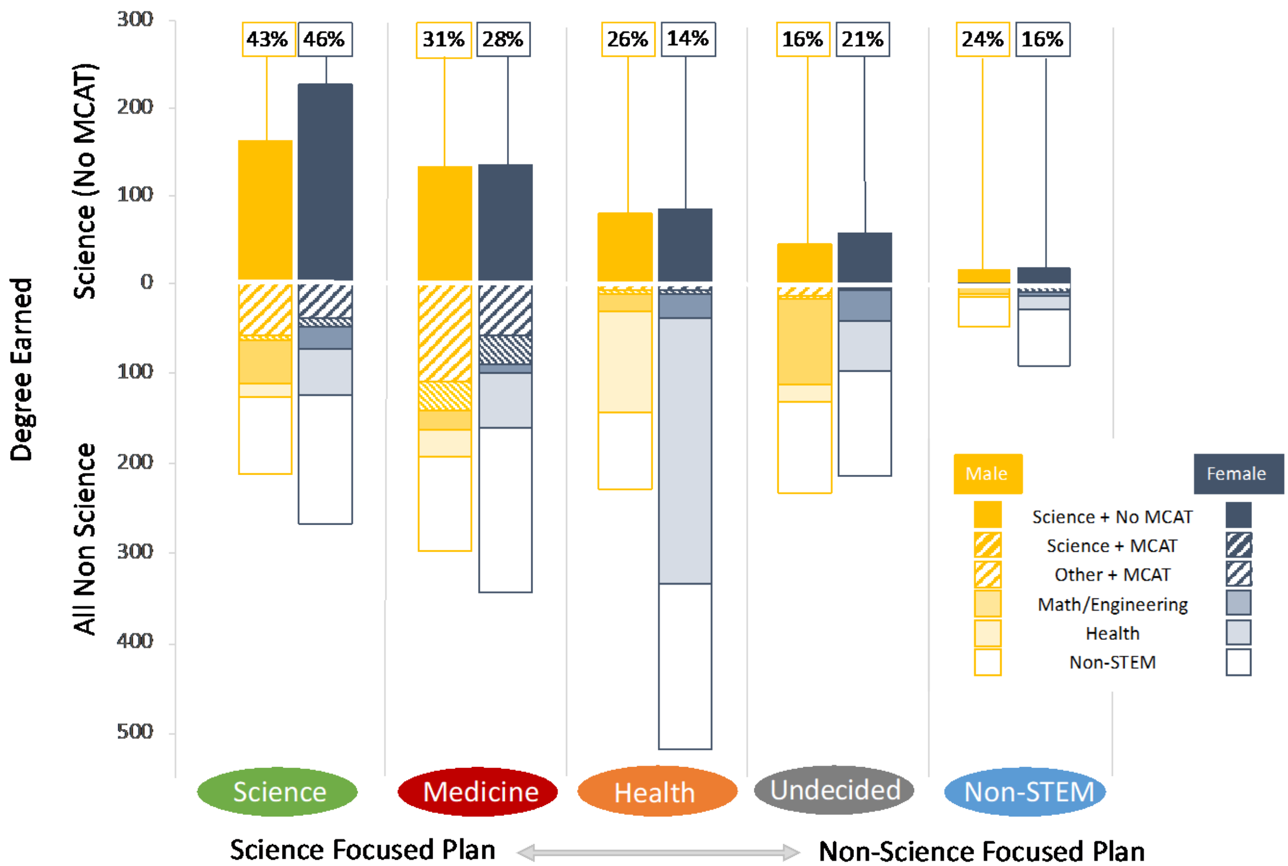


Figure 4. Descriptive statistics of the number of Science degrees earned by academic plan category and gender.

The relative percentage yield to science by gender shown in boxes (top), and the number of disaggregated non-Science degrees earned by academic plan category and gender (bottom).

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 (Cheryan et al., 2016). 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 to students intending Science (86% vs. 79%, $\chi^2(1)=7.21$, $p<.01$). By gender, more women than men earned a degree in the Life Sciences compared to the Physical Sciences (86% vs. 77%, $\chi^2(1)=13.94$, $p<.001$). However, from students intending Medicine and Health, there were no significant differences in the yield to Life Sciences between women and men (88% vs. 84%, $\chi^2(1)=1.18$, $p=.278$). 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% vs. 70%, $\chi^2(1)=13.13$, $p<.001$; see Figure 5), 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).

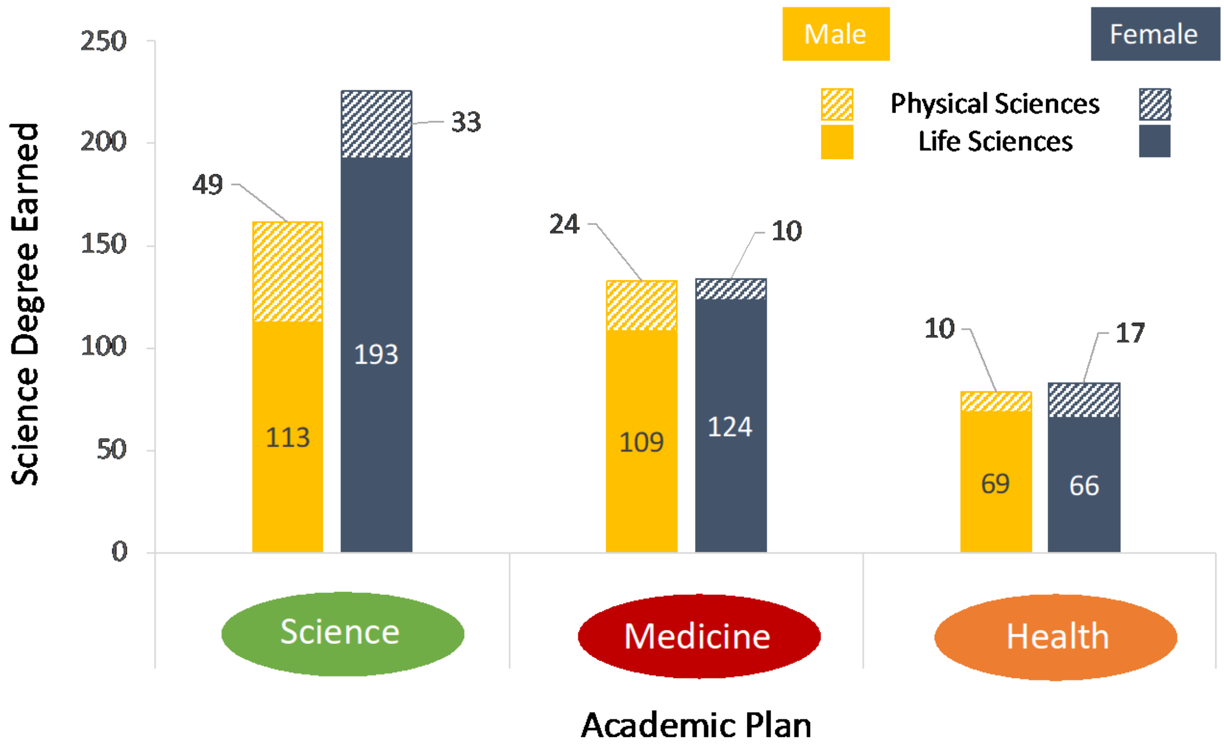


Figure 5. Number of students earning degrees in the Physical and Life Sciences, from Science, Medicine and Health academic plans, by gender.

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 4). 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% vs. 28%, $\chi^2(1)=19.24$, $p<.001$). 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% vs. 24%, $\chi^2(1)=20.30$, $p<.001$) as well as those intending to pursue Science (29% vs. 23%, $\chi^2(1)=4.16$, $p<.05$). 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.

2.4.2 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: 1) Medical academic plans leading to more women than men graduating with entirely non-STEM degrees and 2) 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 6a). 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 (β = .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.

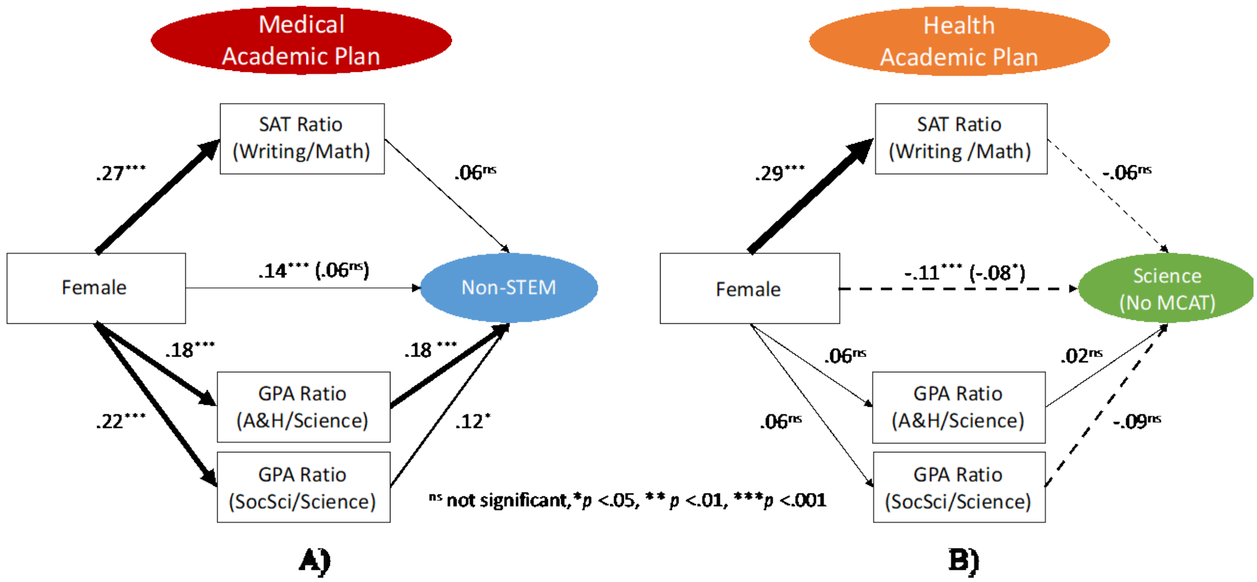


Figure 6. Mediation analyses of the effect of relative SAT scores and relative GPA on gender differences in students

(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 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.)

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 subplan 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 6b). Although the combination of these mediators still contributed a statistically significant indirect effect (β = -.02, p<.05), these factors only mediated a small proportion (20%) of the total effect, and the direct path between women and Science degrees remained significant. 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 the discussion below.

2.5 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 multi-pathways 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 et al., 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; M. Te 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 (1986; 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 (M. Te 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 towards 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.

2.5.1 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-med 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, 2017a).

Second, our current data was 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 social cognitive career theory 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 non-academic 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 towards 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 US, 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 US has one of the lowest proportions of women in medicine among OECD countries (Organization for Economic Co-operation and Development, 2018).

2.6 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 utilize motivational surveys to directly measure non-academic factors that could provide additional explanatory power for these functions. Importantly, identifying these mechanisms can help to target interventions specifically towards these large science-related groups of students which could continue to help remove barriers to gender equity in STEM.

3.0 Study 2: When Making the Grade Isn't Enough: The Gendered Nature of Pre-Med Science Course Attrition

(Manuscript published in Education Researcher; see Witherspoon, Vincent-Ruz, & Schunn, 2019. The final definitive version is available at http://er.aera.net)

Women take qualifying exams and enter medical school at substantially lower levels than predicted by their interest in medical degrees at the end of high school. We examined how science course experiences contribute to gendered attrition in pre-med, using a multi-cohort dataset of *N*=8,190 undergraduates taking the traditional pre-med sequence of introductory science courses at a public research university between 2008 and 2016. Gendered attrition was not based in academic performance, was specific to high-performing women, and yet was grounded in competency beliefs. The result is that high-performing women often graduate with lower-paying, lower-status degrees. Motivational interventions in pre-med science courses will be critical for retaining high-performing women in pre-med, an important outcome with implications for equity and women's health.

3.1 Introduction

Most areas of science, technology, engineering and mathematics (STEM) are dominated by men, significantly contributing to overall gender inequality in pay and positions of influence (Beede et al., 2011; Oh & Lewis, 2011). Medicine, another domain of high pay and status, should be a countervailing force—by the end of high school, girls are much more likely than boys to express interest in medical careers (Sadler et al., 2012), have higher high school grades (Duckworth & Seligman, 2006; Perkins, Kleiner, Roey, & Brown, 2004), and are more likely go to college (Lopez & Gonzalez-Barrera, 2014; Ryan & Bauman, 2016). Other fields where similar asymmetrical interests exist (i.e. teaching, social work) are now female dominant, and 30 of the 34 OECD countries have higher rates of women physicians than men. However, the U.S. Bureau of Labor Statistics (2017b) reports that only about 38% of physicians and surgeons are women, placing the US as 31st out of 34 OECD countries in terms of percentage of women physicians (OECD, 2018).

While gender differences in medical training have also been studied internationally (Kvaerner, Aasland, & Botten, 1999; Riska, 2011), gendered attrition in the pursuit of medical careers may be particularly problematic in countries like the United States, which place medicine as a post-graduate degree. Indeed, national data show that women are not persisting through undergraduate pre-med pathways (Fiorentine & Cole, 1992); women’s overrepresentation in pathways to medicine in high school largely disappears by the time students finish college and apply to medical school. However, despite these recent reports of growing *equality* by gender in medical school matriculation over the past decade (Association of American Medical Colleges, 2017a), this parity may hide that the large attrition from core pre-medical science courses during the undergraduate years is not *equitable* by gender; if early trends in medical interest remained constant, women should be overrepresented in the medical profession, as they are in many other OECD countries. Further, given the documented barriers for women occurring at later stages of medical training and careers, and particularly from administrative positions and male-dominated specialties where women's health issues are often relatively under-studied (Bates et al., 2016; Bickel, 2005; Johnson, Fitzgerald, Salganicoff, Wood, & Goldstein, 2014; Kvaerner et al., 1999), promoting equity in these areas is likely to require encouraging high-performing women to persist at levels above 1:1 gender parity, through and beyond the point of medical school matriculation.

Explanations are unlikely to lie in innate biological or cognitive differences related to competence in science; psychology research shows gender differences to be very small or exhibit relative strengths in both directions (Hyde, 2005; Kilminster, Downes, Gough, Murdoch-Eaton, & Roberts, 2007; Spelke, 2005). Instead, motivational mechanisms or relative academic performance may provide alternative hypotheses. There is a wealth of evidence to suggest that students' academic decision making, including course-taking behaviors, are influenced by expectations of success and valuation of academic pursuits, and that these factors are related to both students' perceptions of self and abilities, as well as affective responses to prior academic experiences (Watt, Eccles, & Durik, 2006). Studies show that as early as middle school, motivational factors like interest, identity and especially competency beliefs in science, are related to girls' participation in and learning of science content (Cromley et al., 2016; Hazari, Sonnert, Sadler, & Shanahan, 2010; Vincent-Ruz & Schunn, 2017). Problematically, women may interpret academic feedback from grades more negatively than their male peers, particularly in domains that are traditionally male-dominated, and may be more likely to rate themselves lower in perceived ability despite similar levels of achievement (Beyer & Bowden, 1997; Eccles, 1994; Kugler et al., 2017).

Eccles' Expectancy-Value Theory (EVT) provides a helpful framework through which to understand how psychological factors such as interest, identity and competency beliefs may interact to introduce gender differences in achievement-related choices (Eccles, 1994; Wigfield & Eccles, 2000). EVT proposes that educational and career decisions are directly influenced by both students' expectations of success on a task, and the subjective value of the task in terms of students' intrinsic interest or enjoyment (*interest value*), their perceived utility of the task (*utility value*), and the relationship between the task and the students' sense of self (*attainment value*; Eccles & Wang, 2016). In course enrollment decisions, for example, students might consider the subjective value of a course based on how much they will enjoy it (i.e., interest; Semsar, Knight, Birol, & Smith, 2011) and whether or not it aligns with their felt sense of self (i.e., identity; Gee, 2001), as well as whether or not they expect to be successful in the course based on beliefs about their abilities in the subject (i.e., competency beliefs; Bauer, 2005). Locating and specifying gender differences in these three indicators of student valuation and expectation of success within specific courses along the pre-med pathway could provide a first step to understanding the features of particular courses that influence women's decision to leave medical careers. For example, are women's decisions to leave certain required course sequences primarily a result of declining interest, perceived incompatibility between the course and their identity, a lack of belief in their ability, or some combination of these factors?

The predicted source of the effect is unclear. On the one hand, there is evidence to suggest that large differences in competency beliefs or interest would not be expected within the high-performing populations of women typically found on a pre-med track (Eccles, 1994). Instead, some have proposed an alternate hypothesis; that some attrition for high-performing women is a function of an increased number of viable and more desirable alternatives available (M. Te Wang et al., 2013). For example, girls perform at slightly higher levels than boys in non-STEM subjects in high-school; this relative academic *advantage* may increase the variety of non-STEM career options for women, which lowers their relative likelihood of pursuing science careers. On the other hand, students' academic self-concept has been shown to be constructed through both external and internal comparisons—that is, perceptions of one's ability may be a function of both a comparison to others' ability in that subject, as well as to one's own relative ability in other subjects (Marsh, 1986). Therefore, high academic achievement in other content areas compared to medicine-related courses, along with a false perception of higher achievement by their male peers, may result in lower competency beliefs even for high performing women.

Gaining a better understanding and addressing the underlying causes of gendered attrition from medical careers will likely require focusing on the science coursework during the years between high school and college graduation (Cromley et al., 2016; Kugler et al., 2017; Morgan et al., 2013). Undergraduate pre-med typically involves four challenging two-course science sequences (Introductory Biology 1 and 2, General Chemistry 1 and 2, Organic Chemistry 1 and 2, and Introductory Physics 1 and 2). Prior studies of pre-med attrition indicate that students perceive Chemistry, Biology, and Physics courses as highly indicative of medical career success; underperformance in those courses may contribute to declining interest in pre-med (Barr, 2010). While prior research has shown variation in gendered attrition broadly within these domains, with some (i.e., Biology, Chemistry) showing relative advantages for women, and other domains (i.e., Physics) showing relative disadvantages for women (Cheryan et al., 2016), little is known about which *specific* pre-med course sequences show the largest gender differences in attrition, and, most importantly, what factors contribute to women leaving these courses. Therefore, our approach uses longitudinal analyses to answer the following research questions: 1) Where do gender differences in attrition appear in the pre-med science sequence, 2) What motivational factors may explain large gender gaps that may appear and, 3) Do these the cumulative effect of attrition over the entire sequence result in gender differences in pre-med persistence?

3.1.1 The Current Study

In this study, we examined institutional data records from 8,190 undergraduate students at a large undergraduate research university, enrolled in core pre-med course sequences within their first two semesters between 2008 and 2016; multiple cohorts across many sections ensures patterns that are not specific to a small number of instructors. Multiple regression analyses were used to observe whether or not women were less likely than men to enroll in the *second* course of an undergraduate pre-med science course sequence, even when successful in the first course. Leaving mid-way through a sequence is a strong and time-specific indicator of attrition, with recent experiences (e.g., first-course performance) offering potential explanations. By contrast, modeling factors that influence the decision to *start* sequences are complicated by the optional order of some courses, making attrition decisions at that time-point somewhat ambiguous. Our primary analyses examine both whether women were more likely than men to drop from these courses, and whether some sequences (either by content domain or timing) showed greater differences in within-sequence attrition by gender.

Finally, we also use graduation records to determine the proportion of degrees eventually earned by students who entered and completed these sequences of pre-med courses. These analyses address the critical outcomes question: if women don’t persist in pre-med pathways, what degrees do they end up pursuing instead?

3.2 Methods

3.2.1 Sample

This retrospective multi-cohort study consisted of *N*=8,253 undergraduate students at a large urban research university in the Northeastern United States (henceforth, "the University"). The University is broadly representative of similar institutions with a relatively selective admission 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 (*Mdn* = $111,000).

The sample included for analyses those students enrolled in General Chemistry 1 within the first two semesters. Importantly, in the sample used here, men were not more likely than women to pass these courses (i.e., receive an A, B, or C); small differences instead favor women in Introductory Biology (94% vs. 92%, $p<.01$) and General Chemistry (96% vs. 94%, $p<.05$). Therefore, differential failure rates would not account for women's higher levels of attrition in the overall pathway. We therefore were interested in whether observed gender differences in attrition for this group of students could be explained by relative academic strengths and weaknesses in STEM a non-STEM disciplines, or by motivational factors such as competency beliefs, science identify and science fascination.

An eight-year window was used to insure generalizability across student cohorts and instructors. The racial and ethnic diversity of our sample roughly mirrored that of the University as a whole; students were predominantly White (71%), with Asian (15%) and Black or Hispanic (9%) students making up the next two largest ethnic groups. The primary predictor variable, gender, was coded as 1 if the student self-identified as female (57%) and 0 if the student self-identified as male (43%). In-course surveys established that, overall, 63% of these students planned on going to medical school. Of those intending to go to medical school, 65% were women. All University data was provided for analysis with Institutional Review Board approval.

3.2.2 Measures

Course variables. Primary outcome variables included four binary measures of enrollment (1=enrolled, 0=not enrolled) in each of the second courses of the four pairs of courses of the core pre-med sequence: Introductory Biology 2, General Chemistry 2, Organic Chemistry 2, and Introductory Physics 2. To analyze how performance in the first course of a sequence was related to students' persistence to the next course, only students also enrolled in the prior course at the University were in analyses of each course sequence. It is important to note here while not mandatory, this series of pre-med science courses represents a progression that is common to pre-med tracks across multiple institutions (Association of American Medical Colleges, 2017b), was highly recommended by pre-health advisors at the University, and was the most commonly observed progression in our data. Therefore, while it is possible that a student, for example, entered Organic Chemistry 1 and Physics 1 without entering Organic Chemistry 2, this occurred in less than 4% of our dataset, and thus the experiences with later sequence courses would rarely influence earlier sequence courses. Further, any effects from gendered selection at earlier points would logically lead to a smaller effect in later courses, as women with a propensity to leave pre-med would not be present in later course sequences. However, the size of the effects in Organic Chemistry and Physics were comparable, arguing against selective attrition as the source of the observed temporal pattern, regardless of the order of these courses.

