

THE UNDERREPRESENTATION OF WOMEN IN SCIENCE: INTERNATIONAL AND CROSS-DISCIPLINARY EVIDENCE AND DEBATE

EDITED BY : Stephen J. Ceci, Wendy M. Williams and Shulamit Kahn
PUBLISHED IN: Frontiers in Psychology





frontiers

Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-434-1

DOI 10.3389/978-2-88945-434-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

THE UNDERREPRESENTATION OF WOMEN IN SCIENCE: INTERNATIONAL AND CROSS-DISCIPLINARY EVIDENCE AND DEBATE

Topic Editors:

Stephen J. Ceci, Cornell University, United States

Wendy M. Williams, Cornell University, United States

Shulamit Kahn, Boston University, United States



"Woman teaching geometry" attributed to Meliacin Master (fl. 13th century–1312). This file has been provided by the British Library from its digital collections.

There is no shortage of articles and books exploring women's underrepresentation in science. Everyone is interested--academics, politicians, parents, high school girls (and boys), women in search of college majors, administrators working to accommodate women's educational interests; the list goes on. But one thing often missing is an evidence-based examination of the problem, uninfluenced by personal opinions, accounts of "lived experiences," anecdotes, and the always-encroaching inputs of popular culture. This is why this special issue of *Frontiers in Psychology* can make a difference. In it, a diverse group of authors and researchers with even more diverse viewpoints find themselves united by their empirical, objective approaches to understanding women's underrepresentation in science today.

The questions considered within this special issue span academic disciplines, methods, levels of analysis, and nature of analysis; what these articles share is their scholarly, evidence-based approach to understanding a key issue of our time.

Citation: Ceci, S. J., Williams, W. M., Kahn, S., eds. (2018). *The Underrepresentation of Women in Science: International and Cross-Disciplinary Evidence and Debate*. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-434-1

Table of Contents

- 06 Editorial: Underrepresentation of Women in Science: International and Cross-Disciplinary Evidence and Debate**

Wendy M. Williams

Sexism in Professorial Hiring

- 09 Women Have Substantial Advantage in STEM Faculty Hiring, Except When Competing Against More-Accomplished Men**

Stephen J. Ceci and Wendy