A binary measure of students sitting the MCAT exam was also included. These analyses focused on students who completed the full combination of all four core sequences, and an additional elective course (either Biochemistry or Chemical Biology). These “Pre-Med Courses” are typically taken by pre-med students, make up the content of the MCAT exam, and were not a required combination to earn any other degree. Enrollment in these courses was highly predictive of taking the MCAT; odds of taking the MCAT in students completing this combination were about 8.6 times higher than those not completing them, OR=8.62, 95% CI [7.10, 10.46], z=21.76, p<.001.

Academic covariates. Academic variables consisted of students’ highest SAT Math, Verbal and Writing scores, Advanced Placement credits and scores, and cumulative high-school GPA. Two ratio variables were calculated to represent strengths in courses outside of the pre-med track, relative to their pre-med science courses: “Social Sciences” and “Arts and Humanities”. Social Science ratios were calculated as the mean GPA of all courses within Anthropology, Psychology, Sociology, Economics and Political Science, divided by the mean GPA of all pre-med science courses. Arts and Humanities courses were calculated as the mean GPA of all courses taken within English, History, African Studies, Arts, Music, Theater and various Languages, divided by the mean GPA of all pre-med science courses. For attrition analyses, these discipline-specific ratios were calculated using only courses prior to the point of pre-med attrition (see Table 5).

Motivational covariates. Motivational data was collected from a subset of N=520 students during the first three weeks of Organic Chemistry 1 (i.e., prior to the first summative assessment) for in-depth analysis of the largest gender effect, and consisted of Chemistry Fascination (e.g., “I want to know everything I can about chemistry”), chemistry Competency Beliefs (e.g., “I can usually figure out a way to solve chemistry problems”), and Science Identity (e.g., “I think of myself as a ‘science person’), adapted from the *Colorado Learning Science Survey for Use in Chemistry* (CLASS-Chem) (Semsar et al., 2011), the *Chemistry Self-Concept Inventory* (CSCI) (Bauer, 2005), and a science identity survey (Hazari et al., 2010). Items were rated on a 4-point Likert scale (Strongly Agree – Agree – Disagree – Strongly Disagree) and calculated as a mean score (see Table 5). Reliability for all items ranged from moderate to good. In terms of discriminability, the highest correlation was between Fascination and Competency Beliefs, showing a moderate correlation of r=.55, consistent with previous reports in the literature (Bauer, 2005; Vincent-Ruz & Schunn, 2017; see Appendix B, Table 9). A single binary indicator of intent to pursue a medical career (“Yes” = 1, “No” = 0) was also collected in General Chemistry 1 and Organic Chemistry 1.

Degrees earned. Bachelor’s degrees earned by students in the sample were gathered from University historical data and coded across N=498 unique degree combinations into seven general degree categories: Health, Social Science, Arts & Humanities, Science, Math, Engineering, and Business (see Appendix B, Table 10 for a detailed coding scheme).

Table 5. Descriptive statistics of all STEM, non-STEM, and motivational covariates by gender.

	Female					Male				
	N	M	SD	Min	Max	N	M	SD	Min	Max
STEM AP Score[a]	4,686	0.28	0.01	0	1	3,567	0.33	0.01	0	1
STEM AP Credit	4,686	1.1	1.4	0	9	3,567	1.3	1.6	0	10
SAT Math	4,411	642	69	410	800	3,399	665	67	440	800
Non-STEM AP Score[a]	4,686	0.3	0.01	0	1	3,567	0.28	0.01	0	1
Non-STEM AP Credit	4,686	1.3	1.6	0	11	3,567	1.1	1.6	0	9
SAT Verbal	4,411	639	73	390	800	3,399	633	72	400	800
SAT Writing	4,643	623	71	380	800	3,524	605	74	390	800
HS GPA	4,676	4.0	0.5	0	9	3,562	3.9	0.5	0	7
Arts & Hum. Ratio	4,494	1.5	0.6	0	9	3,399	1.4	0.6	0	10
Social Sci. Ratio	4,367	1.4	0.5	0	8	3,253	1.3	0.5	0	9
Fascination	322	2.9	0.6	1	4	197	3.1	0.6	2	4
Competency Beliefs	321	2.8	0.5	1	4	197	3.0	0.5	2	4
Science Identity	320	3.6	0.4	2	4	197	3.4	0.5	2	4
Med Career Interest	322	0.59	0.03	0	1	198	0.55	0.03	0	1

[a] AP Score defined as proportion of STEM AP Exams earning > 3 out of 5, a common threshold for acceptance for university course equivalence.

3.2.3 Procedure

For sequence completion and MCAT enrollment analyses, a multi-cohort longitudinal dataset was analyzed using a series of multivariate logistic regressions in *Stata 15*. Large, multi-cohort datasets enable discovery of generalizable patterns unlikely to be specific to particular instructional styles, course structure, teaching assistants, or cohorts of students. However, cohorts may also differ and produce confounds in the analyses; to control for general cohort effects, we also included a model using each students' starting academic term as a single continuous cohort covariate in our analyses. Each regression model was built showing direct, uncontrolled effects of gender on each outcome variable ("Uncontrolled"). Next, two groups of covariates were added one-by-one to determine if different academic factors influenced gender differences: relative STEM academic strengths ("STEM") and non-STEM relative academic strengths ("Non-STEM").

For each model, covariates were only included as potential mediators of the gender effect if they were significantly correlated with gender. An alpha level of $\alpha = .01$ was used for all exploratory analyses of the large dataset.

To understand potential motivational mechanisms, we collected additional data via online surveys from multiple sections of Organic Chemistry, the sequence with the largest gendered attrition. Using mediation analysis, we tested whether the relationship between gender and enrolling in the second course was mediated by each attitudinal factor (i.e., fascination, science identity, competency beliefs), or all three. A generalized linear estimator was implemented using the *lavaan* package in *R* (Rosseel, 2012), to more accurately model binary outcomes. An alpha level of $\alpha = .05$ was used for this focused analysis of the smaller survey dataset.

Finally, because course attrition may be related to career pathways through the type of degree earned, logistic regressions were performed on each degree category using gender as a predictor. In addition, because students enter medical school from many undergraduate degrees, but some may be more common pathways, percentages of students who took the MCAT within each degree category by gender were also examined using logistic regression, with taking the MCAT as the outcome and degree type earned as the predictor.

3.3 Results

Overall, attrition by prior course grade showed the expected trend; both male and female students with higher grades were far more likely overall to continue to the next course, with fewer than .01% of students earning a D or F continuing on to take the second course of any sequence, or the MCAT (see Appendix B, Table 11). Further, across all course sequences, there were no significant differences in attrition by gender for students in these lowest grade ranges; therefore, the analyses focus on students earning a C or higher. No significant differences were found between female and male attrition from course sequences typically taken in the first year, Intro Biology (89% vs. 91%, p=.14; see Figure 7A) and General Chemistry (88% vs. 88%, p=.91; see Figure 7B). However, in course sequences typically taken in the second year (Organic Chemistry) and third year (Physics), significant differences were found between males and females. But this gendered attrition difference was only found in students who received an A or B in the prior course, with women having significantly lower odds than men of continuing to Organic Chemistry 2 (89% vs. 96%, p<.001), and of continuing to Physics 2 (82% vs. 88%, p<.001). That is, women receiving an A or B in advanced courses on the pre-med track were approximately 2.9 times less likely than similarly performing men to continue to the next Organic Chemistry course (see Figure 7C), and about 1.7 times less likely to continue to the next Physics course (see Figure 7D).

There was also large gendered attrition of high-performers taking the MCAT, after having completed all the sequences; the odds of taking the MCAT after having received an A or B in the full set of Pre-Med Courses were about 2 times lower for females than males (32% vs. 47%, p<.001; see Figure 7E). Overall, the gender proportion shifts from almost 2:1 female-to-male students intending med-school at the beginning of their first semester at university, to 3:4 female-to-male taking the MCAT (see Figure 7F). That is, differential losses by gender throughout the pre-med sequence cumulatively produce a large gender effect overall.

Models testing for cohort effects showed a small overall negative linear trend (i.e., overall lower retention for later cohorts), with significant cohort effects found only in General Chemistry, Physics and for taking the MCAT; however, these effects were not strong enough to meaningfully change the estimates for gender differences in these courses (i.e., adjusted effects were not outside the 95% CI of the uncontrolled model) and so we proceeded without this covariate in our subsequent modeling (see Appendix B, Table 12 for model details).

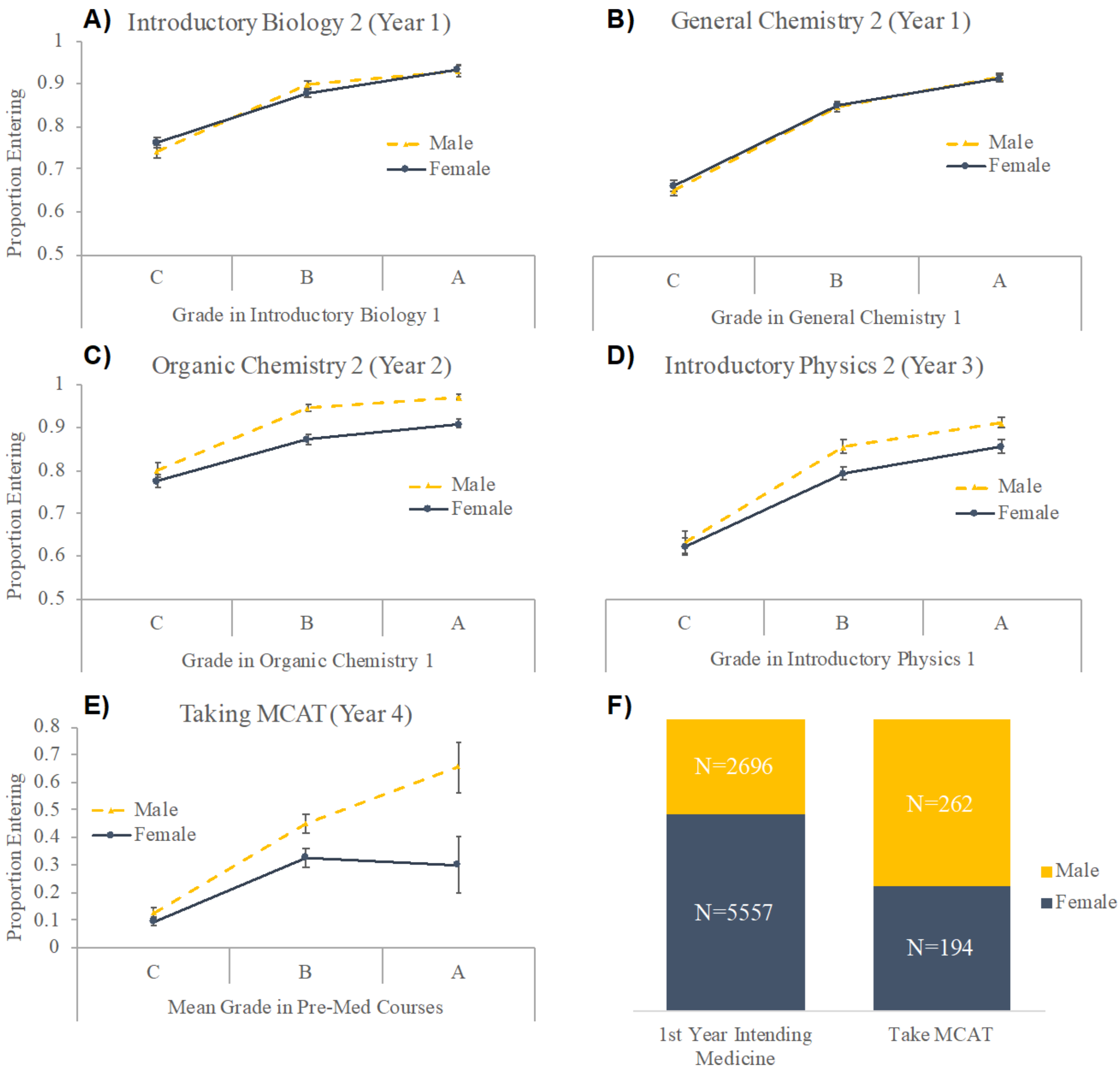


Figure 7. Gendered attrition on the path to medical school.

Proportion of passing students entering each second course in the four core science sequences (A-D) and taking the MCAT (E), by gender and grade in prior course(s), with. (F) shows estimated numbers of entering students intending medical school and taking the MCAT.

To better understand which academic preparation and performance variables might explain persistence differences, we first tested which academic covariates generally predicted persistence in Organic Chemistry, the course sequence with the largest gender gap. All tested covariates were correlated with overall persistence as expected (see Appendix B, Table 13 for the full correlation table). Because strong correlations were shown amongst some of our covariates, we examined Variance Inflation Factors (VIFs) to check for multicollinearity in those predictors; all VIFs were shown to be below 2.5, a conservative threshold for multicollinearity with such a large sample size (O'Brien, 2007; see Appendix B, Table 14). We also tested for gender differences in the extent to which each academic variable predicted persistence, based upon prior research showing differential reactions to negative grade feedback (Beyer & Bowden, 1997)—that is, do some academic variables matter more for women than men?

Analyses showed lower odds for women than men (1:1.3) of continuing to Organic Chemistry 2 in the middle two quartiles of HS GPA range (42% vs. 49%, $p<.001$). However, women in the highest two quartiles of SAT Math had slightly higher odds than men (1.2:1) of continuing to Organic Chemistry (61% vs. 55%, $p< .01$). Across both genders, students with a higher relative GPA in either Social Sciences or Arts and Humanities have lower odds of continuing to Organic Chemistry 2. While there were no gender differences for students with the highest and lowest range of these ratios, women in the 3rd quartile of the Social Sciences ratio showed lower odds (1:1.7) of continuing than men (79% vs. 86%, $p< .01$), and women in the 2nd quartile of the Arts and Humanities ratio had lower odds (1:2.2) of continuing than men (86% vs. 93%, $p< .001$). Therefore, some academic variables favor men persisting in pre-med, but others favor women's persistence (see Appendix B, Figure 14 for a summary).

To then test the extent to which these relative academic strengths and weaknesses explained (i.e., mediated) the differential attrition by gender across all course sequences, we included academic covariates (i.e., AP scores, SAT scores, HS GPA, recent course performance) representing both STEM and non-STEM strengths in each sequence-continuation regression model (see Appendix B, Table 15 for model details). While overall including STEM variables slightly decreased women's relative attrition, and including non-STEM variables slightly increased relative attrition as suggested in the literature (M. Te Wang et al., 2013); these effects did not differ significantly from the uncontrolled model (see Figure 8).

Resulting models still showed lower odds of women continuing to Organic Chemistry 2 when including relative strengths in both STEM (91% vs. 96%, $p < .001$), and non-STEM, (90% vs. 96%, $p < .001$). Similarly, lower odds for women continuing to Physics 2 remained when including relative strengths in STEM (84% vs. 90%, $p < .001$), and non-STEM, (82% vs. 89%, $p < .001$). Finally, women had lower odds relative to men of taking the MCAT exam (including only students who had completed all Pre-Med Courses), both when controlling for relative strengths in STEM, (35% vs. 49%, $p < .001$), and non-STEM, (33% vs. 48%, $p < .001$). In sum, differential attrition effects by gender were not explained by relative academic strengths and weaknesses.

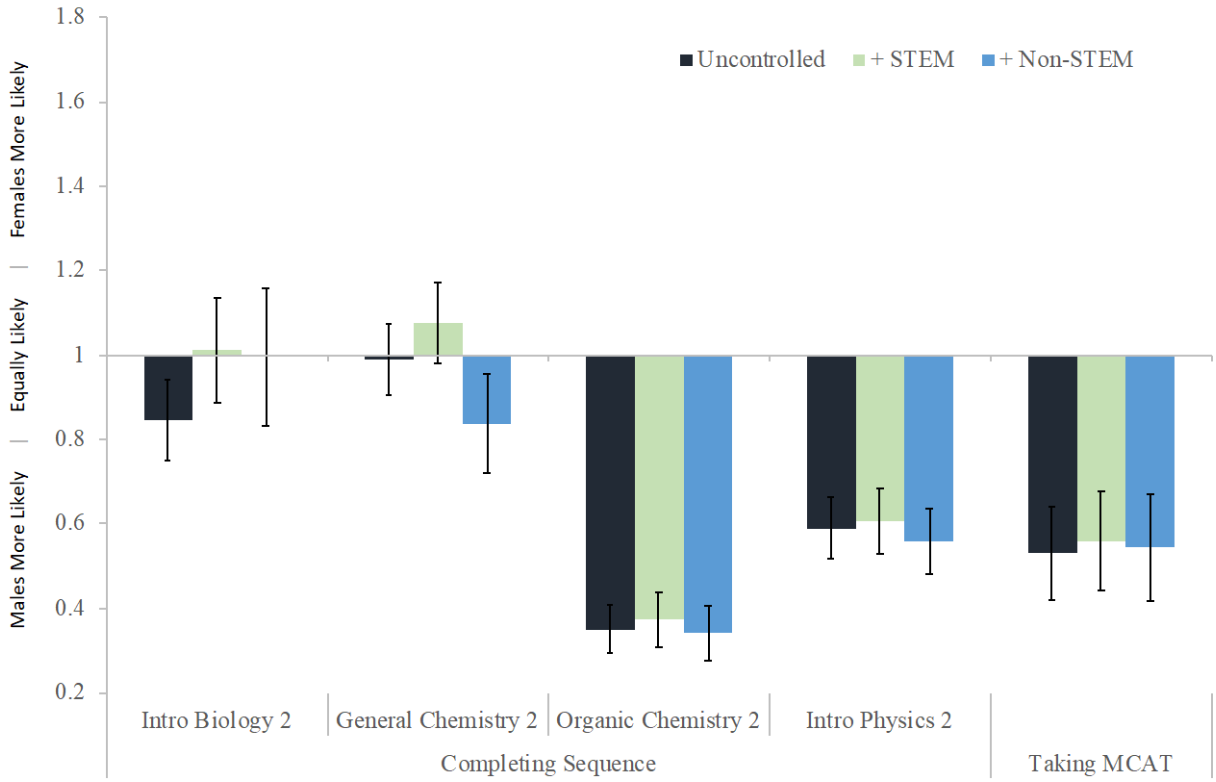


Figure 8. Changes in relative size of gender effect, with explanatory variables.

Odds ratio of women compared to men for entering a second course in sequence after receiving an A or B in the prior course, overall and after controlling for relative strength in STEM and non-STEM academics, with standard error bars shown.

The next set of analyses tested a mediation hypothesis using a subset of A and B students (N=335) surveyed in several course offerings of Organic Chemistry 1. These models included attitudinal survey variables (chemistry fascination, Chemistry Competency Beliefs and Science Identity) as possible mediators between gender and enrolling in Organic Chemistry 2, as well as a binary indicator of intent to pursue a medical career, for students receiving A and B grades in the prior course. Medical career intent was not significantly correlated with either gender (p=.72) or retention to Organic Chemistry 2 (p=.58). Overall, this sample also showed significantly lower odds of women continuing to Organic Chemistry 2 (74% vs. 84%, p<.05). Interestingly, mediation of this gendered attrition through Chemistry Fascination was not significant, while mediation through Science Identity was a smaller effect and in the wrong direction (i.e., predicted greater female enrollment in Organic 2); therefore, these paths were trimmed from the final model.

Mediation analyses revealed that Chemistry Competency Beliefs, which showed the largest gender difference and the strongest connection to Organic Chemistry 2 enrollment, was the primary mediator (see Figure 9). When only Chemistry Competency Beliefs are included in the mediation model (see Figure 9, coefficients in parentheses), the initial direct relationship between gender and enrollment is no longer significant, (76% vs. 84%, p=.10). Further, there is a significant negative correlation between gender and Chemistry Competency Beliefs (p<.001), meaning that women are more likely to respond with lower ratings of their beliefs in their chemistry ability. Also, there is a positive correlation between Chemistry Competency Beliefs and enrollment in Organic Chemistry 2 (p<.001), meaning that higher ratings of Chemistry Competency Beliefs are correlated with a higher likelihood of continuing to Organic Chemistry 2. This significant indirect pathway suggests that the initially observed direct effect of gender on enrollment in Organic

Chemistry 2 is at least partially explained by women's lower competency beliefs in chemistry (Baron & Kenny, 1986).

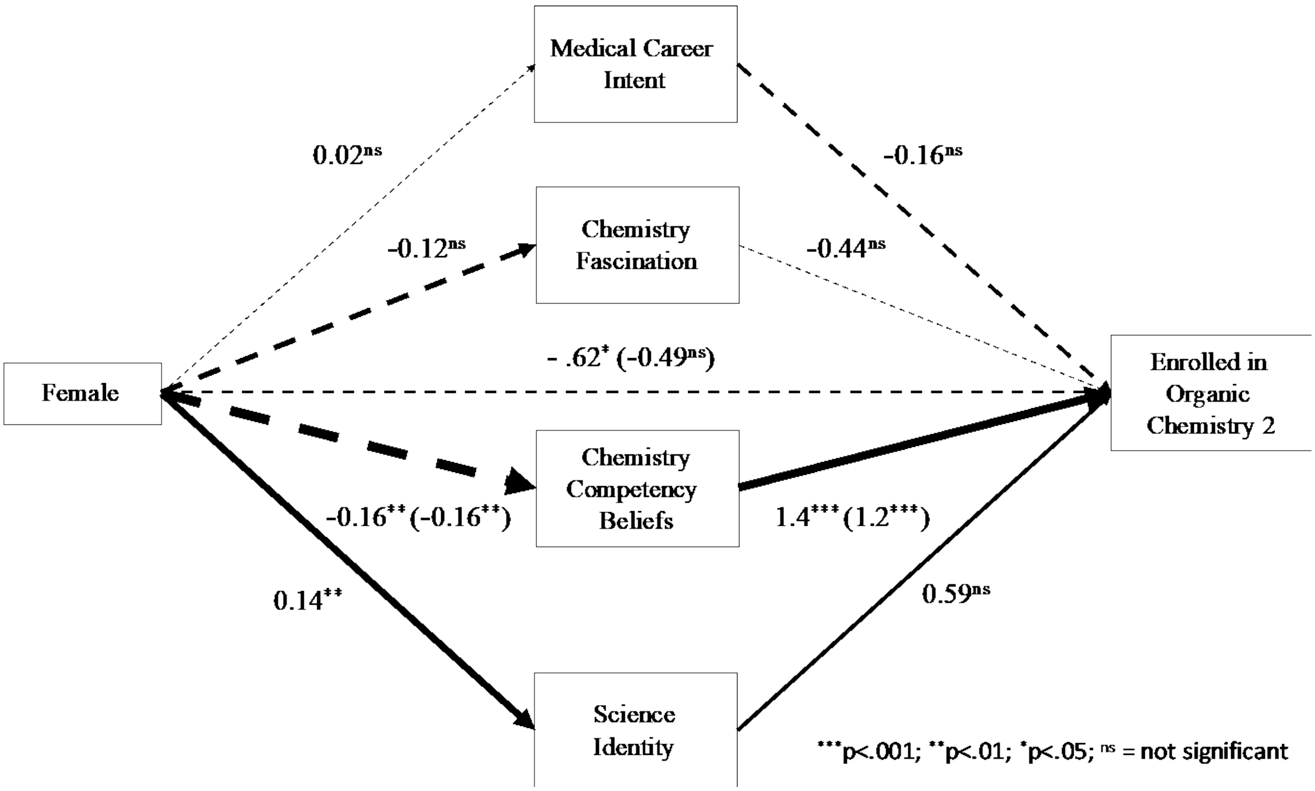
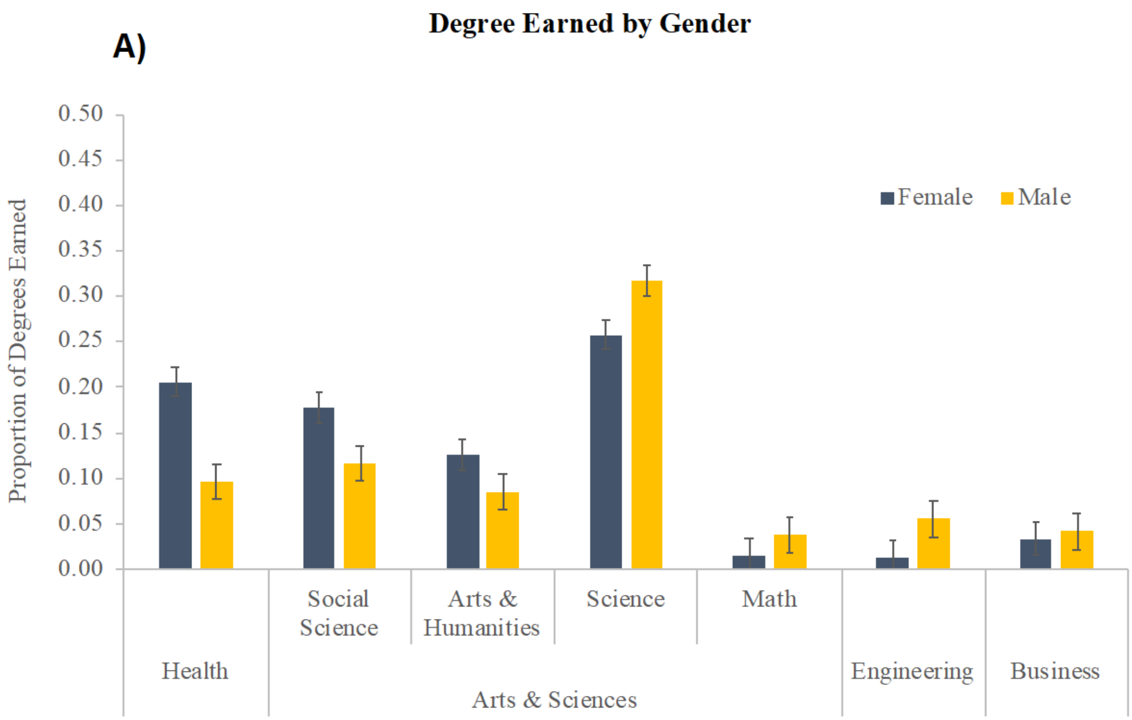


Figure 9. Mediating role of competency beliefs on gendered attrition in organic chemistry. Mediation analysis with logit regression coefficients shown for students earning an A or B in the first course, including all motivation variables (and with competency beliefs only).

Finally, additional logistic regressions were performed to determine whether there were gender differences in the type of undergraduate degree earned (indicative of career pathways directed away from medicine). Similar to broader studies of undergraduate degree earning (Morgan et al., 2013), women in our sample were more likely than men to earn degrees in Health, Social Sciences and Arts & Humanities, but less likely than men overall to earn degrees in Science, Math, Engineering or Business (see Figure 10A). Results showed that even in this large sample of students intending medical school and likely taking many science courses, women's odds were 1.3 times lower than men for earning a Science degree (26% vs. 32%, $p<.001$), and students earning Science degrees made up 77% of MCAT-takers. Instead, women's odds of earning an undergraduate Health degree were 2.4 times higher than men (21% vs. 10%, $p<.001$), one of the least likely groups to take the MCAT (see Figure 10B). Further, even women earning Science degrees were less likely to take the MCAT than men (17% vs. 28%, $p<.001$). Thus, women commonly pursued a career in broader health fields, but in lower-paying, lower status positions relative to their initial medical-school intentions. When examining subcategories of Science (i.e. degrees earned in Biology, Chemistry or Physics) our findings follow national trends, with women equally likely to earn Biology degrees as men (Cheryan et al., 2016); however, those women were still less likely than their male counterparts to continue to take the MCAT (see Appendix B, Table 16 for model details). Similar patterns held when including only students who consistently obtained an A or B across each of the course sequences, ruling out differential attrition from poor course performance (see Appendix B, Table 17 for model details).



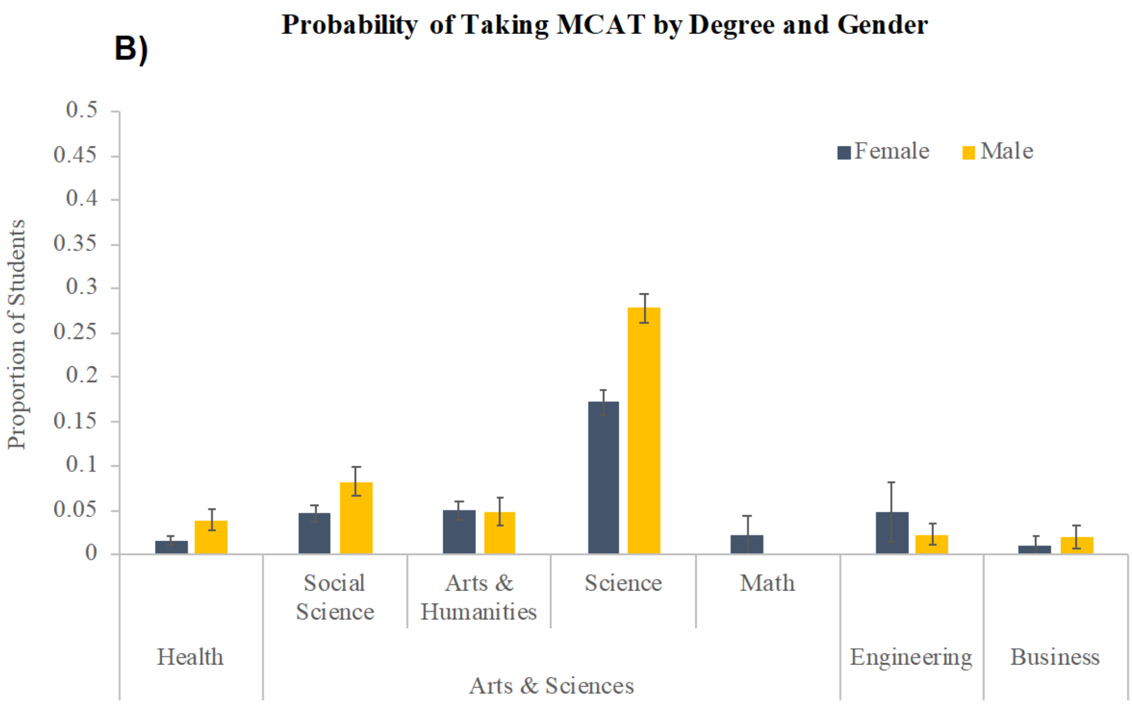


Figure 10. Proportions of A) degrees types earned by gender and B) the probability of taking the MCAT within each degree type by gender, with standard error bars shown.

3.4 Discussion

Gender equity is currently a central problem in the US, particularly in science, but also broadly in all high-status and high-pay positions. This study brings into focus a large contributor to this overall problem that has received relatively little prior attention: gender equity in pre-medicine, a pathway to one of the largest high-status, high-pay science workforce sectors in the US. Prior work often fundamentally mischaracterizes the current situation as one of parity between men and women entering medical school— yet for a number of reasons which we have presented, medicine should be heavily dominated by women. Therefore, parity in medical schools represents an outcome that is far from ideal. Our analyses provide critical insight into the clear trend of gendered attrition that can be observed through the comparison of national career interest data for students exiting high school with national entering demographics of medical schools. We find that only later pre-med courses, and the step of choosing to sit the MCATs after completing all required courses, show large gender biases in attrition. Importantly, these effects were found primarily for high performing women, and these effects were partially mediated by differences in competency beliefs. While many previous studies have highlighted the importance of competency beliefs in STEM persistence, our findings uniquely locate this effect to the group of high performing women in pre-med; the effect found here is therefore unlikely to be one of differential reaction to failure feedback from formal grades, a commonly offered explanation.

These findings have a number of implications. First, for this population we have shown that previously offered hypotheses of gendered attrition as a function of relative successes in non-STEM academic areas, or relative weaknesses in STEM, do not hold for our sample. While both men and women are influenced by relative academic performance, this effect does not disproportionately impact women. Instead the current data supports an alternative hypothesis related to motivational factors. Yet this observed mediation of competency beliefs is also not a simple replication of past research: a general effect of competency beliefs would have predicted gendered attrition that was equally large across performance levels due to pre-existing differences in competency beliefs (Vincent-Ruz et al., 2018) or gendered attrition only in the C and B range due to differential reaction to failure feedback based on grades (Kugler et al., 2017). One possible interpretation of the strong mediation effect of competency beliefs for women who earn an A or B in our data might be related to the perception of the relative effort required to earn those grades in these environments. Especially in science courses, where stereotype threats are often particularly noticeable for women, the level of effort required to achieve high grades in these courses may be highly salient and attributed to a lack of ability. Future work incorporating additional components of Expectancy Value Theory could help to identify these other important factors of women's course-taking decisions. For example, understanding students' interpretation of the various costs (e.g., psychological costs, opportunity costs) associated with continuing through the pre-med track, as well as their perception of gendered sociocultural norms associated with pre-medical study at their institution, could further inform intervention. Indeed, some research has shown evidence that all students are likely to perceive incongruity between STEM careers and family caregiving as they get older; while no gender differences in these evaluations were found for college-aged students, structural factors such as family planning may influence women's decisions to continue in a medical career at much later points (Eccles & Wang, 2016; Weisgram & Diekman, 2017).

Overall, regarding women's pre-med attrition, this study has advanced understanding of gendered attrition in this very large STEM-focused group:

- Previously-offered explanations shown not to hold for our sample:
    - Relative academic weakness in STEM

    - Relative academic strength in non-STEM
    - Differential changes in competency beliefs due to differential reaction to failure feedback based on grades
    - Prior differences in competency beliefs prior to arriving at university
- Alternate explanation supported:
    - Differential change in competency beliefs due to factors other than success

It now becomes important to examine the ways in which this phenomenon may be moderated by instructional and institution factors. For example, the particular institution studied was a moderately selective, large co-ed urban school, with a student body that was relatively homogenous by ethnicity. Looking at pre-med attrition at universities that are more or less selective, located in non-urban areas, more or less ethnically diverse, of a different size, or single-gender may reveal interesting variation in gendered attrition along the pre-med pathway. A multi-institutional approach that intentionally incorporates schools with different pre-med course sequences could also address a potential limitation of the current study, by examining if similar attrition patterns exist in Organic Chemistry and Physics in universities where this particular pre-med course combination is less common. Additionally, as this work points to the importance of motivational factors, rather than relative academic strengths and weaknesses, exploring a wider range of motivational constructs that could be related to gendered attrition might offer a deeper understanding of the underlying mechanisms that contribute to gendered attrition in pre-med, and how these interact with particular pre-med science courses. For example, we hypothesize that the effects found in Organic Chemistry are related to the pervasive perception of this course among students and faculty as a “gatekeeper course” that is predictive of success in medical careers, making gender-based stereotypes particularly salient (Barr et al., 2008). It may also be that for strong students, the particular content of Organic Chemistry 2 is much different from the content of the other Chemistry courses in which they had positive experiences with up to that point, which may further lower their competency beliefs and contribute to attrition decisions. These considerations will be important in the development of targeted interventions that help support women in persisting along this trajectory, which in could additionally provide opportunities to formally test causal hypotheses relating to gendered pre-med attrition.

Maintaining the high ratio of women showing interest in medicine is integral to improving innovation in the medical field, and improving equity within health-related profession (Bates et al., 2016; Bickel, 2005; Kvaerner et al., 1999; Reed & Buddeberg-Fischer, 2001) and society more broadly. There is an increasing need for a larger and more diverse STEM workforce, including more women (Page, 2007). Like gendered attrition found in other fields (i.e. engineering, computer science), losing women from highly influential and higher salaried positions in healthcare represents a loss in potential contributions to the field, and perpetuates problematic wage and power inequities across male-dominated vs. female-dominated health professions (Beede et al., 2011; Oh & Lewis, 2011). Further, medical research often fails to account for sex-based health differences, sometimes leading to misdiagnoses (Johnson et al., 2014; Mazure & Jones, 2015); patients who see female physicians may receive higher quality care in general (Tsugawa et al., 2017), and particularly regarding women's health issues (Siriwardena et al., 2012). Maintaining early medical school interest, particularly for high-performing female students, could address these concerns by increasing the number of female physicians occupying high level positions in various medical specialties.

However, we also acknowledge that while this study attends to a "supply-side" explanation of women's representation in the medical profession, we believe this to be an important but not sufficient condition for alleviating gender inequities in medicine, particularly in specialty fields. There are also significant structural barriers for women within medical institutions on the "demand-side", such as workplace harassment and discriminatory hiring practices, that limit women's advancement in male-dominated specialty areas (Boulis & Jacobs, 2008; Davis & Allison, 2013). Therefore, gender equity will require both increased representation and a concurrent dismantling of discriminatory practices that prevent women from reaching leadership positions in these fields and perpetuate the perception of some medical specialties as inhospitable workplaces for women.

3.5 Conclusion

Our analyses provide evidence from a large, multi-cohort study of undergraduate pre-medical students, using a large institutional dataset to understand where differential losses by gender occur along the undergraduate pre-med pathway. Further, we present a method for applying a broad range of prior and concurrent academic ability and motivation data to characterize the nature of this attrition in depth, demonstrating some common explanations of this phenomenon that fail to hold in our sample, and suggesting alternate models which provide strong explanatory power, and favor a hypothesis offering a greater potential for intervention. This provides a strong and simple metric for other institutions interested in identifying potential sources of gendered attrition in pre-med science courses; namely, focusing on continuation of students who earn an A or B within these core course sequences by gender. Important to note is that these findings also show evidence that this phenomenon has implications for the overall numbers of science undergraduate degrees for women. Further, by identifying courses in which these gaps are most problematic, and providing evidence against absolute and relative academic performance explanations, this study provides a foundation for interventions focused on addressing underlying causes (e.g., gender role models; Rosenthal, Levy, London, Lobel, & Bazile, 2013; Sanfey, 2006), negative instructor and peer messages (Archer et al., 2012; Schunk & Meece, 2006; Vincent-Ruz & Schunn, 2017; Zohar & Bronshtein, 2005) and directly targeting motivational factors that appear to be instrumental in large gender gaps found in Organic Chemistry, in Introductory Physics, and in taking the MCAT.

4.0 Study 3: Sources of Gender Differences in Competency Beliefs and Retention in an Introductory Pre-Medical Science Course

Gender disparities in retention in science pathways continues to vary widely by course. Undergraduates intending to study pre-health and pre-medicine often represent a majority of students enrolled in introductory science courses, contribute to a large number of eventual science degree earners, and are a population that typically includes a high number of women. However, gender differences in attrition, grades, and attitudes persist in the introductory science courses required for undergraduate pre-heath and pre-medical students, particularly within the physical sciences (i.e., Chemistry and Physics). We use structural equation modeling to study 416 undergraduate students across multiple sections of an Algebra-based Physics course, a common course on the pre-health and pre-medical track where large gender differences in grades, retention and competency beliefs have been documented. Our analysis focuses on identifying potential academic and attitudinal sources for gender differences in students' beliefs about their Physics abilities at the end of the course, and retention to the second physics course, which is often influenced by these competency beliefs. Results suggest that while men's ability beliefs in Physics are relatively stable and largely derived from early performance indicators, this is a smaller source of ability beliefs for women. Instead, women's ability beliefs are mediated during the course through their sense of belonging in Physics, and the extent to which they believe that Physics ability is fixed or malleable. These findings can inform the design of interventions in Physics courses that specifically target the development of ability beliefs for women intending medical careers.

4.1 Introduction

Despite recent reports showing an increase in the number of women earning STEM degrees more broadly and near parity in science overall, persistence in specific science majors continues to vary widely by gender. For example, while in general the number of women earning undergraduate degrees in the life sciences has surpassed men in recent years (e.g., Biology and Neuroscience, 72% women), the physical sciences continues to be male-dominated (e.g., Chemistry and Physics, 38% women; National Science Foundation, 2015). This lack of gender diversity has been linked to reduced innovation (Beede et al., 2011). The low participation by women in physical sciences also perpetuates gender wage disparities: Graduates with a Bachelor's degree in the physical sciences have a higher median wage ($65, 000) than those in the life sciences ($54,000), resulting in a $9,000 wage penalty in the science degrees more commonly earned by women (Carnevale et al., 2015). However, many studies of gender differences in STEM continue to treat the sciences as a single discipline, an approach that potentially obscures continued gender disparities between different science disciplines. The current study focuses on attitudes and retention within a physical science course sequence that has previously been shown to have pervasive gender disparities.

Prior work often also conceptualizes retention across science courses as primarily an issue for science majors, even though many introductory science courses are dominated by students intending to pursue pre-health or pre-medical studies. Pre-medical education in the United States takes place at the undergraduate level, and typically requires a rigorous sequence of science courses across a variety of disciplines, including multiple courses in Biology, Chemistry and Physics. At our own institution, approximately two-thirds of all undergraduates in the College of

Arts & Sciences enter the university with the intention to pursue medicine, and these students dominate many of the introductory science courses.

Studies using nationally representative data have also shown that at the end of high school, women are more than twice as likely as men to express interest in pursuing careers in medicine or health (Sadler et al., 2012). Further, students who enter these courses with the intent to pursue medicine are likely to have larger variation in attitudes towards particular science courses than students choosing those courses with a specific interest and intent to major in those particular disciplines (Gasiewski, Eagan, Garcia, Hurtado, & Chang, 2012). For example, while some pre-medical students may be drawn to the human services aspects of a career in medicine, others may be interested in the salary and prestige of such professions, while other still may enter with scientific interests. Women in particular have been shown to be more likely to choose pre-health professions with an express interest in working with people, and therefore may see introductory science courses simply as necessary stepping stones to a health profession (P. H. Miller, Blessing, & Schwartz, 2006).

Importantly for the number of degrees earned in the sciences, while a high number of students in introductory science courses may initially enter university intending pre-health or pre-medicine, relatively few students overall actually persist into those professions. Students intending pre-medicine and pre-health but who do not go on to take the medical school entrance exam have been shown to represent a larger proportion of eventual science degree earners than do students initially intending to major in science; however, women initially intending those professions are often more likely than men to switch into non-STEM fields entirely (Witherspoon & Schunn, in review). Therefore, negative experiences in introductory science courses required for pre-health and pre-medical studies may be particularly important contributors to differential attrition of women from those courses, and act as pathways away from, rather than towards, science-related majors and careers (Barr, 2010).

4.1.1 Sources of Gender Differences and the Primacy of Competency Beliefs

There is overwhelming research from psychology to suggest that observed gender differences in the sciences are socially constructed, rather than derived from biological or cognitive sex differences. Small differences that have been found are often in opposite directions such that the means are just as often higher for women. Further, all differences fall far below a threshold that would explain the large gender differences in participation that are observed (Hyde, 2005).

Therefore, current work focuses more on the underlying and concurrent social and psychological factors. For example, evidence suggests that women are often socialized at a young age to participate in different science-related activities than men, leading to disparities in prior experience that can contribute to differences in both performance and beliefs about ability at later stages (Jones, Howe, & Rua, 2000; Vincent-Ruz & Schunn, 2017). Further, particularly in disciplines where women are underrepresented and where negative stereotypes about women's abilities exist, even high-performing women are more likely than men to attribute difficulties in the subject to inherent skill, rather than the difficulty of the content or aspects of the learning environment (Beyer & Bowden, 1997). Particularly in higher education research, separately understanding the effects of differences in prior experience and preparation as well as the more immediate effects of differences in men and women's experiences of undergraduate science learning environments can help inform targeted interventions in these processes at the undergraduate level.

Specifically, students' competency beliefs (beliefs about their abilities within certain domains) have been shown across a number of studies to be a significant predictor of gender differences in both grades and persistence in STEM fields more broadly, and the physical sciences in particular (Hazari et al., 2010; Huang, 2013; Nissen, 2019; X. Wang, 2012). Competency beliefs (often also called self-efficacy) is grounded within the larger framework of social cognitive theory, and has been defined as the perception of one's ability to learn or perform a particular task to a certain level of proficiency (Bandura, 1989; Schunk & Pajares, 2002). There is abundant empirical support to suggest that competency beliefs in science are related to both academic persistence and career commitment (Lent, Ireland, Penn, Morris, & Sappington, 2017; Multon, Brown, & Lent, 1991), and that gender differences in competency beliefs are likely to be discipline specific (Debacker & Nelson, 2000; Huang, 2013).

While many studies indicate that differences in competency beliefs may be an important factor in gender gaps that appear in introductory science courses, less is known about the particular sources of competency beliefs for pre-medical students in these courses, and whether these are differential by gender. One study of pre-medical students shows that self-efficacy beliefs in Chemistry partially mediate gender differences in decisions to persist to the second Organic Chemistry course, even for men and women who earn the same high grade (i.e., an A or B) in the first Organic Chemistry course (Witherspoon et al., 2019). However, this effect should be replicated in other required physical science courses along the pre-medical pathway. Further, understanding how men and women in pre-medical physical science courses come to develop different beliefs about their abilities can be important to informing interventions that help to mitigate gender differences in competency beliefs, performance, and persistence in these courses.

Thus, the current study will seek to replicate the central role of competency beliefs in grades and retention in a physics course.

4.1.2 Sources of Competency Beliefs

There is a large body of extant research on possible sources of competency beliefs (see Usher & Pajares, 2008 for a review). Bandura (1986) identified four primary sources of competency beliefs: *mastery experiences* (i.e., interpretations of their past academic performance), *vicarious experiences* (i.e., students' observations and comparison of the activities of relevant others in the class), *verbal and social persuasions* (i.e., the feedback that students receive from others regarding their abilities), and *physiological and affective states* (i.e., the physical and mental feelings experienced in a course, and interpretations of those experiences). Each of these could be connected to women's experiences in physical science courses (Zeldin & Pajares, 2000)

Much of the research on sources of competency beliefs have primarily been concerned with which sources are most strongly correlated with competency beliefs (Usher & Pajares, 2008). However, by gender there remains uncertainty in the literature if men and women experience differences in sources of competency beliefs. Some studies claim to find no gender differences in sources of competency beliefs (Matsui, Matsui, & Ohnishi, 1990; Stevens, Wang, Olivárez, & Hamman, 2007) while others propose that competency beliefs are necessarily interpreted in context, thereby suggesting that an interaction between the individual and a particular domain determines the sources of competency beliefs that students attend to (Lent, Lopez, Brown, & Gore, 1996; Pajares, Johnson, & Usher, 2007; Trujillo & Tanner, 2014). For example, studies in science suggest that women are more likely to generate competency beliefs from vicarious experiences and verbal and social persuasions, while men are more likely to draw on mastery experiences

(Sawtelle, Brewe, & Kramer, 2012). This suggests that while prior experiences and performance (i.e., *mastery experiences*) may be the primary source of competency beliefs for men, other environmental factors may be more influential for women in science. While there are relatively few studies that have tested sources of competency beliefs within specific science domains, this earlier work can provide a foothold for identifying constructs that could contribute to the differential development of competency beliefs in science courses by gender.

Sense of belonging. Recent studies have documented how the lack of a "sense of belonging" (i.e., feelings of membership and acceptance in a group) can lead to lower rates of persistence for women in male-dominated physical science domains like Engineering and Physics (Lewis et al., 2017; Walton, Logel, Peach, Spencer, & Zanna, 2015). Research examining sense of belonging in a general pre-medical context suggest it may mediate interest in continued pre-medical study by gender (Rosenthal et al., 2013), while studies in a Calculus-based Physics course have shown belonging to be equally associated with higher grades for both men and women (Stout, Ito, Finkelstein, & Pollock, 2013). Understanding the effects of sense of belonging on grades and persistence may therefore be particularly important in introductory physical sciences courses for non-majors, as these courses are often required for pre-medicine, typically enroll a higher proportion of women, and yet exhibit greater attrition for high-performing women (Witherspoon et al., 2019).

Early research has also established a direct temporal link between feelings of belonging and academic competency beliefs, through primarily these studies have been conducted with middle school students (McMahon & Wernsman, 2008; Roeser, Midgley, & Urdan, 1996). While studies in higher education have also shown a high correlation between belonging and competency beliefs, the directionality of these effects is less clear; some drawing from social cognitive career theory instead model belonging and competency beliefs both as predictors of interest (Tellhed, Bäckström, & Björklund, 2017).

Particularly in an introductory physical science course for non-majors, sense of belonging may be related to two key sources of competency beliefs for women: *vicarious experiences,* and *affective states*. Uncertainty about belonging for underrepresented groups is believed to function in part through inducing stereotype threat (i.e., anxiety about confirming negative stereotypes about one's group), which may unconsciously interfere with cognition and inhibit performance (Steele & Aronson, 1995; Walton & Cohen, 2007). This mechanism also can contribute to gender differences in the interpretation of relative effort in domains like the physical sciences, where women may be more likely to experience belonging uncertainty. If stereotype threat leads to difficulties in performance, women may interpret this mismatch of effort and performance as an indicator of lower abilities (Smith, Lewis, Hawthorne, & Hodges, 2013). Further, students' beliefs about their abilities may be particularly likely to develop through both feedback and comparison to others during new and transitional experiences in science such as the first course in college, as internalized metrics of successful performance have not yet formed and are thus relatively unknown during these periods (Bandura, 1997; Eccles, Midgley, & Adler, 1984). Thus, the current study will examine how this factor predicts changes in competency beliefs for both men and women within introductory physics.

Implicit theories of intelligence. Another separate but related element of the learning environment that may contribute to gender differences in the development of ability beliefs are implicit theories of intelligence that students hold about a particular domain (e.g., theories of intelligence in physics). Adopting theories about the nature of ability in a field as "fixed" (i.e., ability is innate and unchangeable), rather than "malleable" (i.e., ability is changeable and can to be developed through effort), is detrimental to students' future achievement and persistence (Blackwell, Trzesniewski, & Dweck, 2007). This mechanism is thought to operate through influencing students perception of effort; students who believe that ability is linked to effort see challenges as part of growth, while students who believe that ability is fixed see effort as an indicator of lack of ability (Dweck & Leggett, 1988). Therefore, the development of a more fixed belief about the nature of intelligence in a field could provide a lens through which *vicarious experiences* and *social persuasions* are interpreted; students with a fixed mindset may be more likely to interpret their effort relative to others as a lack of inherent ability, or negative feedback about their ability as something that cannot be changed.

Recent studies of students in middle and high school science courses support this hypothesis; latent profile analyses show that students with more malleable beliefs about ability are more likely to accept competency beliefs feedback from multiple sources than those with more fixed mindsets, and as a result showed larger gains in competency beliefs (J. A. Chen & Usher, 2013). Importantly, fixed mindsets may be particularly salient and detrimental to women in historically male-dominated fields like the physical sciences. Studies have shown that faculty perceptions of domains that require "innate brilliance" are negatively correlated with the number of women in those fields, with many of the physical sciences being the most extreme along those dimensions (Leslie et al., 2015).

Further, students' implicit theories of intelligence may contribute separately to their ability beliefs, or in combination with other factors like sense of belonging, in ways that produce gender differences in competency beliefs. For example, research in mathematics suggests that sense of belonging for women may be particular low if they believe negative stereotypes about women mathematics ability, and if they believe that mathematics ability is a natural inherent ability that is difficult to change (Good, Rattan, & Dweck, 2012). Similarly, women in the physical sciences could also interpret a combination of environmental messaging that contribute to beliefs about their ability: the extent to which they interpret the particular domain as requiring an innate set of skills, as well as the extent to which they sense that they are accepted as full members of that domain. Thus, the current study will examine the role of physics theories of intelligence in predicting changes in competency beliefs for both men and women.

4.1.3 Current Study Overview

The current study addresses gaps in the current literature by utilizing a combination of historical institutional data and intensive course-level survey data to examine gender differences in the development of competency beliefs, grades, and retention to the next course, an introductory Algebra-based Physics course sequence for non-Physics majors. This course was selected as it is, along with Organic Chemistry, one of the required introductory physical science courses for pre-medical students that often show large differences in attrition for underrepresented groups (Barr et al., 2010; Witherspoon et al., 2019). Physics in particular is one of the core science domains that has continued to have difficulty attracting and retaining women (Cheryan et al., 2016; Matz et al., 2017). Women in Physics tend to underperform relative to their male peers on Physics assessments, as well as have lower beliefs about their ability to perform in Physics courses (Marshman, Kalender, Schunn, Nokes-Malach, & Singh, 2017). This could in part be related to differences in the perception of the importance of mathematics ability in the physical sciences, and differences in perceived mathematics ability by gender. For example, performance on the SAT Mathematics section may be more strongly related to lower beliefs about Chemistry ability for women than men (Vincent-Ruz et al., 2018). Relatedly, men are more likely to overestimate their performance in mathematics-related fields, which can contribute to women's artificially low perception of their abilities in science courses like Physics, a field which is widely perceived to be both mathematics-related and disproportionately male (Bench, Lench, Liew, Miner, & Flores, 2015). However, much of the research on gender differences in Physics focus on how these differences appear among Physical science or Engineering majors, and therefore focus on the sequence of calculus-based courses.

Understanding gender differences within the algebra-based Physics courses more typically taken by pre-medicine or pre-health majors can help to corroborate or uncover new sources of mechanisms of how discipline-specific academic and attitudinal gender differences manifest along a large pathway to science-related careers. Further, unlike Calculus-based Physics, Algebra-based Physics are typically more equally populated by gender, reducing the potential effects of pure numeric underrepresentation in the classroom for women. Therefore, in this way the current study builds on prior work conducted in Calculus-based Physics classrooms that examines the impact of belonging on gendered grades and participation in Physics (Lewis, Stout, Pollock, Finkelstein, & Ito, 2016; Stout et al., 2013), and expands on it in an Algebra-based Physics course by incorporating measures of students beliefs about the fixed or malleable nature of Physics ability, and contrasting the effects of these broader attitudes about physics with concurrent measures of content knowledge and indicators of prior experience.

Importantly, we also include the longitudinal development of self-efficacy as a key outcome that has been shown to contribute to gender differences in physical sciences courses along the pre-medical track. Such analyses allowed us to test not just cross-sectional associations in mean levels of competency beliefs, but also the relationship between changes in these variables over time, and the mediation of the stability of ability beliefs by gender.

Therefore, the current study addresses the following research questions in an introductory, Algebra-based Physics sequence for non-majors:

1) the relative contributions of prior knowledge and prior competency beliefs to grades and retention by gender;
2) the extent to which differences in the development of competency beliefs is mediated by broader beliefs about belonging and inherent Physics ability; and
3) if the contributions of these broader beliefs about Physics are moderated by gender.

4.2 Methods

4.2.1 Sample

Analyses were conducted with a sample of (N=416) undergraduate students (M_{age}=21, SD_{age}=1.3) enrolled in an Algebra-Based Introductory Physics course sequence at a large research-intensive public institution in the northeastern United States (henceforth, "the University"). Data were gathered through both in-course surveys, as well as from institutional warehouse data provided by the University, both with Institutional Review Board approval.

The University is broadly representative of similar public research institutions with a relatively selective admission rate (approximately 60%): it offers over 100 undergraduate majors, a high percentage (60%) of in-state students, and a small percentage (5%) of international students. While there is large variability in family income (*SD* = $122, 000), students tend to come from upper middle-income brackets (*Mdn* = $111,000). The sample was predominately female (58%), which is typical in the Algebra-Based Introductory Physics courses at this institution, a course which enrolls primarily pre-medical students[1]. The racial and ethnic diversity of the sample roughly mirrored that of the University as a whole; students were predominantly White (67%), with Asian (19%) and Black and Hispanic (15%) students making up the next two largest ethnic groups.

4.2.2 Measures

Attitudes towards physics. All students enrolled in this course were given surveys assessing domain-specific attitudes towards Physics, including competency beliefs, their sense of belonging, and theories of intelligence. All of the instruments were previously developed and validated (Marshman et al., 2017) with introductory physics students at the University through an iterative qualitative (cognitive interviews) and quantitative process (e.g., factor analyses). Item response theory analyses validated the use of means with these Likert-based scales and assured there was no differential scale functioning by gender.

In the current sample, no mean was near the max or min of each scale, which would have limited usefulness in analysis (see Table 6). *N*s varied due to variable attendance variable during recitation for pre- and post- measurements. In terms of discriminability, the highest correlation among these attitudinal variables was between competency beliefs and belonging, showing a moderate correlation of r=.55 (see Table 6). Importantly, the correlations between attitudes and

---

[1] Calculus-Based Introductory Physics courses, not included in these analyses, typically enroll engineering students and are often predominately (~70%) male.

ability in physics were relatively small (i.e., are separable aspects of performance with likely separate drivers).

Physics competency beliefs. The primary attitudinal outcome variable for this study was students' competency beliefs in Physics. It was measured both at the beginning of the semester, and at the end of the first semester, prior to the final exam. This scale contained discipline-specific items measuring students' beliefs about their ability to perform well on physics assessments and to understand physics concepts (e.g., "If I wanted to, I could do well on a physics test") rated on a four-point Likert scale (e.g., 1=NO! to 4=YES!). This scale (α=.79) was computed as the mean across five items, with reliability similar to what has been observed in prior studies in this same context (see Marshman et al., 2017).

Physics belonging. Belonging items measured the extent to which students felt as though they were a member of the particular Physics class (e.g., "Sometimes I worry that I do not belong in this physics class") rated along a 5-point Likert scale (e.g. 1= "Not at all true" to 5="Completely true). This scale (α=.83) was computed as a mean across five items, and it was measured at the end of the semester.

Physics theories of intelligence. Theory of intelligence items measured the extent to which students agreed with statements that described growth mindset (e.g., "Anyone can become good at solving Physics problems through hard work") or fixed mindset (e.g., "To really excel in Physics, a person needs to have a natural ability in Physics") in relation to Physics, rated along a 4-point Likert scale (e.g. 1= "Strongly Disagree" to 4="Strongly Agree). This scale (α=.79) was computed as a mean across seven items, with fixed growth items reverse coded, to generate an overall scale where higher numbers reflect higher endorsement of growth mindset, and lower numbers represent a more fixed mindset. It was measured at the end of the semester.

Table 6. Scale descriptives and Pearson inter-correlations among key academic (top) and motivational (bottom) variables.

				Intercorrelations								
Variables	N	M	SD	1	2	3	4	5	6	7	8	9
1. Physics 2 Retention	416	78%	-	1								
2. Physics 1 GPA	416	3.0	1.0	0.42***	1							
3. Conceptual Understanding Pre	392	37%	17%	0.07	0.32***	1						
4. Conceptual Understanding Post	386	56%	19%	0.20***	0.50***	0.67***	1					
5. SAT Math	323	669	68	0.15**	0.44***	0.45***	0.51***	1				
6. Competency Beliefs Pre	386	2.6	0.6	0.25***	0.42***	0.36***	0.35***	0.29***	1			
7. Competency Beliefs Post	416	2.8	0.6	0.15**	0.23***	0.33***	0.25***	0.23***	0.46***	1		
8. Belonging	414	3.5	0.9	0.24***	0.38***	0.28***	0.33***	0.27***	0.55***	0.40***	1	
9. Growth Mindset	416	2.8	0.5	0.22***	0.14**	0.20***	0.21***	0.21***	0.49***	0.37***	0.47***	1

Note. * p<.05, ** p<.01, *** p<.001

Academic performance and persistence variables. Students' academic performance was operationalized as the grade in their Physics course, measured along a 4-point GPA scale. A measure of students' prior academic performance was also collected from the University data warehouse, focusing on the measure that is most relevant for predicting performance in introductory Physics courses: students highest recorded score on the Math section of the Scholastic Achievement Test (SAT; or the converted score from the math sub-score of the ACT).

Separately, student's Conceptual Understanding of the course content was also measured using the Force Concept Inventory, a commonly-used, research-validated measure of students' deep conceptual understanding of the main content of the first introductory Physics course (Savinainen & Scott, 2002). By contrast, course grades involve a mixture of performance and task completion (e.g., homework points). Conceptual Understanding scores was gathered at the beginning and end of the semester during recitation as a percentage of correct items.

Finally, to examine students' persistence along the Physics course sequence, retention to the second algebra-based Introductory Physics course was measured with a single binary variable from the institutional data warehouse showing whether the student enrolled in Physics 2 at any point after the semester they took Physics 1, with 1=continued to Physics 2 and 0 = did not continue to Physics 2.

4.2.3 Analyses

There was a small amount of missingness (all < 8%) in the attitudinal survey data; a brief correlational analysis showed that a missingness indicator variable was not significantly associated with any of the predictors of interest, and so a maximum likelihood estimator was used to account for missingness in the following analyses. Examining intraclass correlations of all variables showed a large proportion of the variance for one variable (Physics 1 GPA; *ICC* = .52) was at the recitation level; therefore, robust standard errors were calculated to allow correlated errors to be nested within recitation.

The structural equation modeling (SEM) package in Stata 16 was used for all analyses. SEM enables a simultaneous regression analysis of the different attitudes and academic performance outcomes, while accounting for their inter-correlations. This produces a more robust analysis of gender effects within the correlated pathways of attitudes and prior academic factors as drivers of Physics grades and retention. SEM also uses a maximum likelihood estimator which produces less biased standard errors in the presence of missing data on the indicator variables; however, the small amount of missingness in this dataset is not likely to contribute to a significant source of bias in these analyses. Model fit were assessed using a variety of fit statistics, including CFI, TLI, RMSEA and SRMR, using conventional cutoffs as thresholds for good model fit (i.e., CFI/TLI > 0.95, RMSEA <0.06, SRMR <0.08; L. T. Hu & Bentler, 1999).

The evaluated model simultaneously tested two hypothesized sources of students' competency beliefs in Physics: A) an environment and prior beliefs about physics model in which changes in students' competency beliefs were mediated by their sense of belonging and growth mindset in Physics, and B) a performance feedback model in which changes in competency beliefs were explained through prior academic performance and current Physics knowledge, as measured by the SAT math section and the conceptual understanding tests (See Figure 11). For all pathways, we tested a moderated-mediation hypothesis, to understand whether one or both of the pathways were differential strength by students' gender (e.g., did sense of belonging or growth mindset matter more for one gender?).

Model building progressed by first running a fully unconstrained model, both for all students and moderated by gender as a multi-group SEM. Next, all paths that were significant for both men and women were constrained to equality, and a Lagrange multiplier test (Sörbom, 1989) was applied to evaluate if any of these constraints should be relaxed and be freely estimated by gender. Finally, this model was compared to a model with all path coefficients constrained to equality, and a model with intercepts constrained, to see if a model with those equal for men and women showed a better fit.

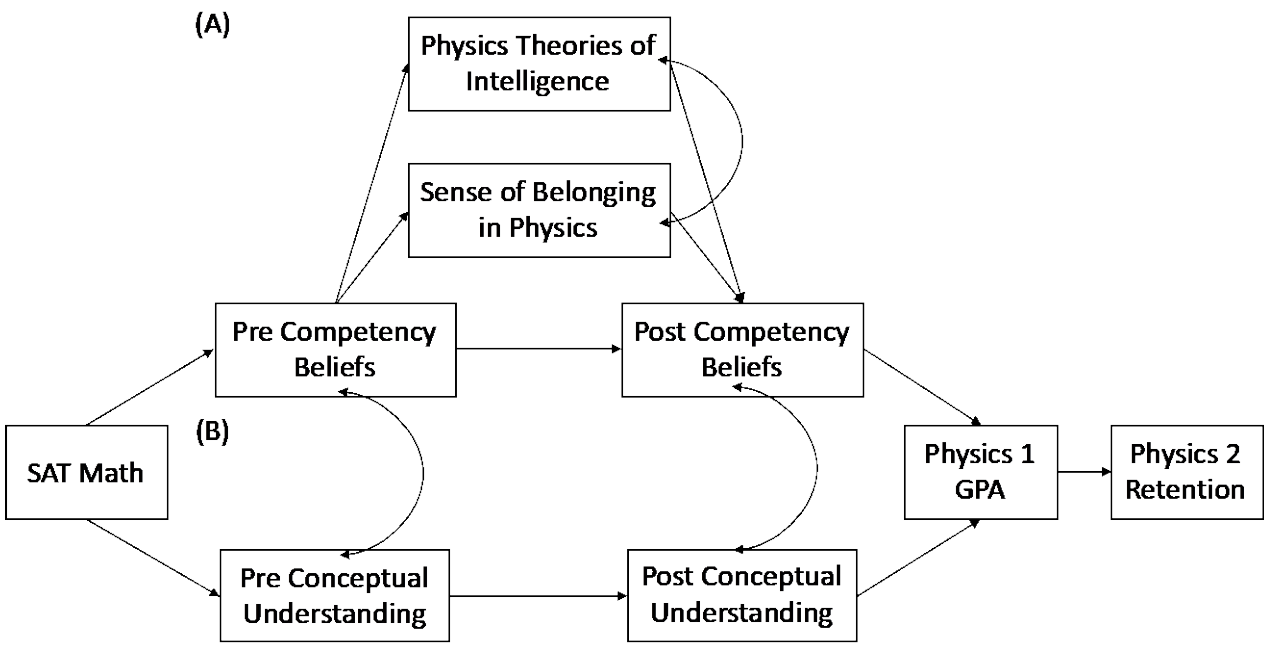


Figure 11. The proposed structural path model

The model tests the contribution of SAT Math to Physics performance and retention through (A) attitudinal factors (i.e., competency beliefs, belonging, theories of intelligence), and (B) conceptual understanding (i.e., FCI), by gender.

4.3 Results and Discussion

Fit statistics suggest that compared to the overall model across all students, a multi-group model by gender showed equally good fit ($\Delta\chi^2(18) = 15.6$, $p = .63$), providing motivation to continue exploration of more specific moderation by gender. Next, all path coefficients that were significant for both men and women were constrained to equality across gender, to test a more parsimonious model. Of those, a Lagrange multiplier test identified that model fit would improve if a single parameter were unconstrained and freely estimated by gender, the path from SAT Math to initial Conceptual Understanding ($\chi^2(1) = 5.9$, $p < .05$), and so the constraint by gender was removed for this one path. The resulting partially constrained model showed no decline in fit and a slight improvement in other fit statistics from the fully unconstrained model ($\Delta\chi^2(9) = 7.2$, $p = .62$). To test if the remaining parameters were indeed different by gender, we also tested the fit of a model with all path coefficients constrained to equality, and a fully constrained model. Both showed a significant decline in fit (see Table 7), suggesting a model allowing for the remaining paths to differ by gender produced the best fit to the data. Further, many paths showed substantial moderation by gender, including cases in which a path was only significant for one gender. Therefore, this model was selected as the final model (see Figure 12).

Table 7. Preliminary fit statistics for the proposed SEM models (acceptable values in bold).

Model	df	RMSEA	CFI	TLI	SRMR	χ^2	$\Delta\chi^2$	Δdf	p
Overall	18	0.06	**0.98**	0.95	**0.07**	37.2	---	---	---
Constraints by Gender									
Unconstrained	36	0.06	**0.98**	0.95	**0.07**	52.7	15.5	18	0.625
Significant Paths *	45	**0.05**	**0.98**	**0.96**	**0.07**	56.0	7.2	9	0.615
All Paths	51	0.06	**0.97**	0.95	0.09	73.7	15.9	6	0.032
All Paths and Intercepts	70	0.10	0.86	0.86	0.14	163.3	89.6	19	0.000

Note. * Final reported model

4.3.1 Predictors of Retention and Physics GPA

Direct effects. Building from the final outcome (Physics 2 retention), path coefficients showed that for both men and women, grade in Physics 1 is the strongest predictor of retention to Physics 2 (β=0.43, p<.001). Interestingly, there is a small direct effect of Physics Theory of Intelligence on retention, which is significant for women (β=0.16, p<.05) but not for men (β=0.10, p=.23; see Figure 12). This moderated path could explain why prior longitudinal work with a large multi-cohort sample (Witherspoon et al., 2019) found that women were less likely to persist to the second algebra-based Physics course, even when they have equivalent grades.

Turning to the penultimate outcome (Physics 1 GPA), for both men and women, there were similarly sized direct associations of grade with knowledge on the conceptual Physics exam (β=0.31, p<.001) and with end-of-course competency beliefs (β=0.27, p<.001; see Figure 12). Prior performance on the SAT Math also contributed to all students' Physics GPA both directly (β=0.22, p<.001), and indirectly as a predictor of Conceptual Understanding at both pre and post. This supports prior work suggesting that Physics is a deeply mathematical science, in which mathematics can act as an important resource for exam performance and conceptual understanding of quantitative laws (Meltzer, 2002).

Indirect effects. The total indirect effects from initial Conceptual Understanding to Physics 1 grade were significant for both men (β=0.23, $p<.001$) and women (β=0.18, $p<.001$); however, for women, a larger proportion of that effect (20% vs. 14%) was through a significant cross-lagged effect of initial performance on end of course Competency Beliefs. This aspect may reflect perceptions of how easily the content was mastered. Such effort / perceived success experiences are inherently ambiguous regarding source, and women may be more likely to attribute difficulty/ease to their inherent skills rather than the general difficulty of the content to be mastered (Beyer & Bowden, 1997).

Initial competency beliefs also significantly contributed to Physics 1 GPA indirectly, for both men (β=0.08, p<.01) and women (β=0.08, p<.01). Overall, the relatively large indirect and direct effects of pre and post competency beliefs on final grades in this course underscores the importance of attending to sources of competency beliefs in this model of student performance, and retention (which is driven by performance).

4.3.2 Sources of Competency Beliefs

Direct effects. Initial Competency Beliefs was significantly associated with SAT Math score for men (β=0.29, p<.001), but for women this direct effect was not significant (β=0.08, p=.27; see Figure 12). Turning to sources of Competency Beliefs at post, students' broader beliefs about physics were shown to differentially contribute to their end-of-course competency beliefs by gender. In particular, for both men and women, their sense of belonging in Physics at the end of the course was a significant contributor to their end-of-course Competency Beliefs (β=0.35, p<.001). By contrast, students' Physics Theory of Intelligence was only a significant contributor to end-of-course Competency Beliefs for women (β=0.30, p<.001), and not for men (β=0.09, p=.30; see Figure 12).

Interestingly, for both men and women, the direct covariations of Competency Beliefs and Conceptual Understanding were small and not statistically significant at pre (β_{Men}=0.18, p=.06; β_{Women}=0.12, p=.10) or post (β_{Men}=-0.02, p=.85; β_{Women}=0.13, p=.08). Overall, with the exception of the one cross-lagged connection for women from pre-Conceptual Understanding to post Competency Beliefs, Competency Beliefs and Conceptual Understanding were largely independent. The moderately-sized concurrent Pearson correlations appear therefore to reflect indirect relations through prior academic experiences and performance: they are useful for doing well and act as sources of initial higher Competency Beliefs, but the Competency Beliefs are not based in any kind of veridical, direct self-assessment.

Mediation effects by gender. Overall, when including these measures of students' broader beliefs in Physics as mediators, the association between initial to end-of-course competency beliefs remained significant for men (β=0.27, p<.001), while these factors partially explain the association of pre- to post-competency beliefs for women, such that the direct association between pre- and post-competency beliefs was no longer significant (β=0.12, p=.053). In other words, broader beliefs about Physics mediated only 30% of the total effect of initial competency beliefs on end-of-course competency beliefs for men, while those same factors mediated 59% of the total effect for women. Importantly, results suggest that while sense of Belonging is important for both men and women (mediating 26% and 43%, respectively) Theories of Intelligence in Physics were only significantly important for maintaining women's competency beliefs, mediating 24% of the total effect on end-of-course competency beliefs for women, compared to only 8% for men.

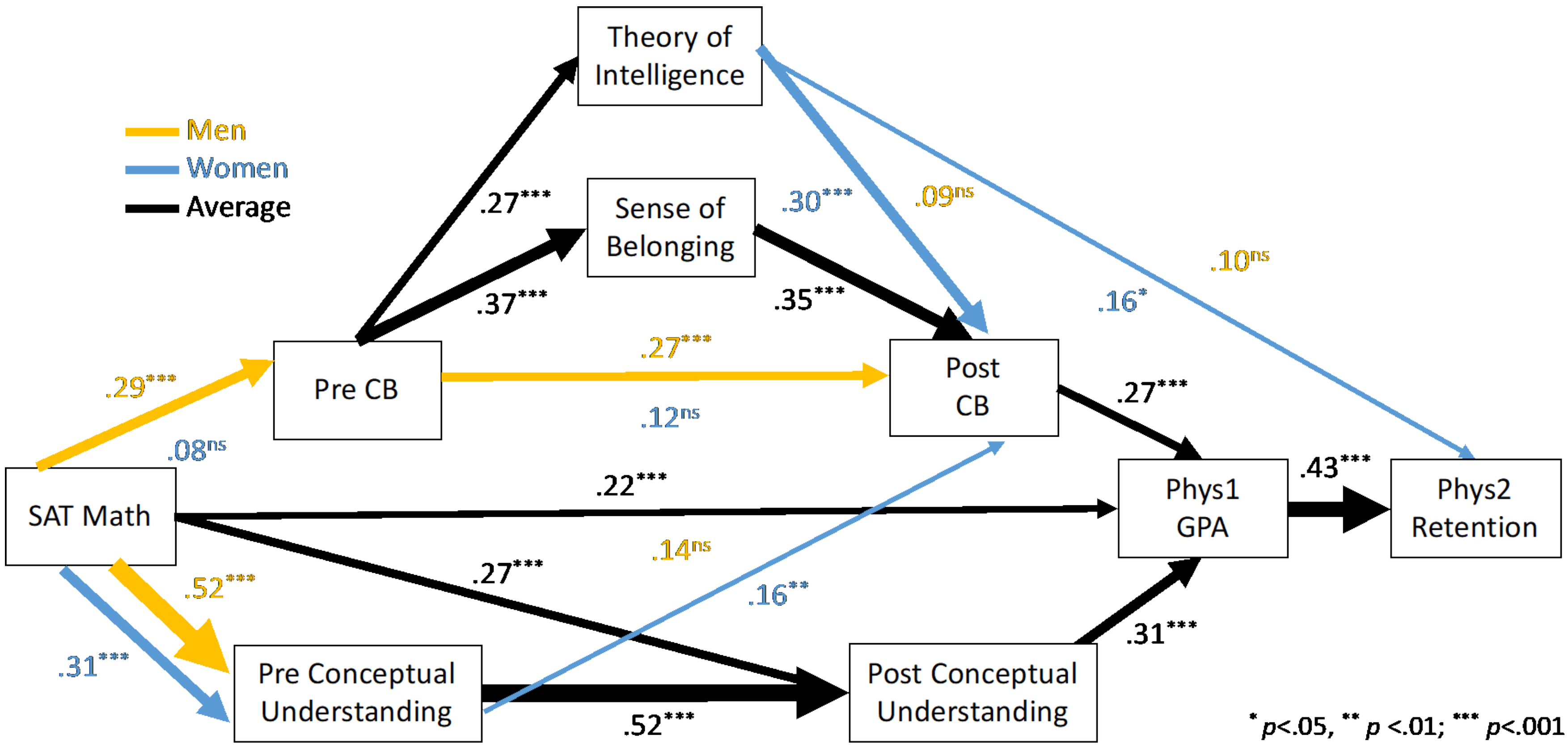


Figure 12. The partially constrained path model, with standardized beta coefficients shown.

Color paths show gender-specific connection strengths, and black show connection strengths averaged across gender. Line thickness indicate relative strength of effect, dashed lines are non-significant. Estimated non-significant paths and covariances (Cov(SoB,ToI) =.42*; Cov(PreCU,PreCB) =.15ns; Cov(PostCU,PostCB) =.08ns) are not shown in the diagram for interpretability.**

4.4 General Discussion

Despite recent reports that suggest growing gender *equality* in science (i.e., equal numbers of men and women earning degrees overall), attending to *equity* in science (i.e., equal opportunities for men and women to earn any degree) must address gender differences in retention across specific science domains, and in particular the physical sciences, where large gender differences persist. In this study, we examine one potential source of continued inequity in a particularly large and gender diverse population of students enrolling in introductory physical science courses: undergraduate pre-medical students enrolled in an introductory Algebra-based Physics course. Our results show that retention in this physical science course for both men and women is primarily driven by prior grades; however, that direct effect is heavily driven by competency beliefs, centering its role in both grades and retention. The finding that actual understanding of physics content is poorly aligned with students' competency beliefs in physics further draws attention to the importance of competency beliefs as its own target of intervention (i.e., it is not enough to just improve understanding of the content; students need to feel competent, too).

Just as importantly in the current study, competency beliefs were found to have both overlapping and non-overlapping sources by gender. Sense of belonging appeared as an important source of ability beliefs for all students (although women were less likely to have a high sense of belonging). However, only for women, the extent to which they believe Physics intelligence to be fixed or malleable acts as a strong predictor of changes in competency beliefs from the beginning to the end of the course.

These results are particularly informative in an introductory physical science course for non-majors, where stereotype threats for women may be especially salient. In a predominately male Physics environment, it may be possible that women who enter those undergraduate courses were able to persist to that point because they have been well-supported and encouraged through earlier educational experiences by high school teachers based on their relatively high aptitude in Mathematics and Science (Seymour, 1995). However, upon entering a more selective pool of students and confronting more difficult content in college, individuals with high initial competency beliefs may experience a “Little Fish, Big Pond” effect (Marsh, Trautwein, Lüdtke, & Köller, 2008), where their prior assessment of their abilities relative to their peers is recalibrated. For women in an historically male-dominated discipline like Physics, this can lead to attribution biases, where women are more likely to attribute a decline in performance to their own lack of ability, whereas men may get their ability beliefs from a broader variety of sources (Beyer, 1990). If sense of belonging and theories of intelligence mediate the stability of competency beliefs for women, and lower competency beliefs are associated with lower grades, attitudes and broader beliefs about Physics can lead to a self-fulfilling prophecy that operates separately from ability. That is, women who feel like they do not belong in Physics and believe that negative performance feedback is a stable indication of their ability could experience declining grades, further solidifying these attitudes and beliefs. At the same time, men who do not experience this mediation effect may experience a relative boost in self-efficacy, which can also continue to perpetuate stereotypes of male-dominance in Physics.

It is interesting that our findings did not show gender differences in effects of belonging on retention, supporting prior work with similar results in Calculus-based Physics (Stout et al., 2013). It may be that as a later course along the pre-med sequence, there has already been a significant “weeding out” process of students with low belonging, who do not persevere to the point of taking Physics. Alternatively, work examining gender differences in retention has proposed a relative strengths hypothesis for high performing students, suggesting that not lack of ability, but instead high ability in other areas, may better predict women’s choices to pursue alternative majors and career trajectories outside of STEM (Breda & Napp, 2019; M. Te Wang et al., 2013). Similarly, future work could test a hypothesis about “relative” belonging, to see if this framing demonstrates gender differences not in overall belonging, but instead relative belonging for men and women in the physical sciences, as compared to their concurrent sense of belonging in other fields (Thoman et al., 2014).

4.4.1 Limitations

It is important to acknowledge certain limitations of the current study. First, the data collected and analyses shown here remain correlational, and therefore strong inference about casual relationships between our variables cannot be drawn. However, the longitudinal data collected and the structural modeling method used do allow for some inference about temporal ordering; for example, it is impossible for the directionality of an effect of later competency beliefs on prior performance on the SAT to be reversed. Interventions with a control group targeting particular points within the model would provide a stronger casual test of the hypotheses derived from the current findings.

Second, while we were able to demonstrate that the selected model was a good fit to the data, structural equation modeling allows for a near infinite number of alternative model specifications, some of which might also adequately fit the data and would therefore provide alternative theoretical explanations. Our model building process addressed and rejected the most obvious alternative explanations for our particular research questions, future work testing similar models will help to corroborate these results. For example, though not reported here, some preliminary analyses conducted to test the influence of recitation-level factors (i.e., variation of the proportion of women in a recitation section) showed inconclusive results, possibly due to a sample size which is lower than required to do these kinds of robust multi-level analyses by gender. However, understanding the effects of variation in course and recitation factors such as gender composition or relative attitudes and achievement could be a fruitful next step in understanding source of students' development of competency beliefs within these introductory courses.

Finally, the current study was conducted with a particular population of mostly pre-medical students in an Algebra-based Physics course, and data was collected at a single institution, limiting the potential generalizability of the results. Therefore, replication at other similar institutions and within other pre-med science courses should be conducted to see if these conditions alter the results found here. In addition, a multi-institutional approach would allow for intentional selection of variation of elements of the learning environment (i.e., proportion of women faculty, size of course sections, selectivity of the institution) to see if these also offer alternative explanations for the results found.

4.4.2 Implications for Practice

Practically, our results provide support for interventions in introductory physical science courses at the undergraduate level that are designed to address both students sense of belonging, and their implicit theories of intelligence about the nature of ability for pre-health and pre-medical students in the physical sciences. While prior preparation does play a role in pre-medical students' grades (and thereby retention) in physical science courses, our findings suggest that the attitudes and beliefs about these domains that are developing concurrent during undergraduate introductory courses contribute an equal amount to these factors. This is important in that it demonstrates that the responsibility of improving achievement and retention for undergraduate students in physical science courses does not solely fall on primary and secondary educators, but also will require input from higher education institutions in creating environments that support all students in developing and maintaining attitudes that contribute to their success in those courses.

For example, instructors of the physical sciences in higher education institutions should consider attending to not only developing understanding of the content, but also the development of competency beliefs in the content, while being aware that these may develop independently from actual content understanding, and differentially for men and women. Specifically, our results show that while interventions targeting improving a sense of belonging are likely to contribute to the development of competency beliefs for all students, this alone is unlikely to address the persistence gender gap found in the physical sciences. Instead, providing experiences that contribute to a greater sense of belonging in the physical sciences, in combination with demonstrating that ability in the physical sciences is not fixed but instead can be improved through effort, could be especially impactful for the development of competency beliefs for women. Examples of such interventions exist, and have demonstrated positive effects on grades and attitudes (Blackwell et al., 2007; Walton et al., 2015). However, interventions may need to be tailored to the unique goals and concerns of non-majors in the physical sciences. For example, the large population of pre-health and pre-medical students who enter these classes, is a particularly important group to consider because interventions targeting them could provide larger effects and contribute to a larger number of students, and women in particular, who choose to continue in the sciences rather than leave to pursue alternative careers outside of STEM.

4.5 Conclusions

Domain-specific investigations of gendered attrition in STEM point out the dangers of the "STEM pipeline" metaphor that is often used to describe the issue of underrepresentation in STEM fields. The pipeline metaphor does not accurately model or locate the discipline-specific sources of differences in grades, ability beliefs, and retention for women. In addition, depicting the issue in undergraduate education as simply a deficit of incoming preparation of underrepresented STEM majors can have the unintended effect of prescribing solutions such as increasing applicants, which do not address the broader range of social and environmental deterrents women experience during their undergraduate education (Blickenstaff, 2005). Longitudinal studies of the interactive effects of grades and motivations for underrepresented groups can help to pinpoint the timing and character of these negative experiences and suggest interventions that may alleviate some of the non-academic deterrents for women in entering various STEM careers.

Further, it is important that research continues to gather data on contextual factors in particular sources of students' sense of belonging and theories of intelligence in the physical sciences. Qualitative studies may be needed to uncover more nuanced understanding of other motivational aspects beyond those measured here that may influence subsequent grades and persistence (Seymour, 1995). For example, studies show that undergraduate teaching faculty are more likely to rate fields with fewer women as domains that require "innate brilliance", with

Physics as one of those domains with both the highest professor ratings of “brilliance required”, and the fewest women who persist to earn degrees (Leslie et al., 2015). Historical underrepresentation of women in these fields can lead to a cementing of these discriminatory associations between talent and gender, and heightened stereotype threat for women in those courses. Importantly, interventions in social belonging and theories of intelligence have shown these beliefs to be highly malleable (Blackwell et al., 2007; Walton et al., 2015), providing an avenue for intervention if they are found to be integral sources of gender differences in self-efficacy. Therefore, rather than continuing to target only improvement in performance, which places students as the primary source of change, it is important that further research in this area is conducted to understand the role faculty (Zohar & Bronshtein, 2005) and institutions can play in mitigating the “chilly climate” (Walton et al., 2015) that deters women from pursuing the physical sciences at the undergraduate level.

5.0 Discussion and Conclusions

Increasing the representation of women in the sciences is important both as a source of broader participation and innovation within an economy that is increasingly reliant on scientific and technical expertise, as well as a mechanism for addressing continued inequities in pay and influence in scientific fields. Decades of research focusing on early predictors of differences in STEM persistence by gender have accomplished important overall increases in the representation of women in STEM more broadly (National Center for Science and Engineering Statistics, 2017). However, this aggregate approach hides the contribution of specific disciplinary contexts that may contribute differentially to men and women's decisions to leave certain science course pathways at differential rates (Cheryan et al., 2016). Therefore, this dissertation examines gender differences in science participation by building on prior work suggesting a shift towards a paradigm considering multiple "pathways" towards and away from science majors and careers (Cannady et al., 2014; D. I. Miller, 2018; PCAST, 2012).

Towards this end, the three studies of this dissertation were organized around a series of interrelated questions addressing differences in gender participation in science. Specifically, I first examined undergraduate students' incoming academic intentions and eventual earned degrees to identify the largest potential pathways to earning a science degree and see if there were differences in these degree pathways by gender. After having identified the largest such pathway, I then examined particular courses along this pathway, to pinpoint the specific courses that may show larger differential attrition by gender, and allow for a better understanding of the extent to which the academic and attitudinal mechanisms that drive these differences are dependent on the context in which they occur. Finally, I conducted “deep dives” into two of the courses in the pathway that were shown to be particularly problematic in terms of gender equity, to compare the direct and indirect effects of academic and attitudinal factors, and began to unpack the ways in which perceptions of the disciplinary environment may lead to either the development or maintenance of gender differences in attitudes, grades, and retention.

5.1 What are the Largest Pathways to Science by Gender?

In each of the three studies presented in this dissertation, I identify the pre-medical pathway as a particularly large population of undergraduate students who enroll in introductory science courses and often end up earning undergraduate science degrees, supporting findings from prior research (Gasiewski et al., 2012). Study 1 showed that students initially intending to pursue pre-medicine and pre-health were more likely to earn degrees in the Physical and Life Sciences than students who initially enter college with the intent to pursue a specific science degree, providing support for a “pathways” rather than a “pipeline” approach to understanding undergraduates’ major and career decision-making processes in the sciences.

However, by gender, the results from Study 1 showed that there were large differences in both pathways into science and pathways away from the sciences for these groups. In terms of pathways into science, the largest population of incoming undergraduate women in the sample, those intending pre-health, were 12 percentage points *less* likely than their male pre-health counterparts to end up earning a science degree. In terms of pathways away from science, women entering with the intent to study pre-medicine were 14 percentage points *more* likely than their male pre-medical counterparts to end up with a degree entirely outside of STEM. By contrast, women who initially intend to study science are actually just as likely to finish with a science degree. Therefore, while pre-health and pre-medicine should be considered as large *potential* on-ramps to science, they cannot be assumed to equitably operate as such for both men and women. In fact, because of the large number of women initially entering pre-medical pathways, these results suggest there should be concern about how negative experiences in introductory science courses act as a dysfunctional science on-ramp for women entering pre-medicine, as they are less likely to go on to earn a science degree. In other words, the greater contributor to the production of gender differences in science degrees may be located within inequities in large on-ramps *to* science, rather than differences in the off-ramps *from* science.

Differentiation by science was an important consideration given the differential representation by women across the sciences. Indeed the results here reflected national trends, with more students earning Life Sciences degrees overall, and women earning more degrees in the Life Sciences than in the Physical Sciences (Cheryan et al., 2016; National Center for Science and Engineering Statistics, 2017). However, the findings here suggest this gender difference is *not* driven by the higher number of women who initially intend pre-health or pre-med. Among those pre-health and pre-medical students who persist to the point of earning a science degree, there are no gender differences in earning a degree in the Life or Physical Sciences. This contradicts arguments claiming that gender equity along the pre-medical pathway would only increase the already high proportion of degrees earned by women in the Life Sciences; instead, the data suggest this could contribute equally to degrees in the Life Sciences and Physical Sciences.

5.2 What are the Courses that Show Patterns of Attrition by Gender?

The pre-medical population often consists of relatively equal numbers of high performing men and women, who enroll in introductory science courses with a variety of interests and career motivations, including but not limited to interest and prior experiences with the core science disciplines required for pre-medical study: Biology, Chemistry, and Physics. Therefore, looking within the pre-medical pathway allowed me to examine how gender differences appear within these different disciplines, and to understand if the mechanisms that drive gender differences in these courses are discipline-specific. Results from Study 2 show that early introductory classes taken by pre-medical students (i.e., Introductory Biology and General Chemistry) show few gender differences in retention. However, women who earn the same high, passing grade in later Physical Science courses (i.e., Organic Chemistry and Introductory Algebra-based Physics) are almost twice as likely to not enroll in the next course of that sequence, and twice as likely to not continue on and take the medical school entrance exam (MCAT).

These results support both prior research and commonly held assumptions among students and faculty that these two courses often act as “gatekeepers” or “weed out” courses that act as barriers to students’ persistence within the pre-medical pathways, and particularly for underrepresented groups (Barr, 2010; Barr et al., 2010). Though beyond the scope of the current study, the consistent identification of these courses as barriers to pre-medical study for underrepresented groups urges the reopening of a conversation questioning the logic of including these courses as requirements for entrance to medical school. First, from a pathways perspective, freeing up pre-medical students from a restrictive sequence of science courses may reduce the number of pre-medical students, and women in particular, who leave STEM entirely. Further, research have demonstrated that the number of undergraduate science courses taken is a poor predictor of medical school and professional performance, and studies of pre-medical programs that have removed these "weed out" courses have shown that their students perform at a level equivalent to their peers taking the more traditional pre-medical courses (Lovecchio & Dundes, 2002; Muller & Kase, 2010). If the only remaining function of these courses in pre-medicine is to cull the number of students, and results show that this function is differentially impacting women and students of color, it is a policy that must be revisited.

5.3 What Academic and Attitudinal Mechanisms Predict Gendered Persistence?

Across each of the three studies presented in this dissertation, I test multiple explanatory mechanisms of observed differences in participation in the sciences. Specifically, I test hypotheses attributing the source of gender differences in grades, retention, and attitudes to the relative contributions of academic experiences and performance both prior to and during the undergraduate years, discipline-specific attitudes that develop during the undergraduate years, and examine the ways in which these two sources may work in combination to perpetuate gender differences. First, these studies examine the "relative-strengths" hypotheses that attribute women's attrition to the higher number of attractive alternative options outside of science, rather than deficits within science (Breda & Napp, 2019; M. Te Wang et al., 2013), and suggest alternative methodology for investigating this phenomenon. For example, Study 1 shows that ratios of more proximal indicators of relative success, comparing the grades received in science with those received in other domains, may provide a stronger indicator of undergraduates' perceptions of relative success than early indicators like the SAT. Second, results suggest that the power of relative strengths in mediating attrition decisions may differ by the particular pathway students are on. For example, Study 1 suggests that while relative strengths explain in part the decision of women intending pre-medicine to pursue a non-STEM degree, this does not explain the decision of women intending pre-health to leave the sciences. This varied role of relative-strengths by pathway indicates that the meaning of relative strengths in non-science domains may be interpreted differently depending on the varied characteristics and career motivations of students in introductory science courses.

The results also highlight the central role of competency beliefs in predicting gender differences in grades and retention in Physical Science courses along the pre-medical pathway, supporting prior findings in the literature (Hazari et al., 2010; Huang, 2013; Nissen, 2019; Vincent-Ruz et al., 2018). While grades in the prior course have the largest direct effect on retention for both men and women, competency beliefs also explain an equally large proportion of the variance in retention. In Study 2, I find that for students' decision to continue on to enroll in Organic Chemistry 2, competency beliefs partially explain the difference in retention between men and women who earn the same high, passing (i.e., A or B) grade in Organic Chemistry 1. In Study 3, results show that in Introductory Algebra-based Physics, competency beliefs play a similar role for men and women influencing grades (and thus indirectly, retention). However, the way in which men and women develop their beliefs about their abilities in this particular domain differ. For both men and women, their incoming beliefs about their abilities in Physics strongly determines the extent to which they feel a sense of belonging in Physics, which in turn contributes to the maintenance of ability beliefs at the end of the course. Further, for women only, their beliefs that intelligence in Physics is innate also contributes to their ability beliefs at the end of the course, and the extent to which they hold these beliefs explains the stability of their ability beliefs throughout the course.

Interestingly, measures of conceptual knowledge and beliefs about ability in Physics remain relatively disconnected for both men and women, emphasizing the importance of interventions targeting students' competency beliefs separately and in addition to interventions that work to improve students actual conceptual understanding. This is different from what has been shown in prior studies, which find that in male-dominated domains, men typically overestimate their abilities, while women underestimate their abilities (Beyer & Bowden, 1997). It may be that in a Physics course for primarily pre-medical students that is no longer male-dominated, the relatively high number of women reduces one source overestimation of abilities for men in that domain.

5.4 Limitations and Implications for Future Research

These studies have some limitations that should be carefully considered when drawing conclusions from the results, and that also provide suggestions for fruitful areas of future research. First, it is important to draw clear boundaries around the generalizability of the findings presented here. These studies were all conducted entirely in the context of the United States, which has a unique structure for medical school in that it requires first obtaining an undergraduate post-baccalaureate degree in a variety of fields (e.g., in neuroscience, in chemistry, in biology, in psychology) with no medicine-specific degree earned at the undergraduate level at most universities, including the one providing the dataset analyzed here. In contrast, undergraduates in many European countries begin medical study directly after the completion of secondary school (Riska, 2010). Therefore, while these studies examine an important and large population for understanding undergraduate science attrition in the U.S., these findings are likely to vary widely in countries where the structure of medical study looks much different.

It is likely not a coincidence then, that among the OECD countries the U.S. has one of the lowest proportion of women practicing medicine (OECD, 2018). It is possible that gender retention in undergraduate medical programs in these countries are more similar to narrowly focused undergraduate training programs like Nursing or Engineering in the United States, where there is likely to be stronger initial commitment and more stable persistence. Following the relative strengths hypothesis, there may be fewer opportunities in these more specialized programs where students are drawn away from science and towards other alternatives where they perceive fewer barriers, unlike liberal arts settings with a variety of general education requirements. In addition, for students self-selecting into these more specialized learning environments, course content may be more tightly aligned with specific career and subject interests, which could make prior experiences more relevant, produce a greater sense of belonging, and thus offer more sources for developing competency beliefs in that domain. While some international studies of medical education exist (Kvaerner et al., 1999; Riska, 2011), future studies testing how the mechanisms here replicate when implemented in these varied learning environments can help to understand ways in which the U.S. pre-medical education structure contributes to women's lower participation in medicine.

Relatedly, these studies draw inferences from data collected entirely at a single institution. While the University studied here is similar to a number of large public research institutions in both its student population and the structure of its introductory pre-medical science courses, it would be important for future studies to intentionally vary these institution-level contextual factors. Particularly, this will allow testing of whether the discipline specific results found here are consistent across institutions, or if there is greater correlation between gender differences and institutional variation in demographics or course structure. For example, some U.S. universities already have students take Organic Chemistry before General Chemistry, yet relatively little research has empirically tested whether these alternative course sequences change the size and location of gender differences in those courses on learning, attitudes, and retention. Including institutional variation will be important for continuing to uncover the specific conditions under which these results may vary; however, the identification here of commonly known and well-documented "gatekeeping" physical science courses like Organic Chemistry and Physics suggests the context studied can provide a good baseline understanding of common processes in a large number of large public institutions with the most common pre-medical course sequence.

Incorporating multiple institutions with different student populations would allow for the possibility of testing if the particular mechanisms modeled here are appropriate for also describing attrition decisions for student of other groups historically minoritized in science (i.e., Black and Latinx students, first-generation students, LGBTQ students) who are pursuing pre-medical study. Critically, multi-institutional data would increase the sample size, allowing for analysis with sufficient power to explore attitudinal mechanisms for intersectional groups (i.e., Black men, first-generation Latinx students) who may have unique experiences in these introductory science courses and were not in sufficiently large numbers to study separately in the dataset investigated in this dissertation. For example, some studies suggest that sources of self-efficacy like vicarious learning experiences typically ascribed to women may be more salient to Black men, as their relative underrepresentation contributes to a particularly lower sense of belonging in STEM

(Byars-Winston & Rogers, 2019). Encouraging academic independence and competition over interdependence, a socialization process common in higher education and particularly in medical training (Lempp & Seale, 2004), may also lead to the uncoupling of performance and perceptions of ability for first-generation students and students from lower socioeconomic households, by discouraging support-seeking behaviors (Calarco, 2014; Stephens, Fryberg, Markus, Johnson, & Covarrubias, 2012). Recent research has also suggested that the lower sense of belonging experienced by women in STEM may extend to sexual minority students (e.g., lesbian, gay, bisexual or queer students); these students are up to 10% less likely to be retained in STEM majors when controlling for equivalent prior experiences, implying the importance of non-academic factors. However, by gender there was a reversal effect for these students, such that sexual minority men were less likely to persist than heterosexual men, while sexual minority women were more likely to persist than heterosexual women (Hughes, 2018). Understanding how gender, race, class and sexual stereotypes contribute individually and in combination to students' sense of belonging, discipline-specific beliefs, and the development of competency beliefs in different science disciplines along the pre-medical track can further refine the constructs modeled here.

Another limitation to the design of the studies contained in this dissertation is that while I operationalize students' perceptions of their learning environment through self-reports of their experiences of belonging and beliefs about intelligence in that discipline, I was not able to collect direct measures of the environment to triangulate these self-reports or include them directly as predictors in the models. Therefore, while the constructs used provide insight into the way in which men and women perceive the environment differently, the specific aspects of the environment that may be influencing these perceptions remain unspecified. A variety of environmental sources, such as the characteristics of the instructors, the physical environment, and the class composition could all contribute to the differences observed; future research isolating these could continue to contribute to our understanding of how science learning environments can be shifted to create more equitable opportunities for all students. Hierarchical analyses can be one way to begin to separate out the individual and course level variation in some of the phenomena uncovered here; however, often institutions are hesitant to agree to release data that could be used to conduct evaluations at the instructor level, and at single institutions there are not enough individual courses to provide sufficient power for these types of multi-level analyses, as was the case with the University studied here. Some preliminary analyses suggest looking at variation at the recitation level (rather than larger lecture level) could provide a promising way forward, as the larger number of sections (and fewer students within each) provide greater variation in potential predictors (e.g., the proportion of men and women in a recitation section). However, in the data used in the current studies, diagnostic tests demonstrate that except in the stated case from Study 3 where robust standard errors were used to account for clustering, design effects for all variables were less than 2, and therefore the single level analyses used were unlikely to produce biased results (Maas & Hox, 2005; Muthen & Satorra, 1995).

Finally, while significant efforts were made to control for endogeneity, and the data collected contains some temporal precedence which allows for quasi-experimental designs and mediation analyses, the studies conducted here are correlational and descriptive, therefore providing a relatively low level of casual inference. Although we incorporate a pre- and post-test design for some measures, including additional timepoints, particularly for the measures of belonging and theories of intelligence, would allow for the use of additional quasi-experimental designs like cross-lagged models which would allow a stronger case for causality. Alternatively, future intervention work with control and treatment groups could provide the most direct test of the mechanistic processes modeled here.

5.5 Conclusions and Synthesis of Contributions Across Studies

Overall, this dissertation provides an approach to studying contextually situated mechanisms for the development of under-represented group differences in attitudes, grades and retention in the sciences. The dissertation focused on gender, but a similar approach could also be used for understanding other dimensions frequently discussed as being underrepresented in STEM such as race/ethnicity, first generation college attendees, and low-income students. First, by focusing on pathways to science rather than the traditional pipeline model, I allow for the incorporation of the wide variety of trajectories and timepoints students may take toward earning undergraduate science degrees. Second, I involve the disciplinary context as an important factor when considering how students make decisions regarding their undergraduate course-taking and majors. Third, within this context, I acknowledge both the direct and indirect influence of prior academic experiences and concurrently developing attitudes towards that discipline as important mechanisms that can perpetuate and influence gender differences beyond early indicators of interest or ability.

An important contribution of this approach is providing a preliminary step in shifting the focus of research on retention in science education away from characterizing between-group differences as “gaps”, and towards instead understanding mechanisms that account for the interaction between specific identities and contexts that may lead to differences in academic decision-making. For example, understanding the contexts in which there are gender similarities, as well as gender differences, can be an equally valuable contribution to our understanding of the sources of these differences and the way in which they are influenced by the science discipline, as well as the particular career and major pathways of particular students (see Table 8). In taking this approach, this dissertation attempts to turn away from the deficit framing that can often be present in many "gender gap" studies, while also providing suggestions as to a method through which research looking at group differences can move beyond "gap gazing" (Rodriguez, 2001) and begin to understand how structural and contextual changes can contribute to improving equity in the sciences.

Table 8. Similarities and differences by gender discovered in the current set of studies, across three levels of analysis.

	Similar by Gender	Different by Gender
Pathways to Science	• From intending science, to earning a science degree	• From intending pre-health and pre-medicine, to earning a science degree
Discipline-Specific Courses	• Retention from Introductory Biology 1 to 2 • Retention from General Chemistry 1 to 2	• Retention from Organic Chemistry 1 to 2 • Retention from Algebra-based Physics 1 to 2
Academic and Attitudinal Mechanisms	• Relative earlier academic strengths as a source of course-specific attrition • Competency beliefs as a contributor to grades • Belonging as a source of competency beliefs • Correlation between understanding and ability beliefs	• Relative proximal grades as a source of science degree attrition from pre-med • Competency beliefs as a mediator of retention for high-performing women • Theories of intelligence as a source of competency beliefs • Theories of intelligence as a source of retention

Specifically, the current collection of studies suggests that interventionists working in introductory science courses in the United States must be aware of the size and unique character of the population of students in these courses who are intending pre-medical study. While surveys of social cognitive constructs can provide important information about students' decision-making, certain items (e.g., discipline-specific belonging, theories of intelligence, and competency beliefs) may elicit different responses from men and women enrolling in those courses with the intent to pursue medicine, compared with those intending to pursue specific science degrees. Further, the studies in this dissertation identify particular areas that may provide the richest areas for effective interventions targeting gender inequities, both in terms of the location of particular courses, as well as particular constructs. For example, because large gender differences are not detected in earlier introductory science courses like Introductory Biology and General Chemistry, it may be more effective to target interventions for pre-medical students in Physical Science courses that are typically offered later in the undergraduate career: Organic Chemistry and Introductory Physics.

Finally, while this dissertation does not contain an intervention as a part of the design of the current studies, the findings here do lay some foundational groundwork for understanding the mechanisms and thereby informing the design of interventions that target gender equity in introductory science courses. Earlier work with similar social psychological interventions caution that such interventions should not be treated as a "magic bullet" or be "scaled up" without attending to important factors like the timing and particular context of the learning environment in which interventions are conducted (Cohen & Garcia, 2014; Sherman et al., 2013). Importantly, this work identifies the location and outlines the mechanisms of such interventions, suggesting a focus on academic ability, as well as the development of student attitudes towards the discipline, are both directly and indirectly involved in grades and attrition in these courses. Further, because vicarious experiences and social persuasions of others in the class are likely to be sources of core attitudes like competency beliefs, changing these attitudes will require more than attempts to directly influence individual attitudes. Instead, fostering a greater sense of belonging within disciplines with a history of negative stereotypes towards women, and shifting the classroom culture away from a belief that ability in the discipline is innate, and towards a belief that disciplinary knowledge can be learned and improved through effort, could contribute to the maintenance of ability beliefs for all students, and particularly for women.

Appendix A Supplemental Materials for Study 1

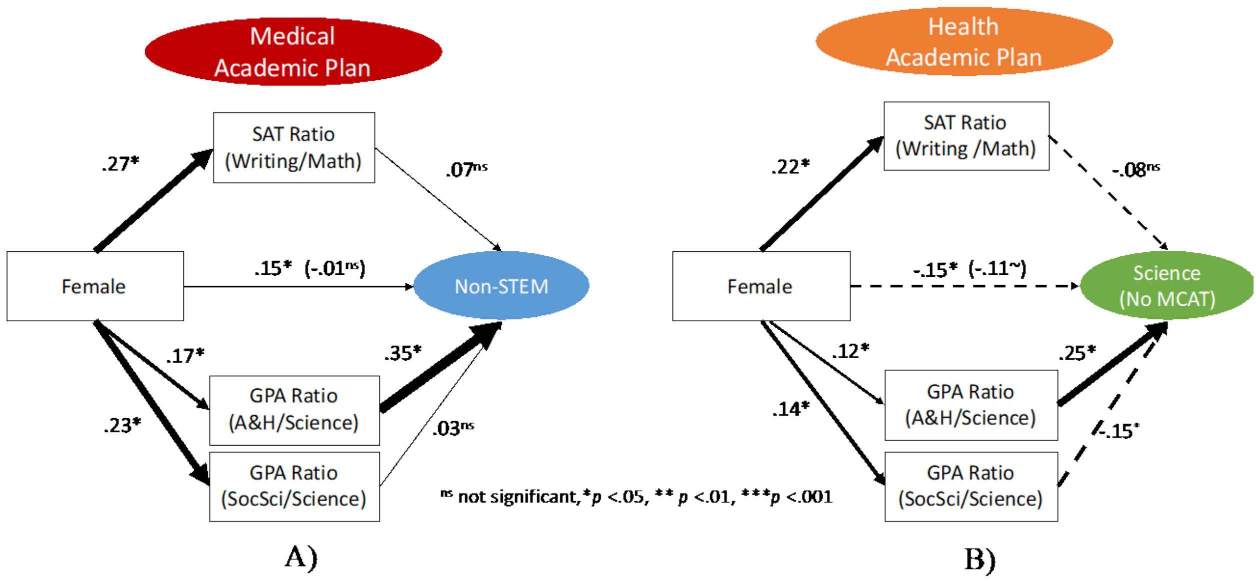


Figure 13. Mediation models of the (a) Medical and (b) Health pathways, using GPA ratios up to and including Organic Chemistry 1.

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.)

Appendix B Supplemental Materials for Study 2

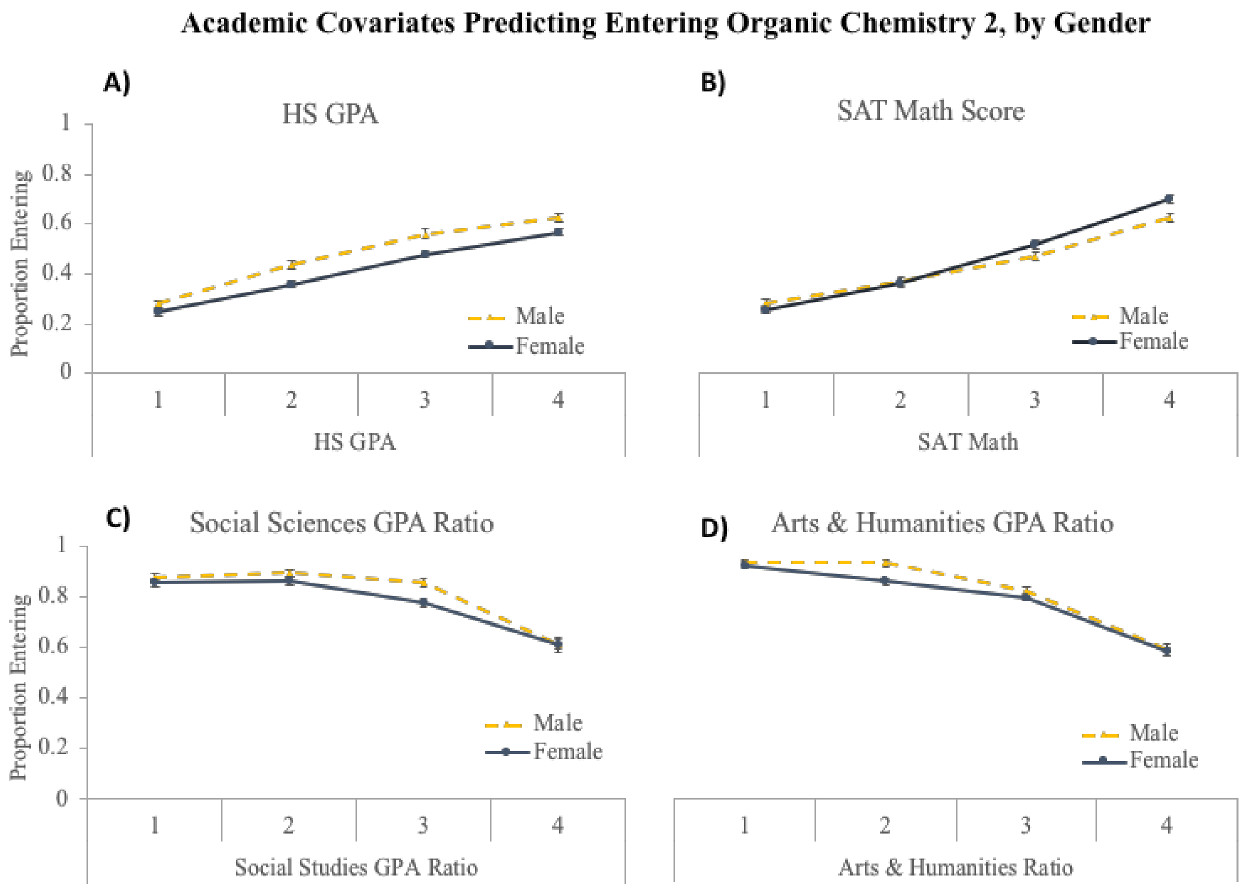


Figure 14. Proportion of students entering Organic Chemistry 2 (with SE bars), at quartiles of covariates. A) high school GPA; B) SAT Math score; C) Social Science ratio and D) Arts and Humanities ratio.

Table 9. Descriptive statistics for each motivation measure by gender for students earning an A or B, including number of items, reliability (Cronbach's alpha), t-tests, and point-biserial inter-correlations.

	# items	α	Female		Male		*t*	*p*	1	2	3	4
			N	M (SD)	N	M (SD)						
1. Fascination	3	0.77	200	3.0 (.62)	134	3.1 (.55)	1.85	0.065	1			
2. Competency Beliefs	4	0.85	199	2.9 (.50)	134	3.0 (.44)	3.04	0.003	0.55***	1		
3. Science Identity	4	0.89	198	3.6 (.46)	134	3.5 (.48)	-2.67	0.008	0.27**	0.19***	1	
4. Med Career Intent	1	-	200	59%	135	61%	-0.36	0.718	-0.10	-0.13*	0.20***	1

Note. **p*<.05; ***p*<.01; ****p*<.001

Table 10. The seven undergraduate degree categories, with examples of specific degrees that were included within each category.

Category	**Examples**
Health Professions	*Nursing, Dental Hygiene, Pharmaceutical Sciences, Sports Medicine, Nutrition and Dietetics, Emergency Medicine, etc.*
Social Sciences	*Anthropology, Economics, Political Science, Psychology, Sociology, etc.*
Arts and Humanities	*English, History, Philosophy, African Studies, Arts, Music Theater, Music, Languages, Religious Studies, etc.*
Science	*Biology, Microbiology, Chemistry, Neuroscience, Physics, Geology, etc.*
Math	*Mathematics, Accounting, Applied Math, Computer Science, Statistics, etc.*
Engineering	*Civil Engineering, Mechanical Engineering, Electrical Engineering, Chemical Engineering, etc.*
Business	*Business Accounting, Finance, Marketing, Human Resources and Management, etc.*

Table 11. Distributions of grades for each core pre-med science course, as percentages by gender.

	Grades									
	F		D		C		B		A	
	M	F	M	F	M	F	M	F	M	F
Introductory Biology 1	1.7	1.5	3.8	3.4	20.1	26.6	28.3	32.9	46.1	35.6
Introductory Biology 2	1.2	0.9	2.6	2.8	16.2	20.7	26.0	29.9	54.0	45.7
General Chemistry 1	2.3	1.4	3.4	3.0	28.6	31.1	37.1	37.6	28.7	26.9
General Chemistry 2	1.2	0.6	2.9	2.2	22.8	23.4	26.6	28.9	46.4	44.9
Organic Chemistry 1	1.5	1.4	1.9	1.9	14.1	17.5	19.7	19.3	62.9	59.9
Organic Chemistry 2	1.0	0.5	1.6	1.9	13.2	14.1	15.6	15.0	68.7	68.6
Introductory Physics 1	0.5	0.3	0.8	0.9	9.6	14.6	13.7	15.6	75.4	68.6
Introductory Physics 2	0.2	0.1	0.6	0.7	7.9	8.8	11.3	10.9	80.0	79.6

Table 12. Logistic regression models for continuing to each course after receiving an A or B in the previous course, controlling for cohort.

	Cohort			
Model	**OR**	**95% CI**	***z***	***p***
Introductory Biology 2				
Gender	0.85	0.68, 1.07	-1.40	0.162
Cohort	0.96	0.91, 1.02	-1.22	0.224
General Chemistry 2				
Gender	1.02	0.86, 1.21	0.27	0.789
Cohort	0.93	0.89, 0.97	-3.41	0.001
Organic Chemistry 2				
Gender	0.35	0.25, 0.48	-6.33	0.000
Cohort	1.00	0.93, 1.08	0.10	0.921
Introductory Physics 2				
Gender	0.62	0.49, 0.79	-3.90	0.000
Cohort	0.90	0.85, 0.96	-3.27	0.001
Take MCAT				
Gender	0.58	0.37, 0.91	-2.35	0.019
Cohort	0.56	0.49, 0.65	-7.87	0.000

Table 13. Correlations of continuing to Introductory Biology 2, General Chemistry 2, Organic Chemistry 2, Physics 2, and taking the MCAT with each other (upper left) and with academic covariates (lower left), along with inter-correlations among academic covariates (lower right).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Enter Bio2	1															
2. Enter Gen2	0.48***	1														
3. Enter Org2	0.38***	0.40***	1													
4. Enter Phys2	0.32***	0.30***	0.53***	1												
5. Take MCAT	0.09***	0.10***	0.23***	0.28***	1											
6. STEM AP Score	0.07***	0.19***	0.24***	0.12***	0.11***	1										
7. STEM AP Cred	0.09***	0.19***	0.21***	0.10***	0.08***	0.61***	1									
8. SAT Math	0.12***	0.20***	0.26***	0.13***	0.11***	0.47***	0.48***	1								
9. Non-STEM AP Score	0.09***	0.13***	0.17***	0.10***	0.08***	0.43***	0.42***	0.33***	1							
10. Non-STEM AP Cred	0.07***	0.10***	0.13***	0.06***	0.06***	0.37***	0.52***	0.28***	0.62***	1						
11. SAT Verbal	0.10***	0.11***	0.16***	0.08***	0.08***	0.37***	0.34***	0.49***	0.48***	0.41***	1					
12. SAT Writing	0.14***	0.15***	0.21***	0.13***	0.10***	0.37***	0.36***	0.50***	0.45***	0.39***	0.75***	1				
13. HS GPA	0.14***	0.17***	0.17***	0.08***	0.05***	0.23***	0.25***	0.21***	0.24***	0.25***	0.25***	0.26***	1			
14. Arts Ratio	-0.24***	-0.32***	-0.28***	-0.19***	-0.10***	-0.24***	-0.21***	-0.29***	-0.17***	-0.10***	-0.15***	-0.16***	-0.18***	1		
15. Social Sci Ratio	-0.19***	-0.27***	-0.23***	-0.16***	-0.09***	-0.20***	-0.19***	-0.26***	-0.11***	-0.06***	-0.09***	-0.11***	-0.12***	0.80***	1	
16. Cohort	-.02	-.07***	-.04***	-.07***	-0.17***	0.04	0.09***	0.09***	0.04	0.09***	0.18***	0.05***	0.08***	0.04***	0.06***	1

Note. **p<.05***, p<.01, ***p<.001*

Note. Ratio variable correlations shown are all courses; individual course ratios are used for attrition analyses.

Table 14. Table of Variance Inflation Factors (VIF) for regression models with STEM and Non-STEM covariates.

Variable		VIF	Tolerance
STEM	Introductory Biology 2	1.47	0.68
	General Chemistry 2	1.48	0.67
	Organic Chemistry 2	1.77	0.57
	Introductory Physics 2	1.52	0.69
	Took MCAT	1.14	0.89
	Female	1.05	0.96
	AP STEM Score	1.73	0.58
	AP STEM Credit	1.74	0.57
	SAT Math	1.49	0.67
	Cohort	1.05	0.95
	Mean VIF	1.44	
Non-STEM	Introductory Biology 2	1.62	0.62
	General Chemistry 2	1.45	0.69
	Organic Chemistry 2	1.52	0.66
	Introductory Physics 2	1.42	0.70
	Took MCAT	1.11	0.90
	Female	1.06	0.94
	AP Non-STEM Score	1.69	0.59
	AP Non-STEM Credit	1.56	0.64
	SAT Verbal	2.42	0.41
	SAT Writing	2.18	0.46
	HS GPA	1.13	0.89
	Arts GPA Ratio	2.06	0.48
	Social Science GPA Ratio	1.90	0.53
	Cohort	1.15	0.87
	Mean VIF	1.59	

Table 15. Logistic regression models for continuing to each course after receiving an A or B in the previous course, with all STEM and Non-STEM covariates.

	Uncontrolled				STEM				Non-STEM			
Model	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***
Introductory Biology 2												
Gender	0.85	0.68, 1.06	-1.46	0.143	1.01	0.80, 1.29	0.09	0.926	1.00	0.72, 1.37	-0.03	0.977
STEM AP Score					1.31	1.11, 1.54	3.22	0.001				
STEM AP Credit					1.10	0.93, 1.30	1.08	0.280				
SAT Math					1.42	1.24, 1.63	4.97	0.000				
Non-STEM AP Score									1.27	1.03, 1.57	2.27	0.023
Non-STEM AP Credit									0.88	0.72, 1.08	-1.24	0.214
SAT Verbal									1.07	0.85, 1.35	0.60	0.548
SAT Writing									1.07	0.85, 1.35	0.59	0.554
HS GPA									1.27	1.08, 1.50	2.82	0.005
Soc Sci Ratio									0.33	0.20, 0.52	-4.69	0.000
Arts Hum Ratio									0.67	0.44, 1.01	-1.91	0.056
General Chemistry 2												
Gender	0.99	0.84, 1.17	-0.11	0.910	1.08	0.90, 1.28	0.81	0.418	0.84	0.63, 1.10	-1.27	0.203
STEM AP Score					1.19	1.06, 1.33	2.96	0.003				
STEM AP Credit					1.16	1.03, 1.31	2.42	0.016				
SAT Math					1.13	1.02, 1.26	2.37	0.018				
Non-STEM AP Score									1.07	0.90, 1.28	0.82	0.414
Non-STEM AP Credit									0.98	0.83, 1.15	-0.26	0.793
SAT Verbal									0.93	0.76, 1.13	-0.75	0.451
SAT Writing									1.15	0.94, 1.40	1.33	0.183
HS GPA									1.24	1.07, 1.43	2.81	0.005
Soc Sci Ratio									0.44	0.30, 0.65	-4.18	0.000
Arts Hum Ratio									1.02	0.73, 1.42	0.12	0.905

Table 15, (cont.) Logistic regression models for continuing to each course after receiving an A or B in the previous course, with all STEM and Non-STEM covariates.

	Uncontrolled				STEM				Non-STEM			
Model	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***
Organic Chemistry 2												
Gender	0.35	0.26, 0.48	-6.39	0.000	0.37	0.27, 0.53	-5.66	0.000	0.34	0.24, 0.50	-5.66	0.000
STEM AP Score					1.05	0.87, 1.26	0.51	0.613				
STEM AP Credit					1.20	0.96, 1.49	1.59	0.113				
SAT Math					1.70	1.42, 2.03	5.85	0.000				
Non-STEM AP Score									0.77	0.63, 0.95	-2.39	0.017
Non-STEM AP Credit									1.32	1.05, 1.68	2.33	0.020
SAT Verbal									1.23	0.97, 1.57	1.68	0.092
SAT Writing									1.23	0.96, 1.57	1.64	0.101
HS GPA									1.09	0.91, 1.30	0.96	0.336
Soc Sci Ratio									0.71	0.52, 0.97	-2.15	0.031
Arts Hum Ratio									0.72	0.54, 0.97	-2.18	0.029
Introductory Physics 2												
Gender	0.59	0.46, 0.75	-4.36	0.000	0.61	0.47, 0.78	-3.90	0.000	0.56	0.42, 0.74	-4.14	0.000
STEM AP Score					1.05	0.90, 1.23	0.63	0.530				
STEM AP Credit					1.52	1.26, 1.83	4.31	0.000				
SAT Math					1.20	1.04, 1.39	2.49	0.013				
Non-STEM AP Score									1.10	0.92, 1.32	1.06	0.290
Non-STEM AP Credit									1.15	0.95, 1.40	1.46	0.145
SAT Verbal									1.17	0.97, 1.43	1.63	0.103
SAT Writing									1.13	0.92, 1.38	1.19	0.236
HS GPA									0.95	0.83, 1.08	-0.81	0.416
Soc Sci Ratio									0.66	0.49, 0.89	-2.74	0.006
Arts Hum Ratio									1.01	0.79, 1.29	0.08	0.932

Table 15, (cont.) Logistic regression models for continuing to each course after receiving an A or B in the previous course, with all STEM and Non-STEM covariates.

	Uncontrolled				STEM				Non-STEM			
Model	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***	**OR**	**95% CI**	***z***	***p***
Take MCAT												
Gender	0.53	0.35, 0.80	-3.06	0.000	0.56	0.37, 0.85	-2.75	0.006	0.54	0.34, 0.86	-2.60	0.009
STEM AP Score					1.18	0.93, 1.51	1.37	0.172				
STEM AP Credit					0.96	0.74, 1.24	-0.31	0.758				
SAT Math					0.86	0.67, 1.10	-1.22	0.222				
Non-STEM AP Score									1.05	0.80, 1.37	0.34	0.737
Non-STEM AP Credit									0.91	0.70, 1.18	-0.71	0.478
SAT Verbal									1.05	0.76, 1.44	0.29	0.774
SAT Writing									0.92	0.65, 1.29	-0.50	0.620
HS GPA									1.06	0.86, 1.31	0.54	0.592
Soc Sci Ratio									0.81	0.42, 1.59	-0.60	0.547
Arts Hum Ratio									0.95	0.51, 1.74	-0.18	0.858

Note.

Table 16. Proportion of students earning each category of degree by gender, and proportion of those degree earners taking the MCAT by gender, with odds ratios, 95% CI, z-test statistic and p-value shown.

	Female (N=3,048)		**Male (N=2,462)**					
	N	**%**	**N**	**%**	**OR**	**95% CI**	***z***	***p***
Degree earned								
Health	628	0.21	233	0.10	2.44	2.08, 2.87	10.87	0.000
Social Sciences	541	0.18	283	0.12	1.63	1.40, 1.91	6.21	0.000
Arts & Humanities	386	0.13	206	0.08	1.56	1.31, 1.87	4.91	0.000
Science	786	0.26	769	0.32	0.75	0.67, 0.84	-4.81	0.000
Bio	*389*	*0.13*	*341*	*0.14*	*0.89*	*0.77, 1.05*	*-1.40*	*0.162*
Chem	*72*	*0.02*	*111*	*0.05*	*0.50*	*0.37, 0.68*	*-4.45*	*0.000*
Physics	*4*	*0.00*	*10*	*0.00*	*0.31*	*0.10, 1.01*	-	-
Other	*321*	*0.11*	*307*	*0.13*	*0.81*	*0.69, 0.96*	*-2.45*	*0.014*
Math	46	0.02	92	0.04	0.39	0.27, 0.56	-5.17	0.000
Engineering	41	0.01	136	0.06	0.23	0.16, 0.33	-8.16	0.000
Business	102	0.03	103	0.04	0.78	0.59, 1.03	-1.74	0.082
Took MCAT								
Health	10	0.02	9	0.04	0.40	0.16, 1.00	-1.95	0.051
Social Sciences	25	0.05	23	0.08	0.55	0.30, 0.98	-2.02	0.044
Arts & Humanities	19	0.05	10	0.05	1.01	0.46, 2.22	0.04	0.971
Science	135	0.17	214	0.28	0.54	0.42, 0.69	-5.00	0.000
Bio	*64*	*0.08*	*91*	*0.12*	*0.54*	*0.38, 0.78*	*-3.35*	*0.001*
Chem	*9*	*0.01*	*18*	*0.02*	*0.73*	*0.31, 1.75*	-	-
Physics	*0*	*0.00*	*0*	*0.00*	-	-	-	-
Other	*62*	*0.08*	*105*	*0.14*	*0.46*	*0.32, 0.66*	*-4.18*	*0.000*
Math	1	0.02	0	0.00	-	-	-	-
Engineering	2	0.05	3	0.02	-	-	-	-
Business	1	0.01	2	0.02	-	-	-	-

Note. Proportions for Took MCAT calculated using total students taking the MCAT; Male (*N*=261) and Female (*N*=193). A dash (-) is shown in place of *z*-statistics and *p*-values for rows where *N*<10 for male or female.

Table 17. Proportion of students earning each category of degree by gender after receiving and A or B in all pre-med courses, and proportion of those degree earners taking the MCAT by gender, with odds ratios, 95% CI, z-test statistic and p-value shown.

	Female (N=143)		**Male (N=177)**					
	N	**%**	**N**	**%**	**OR**	**95% CI**	***z***	***p***
Degree earned								
Health	4	0.03	3	0.02	1.67	0.37, 7.58	-	-
Social Sciences	16	0.11	16	0.09	1.27	0.61, 2.63	0.64	0.525
Arts & Humanities	9	0.06	2	0.01	5.88	1.25, 27.65	-	-
Science	114	0.80	154	0.87	0.59	0.32, 1.07	-1.74	0.081
Bio	*62*	*0.43*	*65*	*0.37*	*1.32*	*0.84, 2.07*	*1.20*	*0.228*
Chem	*6*	*0.04*	*9*	*0.05*	*0.82*	*0.28, 2.35*	-	-
Physics	*0*	*0.00*	*0*	*0.00*	-	-	-	-
Other	*46*	*0.32*	*80*	*0.45*	*0.58*	*0.36, 0.91*	*-2.36*	*0.018*
Math	0	0.00	0	0.00	-	-	-	-
Engineering	0	0.00	0	0.00	-	-	-	-
Business	0	0.00	1	0.01	-	-	-	-
Took MCAT								
Health	1	0.25	2	0.67	0.17	0.01, 4.51	-	-
Social Sciences	9	0.56	9	0.56	1.00	0.25, 4.04	-	-
Arts & Humanities	6	0.67	1	0.50	2.00	0.09, 44.35	-	-
Science	44	0.39	90	0.58	0.45	0.27, 0.73	-3.19	0.001
Bio	*21*	*0.18*	*32*	*0.21*	*0.53*	*0.26, 1.08*	*-1.75*	*0.081*
Chem	*2*	*0.02*	*7*	*0.05*	*0.14*	*0.01, 1.44*	-	-
Physics	*0*	*0.00*	*0*	*0.00*	-	-	-	-
Other	*21*	*0.18*	*51*	*0.33*	*0.48*	*0.23, 1.00*	*-1.96*	*0.050*
Math	0	0.00	0	0.00	-	-	-	-
Engineering	0	0.00	0	0.00	-	-	-	-
Business	0	0.00	1	0.01	-	-	-	-

Note. Proportions for Took MCAT calculated using total students taking the MCAT; Male (*N*=261) and Female (*N*=193). A dash (-) is shown in place of *z*-statistics and *p*-values for rows where *N*<10 for male or female.

References

Ainley, M., & Ainley, J. (2019). Motivation and Learning. In *The Cambridge Handbook of Motivation and Learning* (pp. 665–688). Cambridge University Press. https://doi.org/10.1017/9781316823279.028

Alper, J. (1993). Science-Education: the Pipeline Is Leaking Women All the Way Along. *Science*, *260*(#5106), 409–411.

Archer, L., Dewitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). "Balancing acts''': Elementary school girls' negotiations of femininity, achievement, and science." *Science Education*, *96*(6), 967–989. https://doi.org/10.1002/sce.21031

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

Association of American Medical Colleges. (2017b). Required Premedical Coursework and Competencies. Retrieved from https://students-residents.aamc.org/choosing-medical-career/article/required-premedical-coursework-and-competencies/

Atkinson, R. D., & Mayo, M. (2011). *Refueling the U.S. Innovation Economy: STEM Reform*. Washington, DC: The Information Technology and Innovation Foundation. Retrieved from http://www.itif.org/files/2010-refueling-innovation-economy.pdf

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Bandura, A. (1989). Social Cognitive Theory. In R. Vasta (Ed.), *Annals of child development* (Vol. 6, pp. 1–60). Greenwich, CT: JAI Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

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 : Journal of the Association of American Medical Colleges*, *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

Bates, C., Gordon, L., Travis, E., Chatterjee, A., Chaudron, L., Fivush, B., … Moses, A. (2016). Striving for Gender Equity in Academic Medicine Careers. *Academic Medicine*, *91*(8), 1050–1052. https://doi.org/10.1097/ACM.0000000000001283

Bauer, C. F. (2005). Beyond "Student Attitudes": Chemistry Self-Concept Inventory for Assessment of the Affective Component of Student Learning. *Journal of Chemical Education*, *82*(12), 1864. https://doi.org/10.1021/ed082p1864

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

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. (2011). *Women in STEM : A Gender Gap to Innovation* (No. Issue Brief 04-11, U.S. Dept. of Comm., ESA). Retrieved from http://doi.org/10.2139/ssrn.1964782

Bench, S. W., Lench, H. C., Liew, J., Miner, K., & Flores, S. A. (2015). Gender Gaps in Overestimation of Math Performance. *Sex Roles*, *72*(11–12), 536–546. https://doi.org/10.1007/s11199-015-0486-9

Berryman, S. E. (1983). Who Will Do Science? Trends, and Their Causes in Minority and Female Representation among Holders of Advanced Degrees in Science and Mathematics. A Special Report. *A Special Report: The Rockefeller Foundation*, 148. Retrieved from http://files.eric.ed.gov/fulltext/ED245052.pdf

Beyer, S. (1990). Gender Differences in the Accuracy of Self- Evaluations of Performance. *Journal of Personality and Social Psychology*, *59*(5), 960–970. https://doi.org/10.1037/0022-3514.59.5.960

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

Bickel, J. (2005). Gender Equity in Undergraduate Medical Education : A Status Report. *Journal of Women's Health & Gender-Based Medicine*, *10*(3), 261–270.

Blackwell, K. L., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

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

Bollen, K. A., & Pearl, J. (2013). Eight Myths About Causality and Structural Equation Models, (April 2012), 301–328. https://doi.org/10.1007/978-94-007-6094-3_15

Boulis, A. K., & Jacobs, J. A. (2008). *The Changing Face of Medicine: Women Doctors and the Evolution of Health Care in America*. Ithica, NY: Cornell University Press. https://doi.org/10.1353/bhm.0.0285

Breda, T., & Napp, C. (2019). Girls ' comparative advantage in reading can largely explain the gender gap in math-related fields, *116*(31), 15435–15440. https://doi.org/10.1073/pnas.1905779116

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. *AJP: Advances in Physiology Education*, *37*(1), 70–79. https://doi.org/10.1152/advan.00138.2012

Byars-Winston, A., & Rogers, J. G. (2019). Testing intersectionality of race/ethnicity × gender in a social-cognitive career theory model with science identity. *Journal of Counseling Psychology*, *66*(1), 30–44. https://doi.org/10.1037/cou0000309.supp

Calarco, J. M. C. (2014). The Inconsistent Curriculum: Cultural Tool Kits and Student Interpretations of Ambiguous Expectations. *Social Psychology Quarterly*. https://doi.org/10.1177/0190272514521438

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

Cannady, M. A., Moore, D., Votruba-Drzal, E., Greenwald, E., Stites, R., & Schunn, C. D. (2017). How personal, behavioral, and environmental factors predict working in STEMM vs non-STEMM middle-skill careers. *International Journal of STEM Education*, *4*(22), 1–16. https://doi.org/10.1186/s40594-017-0079-y

Carnevale, A. P., Cheah, B., & Hanson, A. R. (2015). *The Economic Value of College Majors*. Washington, DC.

Chen, J. A., & Usher, E. L. (2013). Profiles of the sources of science self-efficacy. *Learning and Individual Differences*, *24*, 11–21. https://doi.org/10.1016/j.lindif.2012.11.002

Chen, X., & Soldner, M. (2013). STEM Attrition: College Students' Path Into and Out of STEM

Fields. *National Center for Education Statistics*, 2014–001. https://doi.org/http://nces.ed.gov/pubs2014/2014001rev.pdf

Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2016). Why are some STEM field more gender balanced than others? *Psychological Bulletin*, *142*(206), 1–131. https://doi.org/10.1037/bul0000052

Cohen, G. L., & Garcia, J. (2014). Educational Theory, Practice, and Policy and the Wisdom of Social Psychology. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 13–20. https://doi.org/10.1177/2372732214551559

Collett, D. (1994). *Modelling Survival Data in Medical Research*. London, UK: Chapman & Hall.

Confrey, J., Castro-Filho, J., & Wilhelm, J. (2000). Content-Based Collaborative Inquiry. *Educational Psychologist*, *35*(3), 193–206. https://doi.org/10.1207/S15326985EP3503

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

Crowley, K., Barron, B. J., Knutson, K., & Martin, C. K. (2014). Interest and the Development of Pathways to Science. *Interest in Mathematics and Science Learning and Related Activity*, 1–32.

Davis, G., & Allison, R. (2013). Increasing representation, maintaining hierarchy: An assessment of gender and medical specialization. *Social Thought and Research*, *32*, 17–45.

Debacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: Differences related to gender, class type, and ability. *Journal of Educational Research*, *93*(4), 245–254. https://doi.org/10.1080/00220670009598713

Dorph, R., & Bathgate, M. E. (2017). When I grow up : the relationship of science learning activation to STEM career preferences. *International Journal of Science Education*, *0*(0), 1–24. https://doi.org/10.1080/09500693.2017.1360532

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, *98*(1), 198–208. https://doi.org/10.1037/0022-0663.98.1.198

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

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

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment. In J. G. Nicholls (Ed.), *The Development of Achievement Motivation, Vol. 3* (pp. 238–331). Greenwich, CT: JAI Press, Inc.

Eccles, J. S., & Wang, M. Te. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, *40*(2), 100–106. https://doi.org/10.1177/0165025415616201

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

Fiorentine, R. (1987). Men, Women, and the Premed Persistence Gap: A Normative Alternatives Approach. *American Journal of Sociology*, *92*(5), 1118–1139. Retrieved from http://www.jstor.org/stable/2779998

Fiorentine, R., & Cole, S. (1992). Why fewer women become physicians: Explaining the premed persistence gap. *Sociological Forum*, *7*(3), 469–496. https://doi.org/10.1007/BF01117557

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From Gatekeeping to Engagement: A Multicontextual, Mixed Method Study of Student Academic Engagement in Introductory STEM Courses. *Research in Higher Education*, *53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

Gee, J. P. (2001). Identity as an Analytic Lens for Research in Education. *Review of Research in Education*, *25*(May), 99–125.

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. https://doi.org/10.1037/a0026659

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, *47*(8), 978–1003. https://doi.org/10.1002/tea.20363

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

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y

Hughes, B. E. (2018). Coming out in STEM: Factors affecting retention of sexual minority STEM students. *Science Advances*, *4*(3), 1–6. https://doi.org/10.1126/sciadv.aao6373

Husbands Fealing, K., & Myers, S. (2012). Pathways v. Pipelines to Broadening Participation in the Stem Workforce. *Ssrn*. https://doi.org/10.2139/ssrn.2020504

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

Institute of Education Sciences. (2018). *Undergraduate Enrollment*. *The Condition of Education*. Retrieved from https://nces.ed.gov/programs/coe/indicator_cha.asp

Johnson, P. A. ., Fitzgerald, T., Salganicoff, A., Wood, S., & Goldstein, J. (2014). *Sex-Specific Medical Research: Why women's health can't wait*. Boston, MA. Retrieved from https://www.brighamandwomens.org/Departments_and_Services/womenshealth/ConnorsCenter/Policy/ConnorsReportFINAL.pdf

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<180::AID-SCE3>3.0.CO;2-X

Kilbourne, A. M., Bauer, M. S., Damschroder, L., Hagedorn, H., & Smith, J. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology*, *3*(1), 1–12. https://doi.org/10.1186/s40359-015-0089-9

Kilminster, S., Downes, J., Gough, B., Murdoch-Eaton, D., & Roberts, T. (2007). Women in medicine - is there a problem? A literature review of the changing gender composition, structures and occupational cultures in medicine. *Medical Education*, *41*(1), 39–49. https://doi.org/10.1111/j.1365-2929.2006.02645.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.

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. *College Board Research Report No. 08-5*, *5*, 10.

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?* (No. Working Paper #23735, National Bureau of Economic Research). Retrieved from http://www.nber.org/papers/w23735

Kvaerner, K. J., Aasland, O. G., & Botten, G. S. (1999). Female medical leadership: Cross sectional study. *BMJ (Clinical Research Ed.)*, *318*(7176), 91–94. https://doi.org/10.1136/bmj.318.7176.91

Lempp, H., & Seale, C. (2004). The hidden curriculum in undergraduate medical education: qualitative study of medical students' perceptions of teaching. *BMJ*, *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*. https://doi.org/10.1006/jvbe.1994.1027

Lent, R. W., Brown, S., & Hackett, G. (2000). Contextual supports and barriers to career choice: A social cognitive analysis. *Journal of Counseling Psychology*, *47*(1), 36–49. https://doi.org/http://dx.doi.org/10.1037/0022-0167.47.1.36

Lent, R. W., Ireland, G. W., Penn, L. T., Morris, T. R., & Sappington, R. (2017). Sources of self-efficacy and outcome expectations for career exploration and decision-making: A test of the social cognitive model of career self-management. *Journal of Vocational Behavior*, *99*, 107–117. https://doi.org/10.1016/j.jvb.2017.01.002

Lent, R. W., Lopez, F. G., Brown, S. D., & Gore, P. A. (1996). Latent structure of the sources of mathematics self-efficacy. *Journal of Vocational Behavior*, *49*(3), 292–308. https://doi.org/10.1006/jvbe.1996.0045

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

Lewis, K. L., Stout, J. G., Finkelstein, N. D., Pollock, S. J., Miyake, A., Cohen, G. L., & Ito, T. A. (2017). Fitting in to move forward: Belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychology of Women Quarterly*, *41*(4), 420–436. https://doi.org/10.1177/0361684317720186

Lewis, K. L., Stout, J. G., Pollock, S. J., Finkelstein, N. D., & Ito, T. A. (2016). Fitting in or opting out: A review of key social-psychological factors influencing a sense of belonging for women in physics. *Physical Review Physics Education Research*, *12*(2), 1–10. https://doi.org/10.1103/PhysRevPhysEducRes.12.020110

Lindemann, D., Britton, D., & Zundl, E. (2016). “I Don’t Know Why They Make It So Hard Here”: Institutional Factors and Undergraduate Women’s STEM Participation. *International Journal of Gender, Science and Technology*, *8*(2), 21.

Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., & Hill, N. E. (2018). Multiple Pathways to Success: An Examination of Integrative Motivational Profiles Among Upper Elementary and College Students. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000245

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

Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables (Advanced Quantitative Techniques in the Social Sciences)*. London, UK: Sage Publications.

Lopez, M. H., & Gonzalez-Barrera, A. (2014). *Women’s College Enrollment Gains Leave Men Behind*. *Pew Research Center*. Retrieved from http://pewrsr.ch/1qckLFE

Lovecchio, K., & Dundes, L. (2002). Premed survival: understanding the culling process in premedical undergraduate education. *Academic Medicine : Journal of the Association of American Medical Colleges*, *77*, 719–724. https://doi.org/10.1097/00001888-200207000-00016

Maas, C. J., & Hox, J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-1881.1.3.86

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

Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social Comparison and Big-Fish-Little-Pond Effects on Self-Concept and Other Self-Belief Constructs: Role of Generalized and Specific Others. *Journal of Educational Psychology*, *100*(3), 510–524. https://doi.org/10.1037/0022-0663.100.3.510

Marshman, E., Kalender, Z. Y., Schunn, C., Nokes-Malach, T., & Singh, C. (2017). A longitudinal analysis of students’ motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics*, cjp-2017-0185. https://doi.org/10.1139/cjp-2017-0185

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, *37*(2), 225–238. https://doi.org/10.1016/0001-8791(90)90042-Z

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*. https://doi.org/10.2307/257084

Matz, R. L., Koester, B. P., Fiorini, S., Grom, G., Shepard, L., Stangor, C. G., … McKay, T. A. (2017). Patterns of Gendered Performance Differences in Large Introductory Courses at Five Research Universities. *AERA Open*, *3*(4), 1–12. https://doi.org/10.1177/2332858417743754

Mazure, C. M., & Jones, D. P. (2015). Twenty years and still counting: including women as participants and studying sex and gender in biomedical research. *BMC Women's Health*, *15*(1), 94. https://doi.org/10.1186/s12905-015-0251-9

McMahon, S. D., & Wernsman, J. (2008). The Relation of Classroom Environment and School Belonging to Academic Self-Efficacy among Urban Fourth- and Fifth-Grade Students. *The Elementary School Journal*, *109*(3), 267–281. https://doi.org/10.1086/592307

Meltzer, D. E. (2002). The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores. *American Journal of Physics*, *70*(12), 1259–1268. https://doi.org/10.1119/1.1514215

Mervis, J. (2012). Training and workforce. What if the science pipeline isn't really leaking? *Science (New York, N.Y.)*, *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*. Northwestern University.

Miller, J. D., & Kimmel, L. G. (2010). *Pathways to a STEMM Career*.

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., Blessing, J. S., & 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

Milsom, A., & Coughlin, J. (2016). Satisfaction With College Major: A Grounded Theory Study. *NACADA Journal*, *35*(2), 5–14. https://doi.org/10.12930/nacada-14-026

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

Muller, D., & Kase, N. (2010). Challenging traditional premedical requirements as predictors of success in medical school: The mount sinai school of medicine humanities and medicine program. *Academic Medicine*, *85*(8), 1378–1383. https://doi.org/10.1097/ACM.0b013e3181dbf22a

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, *38*(1), 30–38. https://doi.org/10.1037/0022-0167.38.1.30

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

Muthen, B., & Satorra, A. (1995). Complex Sample Data in Structural Equation Modeling. *Sociological Methodology*, *25*, 267–316. Retrieved from papers2://publication/uuid/4D3BFE88-D03F-48BF-8654-9F31943188B0

National Center for Science and Engineering Statistics. (2017). *Women, minorities, and persons with disabilities in science and engineering*. *Special Report NSF 17-310*. https://doi.org/Special Report NSF 17-310

National Science Board. (2015). *Revisiting the STEM Workforce, A Companion to Science and Engineering Indicators 2014*.

National Science Board. (2018). *Science and Engineering Indicators 2018*. *NSB-2018-1*. Alexandria, VA. https://doi.org/10.1016/0040-1625(91)90008-4

National Science Foundation. (2011). Women, Minorities, and Persons with Disabilities in Science and Engineering. *Special Report NSF 15-311*. Retrieved from http://www.nsf.gov/statistics/wmpd/

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

New American Economy Research Fund. (2017). *Sizing Up the Gap in our Supply of STEM Workers: Examining Job Postings and Unemployment Data from 2010-2016*. Retrieved from http://research.newamericaneconomy.org/report/sizing-up-the-gap-in-our-supply-of-stemworkers/

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

Nissen, J. M. (2019). Gender differences in self-efficacy states in high school physics. *Physical Review Physics Education Research*, *15*(1), 13102. https://doi.org/10.1103/PhysRevPhysEducRes.15.013102

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Oh, S. S., & Lewis, G. B. (2011). Stemming inequality? Employment and pay of female and minority scientists and engineers. *Social Science Journal*, *48*(2), 397–403. https://doi.org/10.1016/j.soscij.2010.11.008

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#

Page, S. (2007). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton, NJ: Princeton University Press.

Pajares, F., Johnson, M. J., & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*, *42*(1), 104–120.

PCAST. (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Sicence, Technology, Engineering and Mathemathics.* Washington, DC.

Peng, C. J., So, T. H., Stage, F. K., & St. John, E. P. (2002). The Use and Interpretation of Logistic Regression in Higher Education Journals : 1988-1999. *Research in Higher Education*, *43*(3), 259–293.

Penuel, W. R., & Fishman, B. J. (2012). Large-scale science education intervention research we can use. *Journal of Research in Science Teaching*, *49*(3), 281–304. https://doi.org/10.1002/tea.21001

Perkins, R., Kleiner, B., Roey, S., & Brown, J. (2004). *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990-2000*. *National Center for Education Statistics*. Retrieved from https://nces.ed.gov/pubs2004/2004455.pdf

Reed, V., & Buddeberg-Fischer, B. (2001). Career obstacles for women in medicine: An overview. *Medical Education*, *35*(2), 139–147. https://doi.org/10.1046/j.1365-2923.2001.00837.x

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

Riska, E. (2011). Gender and medical careers. *Maturitas*, *68*(3), 264–267. https://doi.org/10.1016/j.maturitas.2010.09.010

Rodriguez, A. J. (2001). From Gap Gazing to Promising Cases: Moving toward Equity in Urban Education Reform. *Journal of Research in Science Teaching*, *38*(10), 1115–1129. https://doi.org/10.1002/tea.10005

Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the School Psychological Environment and Early Adolescents' Psychological and Behavioral Functioning in School: The Mediating Role of Goals and Belonging. *Journal of Educational Psychology*, *88*(3), 408–422. https://doi.org/10.1037/0022-0663.88.3.408

Rosenthal, L., Levy, S. R., London, B., Lobel, M., & Bazile, C. (2013). In Pursuit of the MD: The Impact of Role Models, Identity Compatibility, and Belonging Among Undergraduate Women. *Sex Roles*, *68*(7–8), 464–473. https://doi.org/10.1007/s11199-012-0257-9

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modelingle. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ryan, C. L., & Bauman, K. (2016). *Educational attainment in the United States: 2015 population characteristics* (No. P20-578). *United States Census Bureau* (Vol. March). Retrieved from https://www.census.gov/content/dam/Census/library/publications/2016/demo/p20-578.pdf

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

Sanfey, H. A. (2006). Influences on Medical Student Career Choice. *Archives of Surgery*, *141*(11), 1086. https://doi.org/10.1001/archsurg.141.11.1086

Savinainen, A., & Scott, P. (2002). The Force Concept Inventory: a tool for monitoring student learning. *Physics Education*, *37*(1), 45–52. https://doi.org/10.1088/0031-9120/37/1/306

Sawtelle, V., Brewe, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching*, *49*(9), 1096–1121. https://doi.org/10.1002/tea.21050

Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating Causal Effects: Using Experimental and Observational Designs*. Washington, DC, DC: American Educational Research Association.

Schunk, D. H., & Meece, J. L. (2006). Self-efficacy development in adolescence. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (V, pp. 71–96). Greewich, CT, CT: Information Age Publishing. https://doi.org/10.1017/CBO9781107415324.004

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. *Development of Achievement Motivation*, *1446*, 15–31. https://doi.org/10.1016/b978-012750053-9/50003-6

Semsar, K., Knight, J. K., Birol, G., & Smith, M. K. (2011). The colorado learning attitudes about science survey (class) for use in biology. *CBE Life Sciences Education*, *10*(3), 268–278. https://doi.org/10.1187/cbe.10-10-0133

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

Seymour, E., & Hewitt, N. M. (1996). *Talking About Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO, CO: Westview Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company.

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

Sherman, D. K., Hartson, K. a., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., … Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, *104*(4), 591–618. https://doi.org/10.1037/a0031495

Siriwardena, A. N., Irish, B., Asghar, Z. B., Dixon, H., Milne, P., Neden, C., … Blow, C. (2012). Comparing performance among male and female candidates in sex-specific clinical knowledge in the MRCGP. *British Journal of General Practice*, *62*(599), 446–450. https://doi.org/10.3399/bjgp12X649142

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

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When Trying Hard Isn't Natural. *Personality and Social Psychology Bulletin*, *39*(2), 131–143. https://doi.org/10.1177/0146167212468332

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*(3), 371–384.

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

Spillane, J. P., Reiser, B. J., & Reimer, T. (2007). Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research*, *72*(3), 387–431. https://doi.org/10.3102/00346543072003387

Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Performance of African Americans. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.69.5.797

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

Stevens, T., Wang, K., Olivárez, A., & Hamman, D. (2007). Use of self-perspectives and their sources to predict the mathematics enrollment intentions of girls and boys. *Sex Roles*, *56*(5–6), 351–363. https://doi.org/10.1007/s11199-006-9180-2

Stout, J. G., Ito, T. A., Finkelstein, N. D., & Pollock, S. J. (2013). How a gender gap in belonging contributes to the gender gap in physics participation. In *AIP Conference Proceedings* (pp. 402–405). AIP.

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I Fit in and Do Well? The Importance of Social Belongingness and Self-Efficacy for Explaining Gender Differences in Interest in STEM and HEED Majors. *Sex Roles*, *77*(1–2), 86–96. https://doi.org/10.1007/s11199-016-0694-y

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

Trujillo, G., & Tanner, K. D. (2014). Considering the Role of Affect in Learning: Monitoring Students' Self-Efficacy, Sense of Belonging, and Science Identity. *CBE—Life Sciences Education*, *13*(1), 6–15. https://doi.org/10.1187/cbe.13-12-0241

Tsugawa, Y., Jena, A. B., Figueroa, J. F., Orav, E. J., Blumenthal, D. M., & Jha, A. K. (2017). Comparison of hospital mortality and readmission rates for medicare patients treated by male vs female physicians. *JAMA Internal Medicine*, *177*(2), 206–213. https://doi.org/10.1001/jamainternmed.2016.7875

U.S. Bureau of Labor Statistics. (2017a). *Employment Projections: Occupational projections and worker characteristics*. Retrieved from https://www.bls.gov/emp/ep_table_107.htm

U.S. Bureau of Labor Statistics. (2017b). *Women in the Labor Force : A Databook* (No. Rep. No. 1065, BLS). Retrieved from http://www.bls.gov/cps/wlf-databook-2010.pdf

Usher, E. L., & Pajares, F. (2008). Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions. *Review of Educational Research*, *78*(4), 751–796. https://doi.org/10.3102/0034654308321456

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

Vincent-Ruz, P., & Schunn, C. D. (2017). The increasingly important role of science competency beliefs for science learning in girls. *Journal of Research in Science Teaching*, *54*(6), 790–822. https://doi.org/10.1002/tea.21387

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96. https://doi.org/10.1037/0022-3514.92.1.82

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two Brief Interventions to Mitigate a "Chilly Climate" Transform Women's Experience, Relationships, and Achievement in Engineering. *Journal of Educational Psychology*, *107*(2), 468–485. https://doi.org/10.1037/a0037461

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

Wang, X. (2012). *Modeling Student Choice of STEM Fields of Study: Testing a Conceptual Framework of Motivation , High School Learning , and Postsecondary Context of Support*. Wisconsin Center for the Advancement of Postsecondary Education, University of Wisconsin - Madison.

Watt, H. M. G., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls. *Equal Opportunities International*, *25*(8), 642–659. https://doi.org/10.1108/02610150610719119

Weisgram, E. S., & Diekman, A. B. (2017). Making STEM "Family Friendly": The Impact of Perceiving Science Careers as Family-Compatible. *Social Sciences*, *6*(2), 61. https://doi.org/10.3390/socsci6020061

Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Williams, D. R. (2008). Racial/ethnic variations in women's health: the social embeddedness of health. *American Journal of Public Health*, *98*(4), S38-47. https://doi.org/10.2105/AJPH.98.Supplement_1.S38

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), 1–12. https://doi.org/10.3102/0013189X19840331

Xie, Y., & Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA, MA: Harvard University Press.

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

Zeldin, A. L., & Pajares, F. (2000). Against the Odds: Self-Efficacy Beliefs of Women in Mathematical, Scientific, and Technological Careers. *American Educational Research*

Journal, *37*(1), 215–246. https://doi.org/10.3102/00028312037001215

Zohar, A., & Bronshtein, B. (2005). Physics teachers' knowledge and beliefs regarding girls' low participation rates in advanced physics classes. *International Journal of Science Education*, *27*(1), 61–77. https://doi.org/10.1080/0950069032000138798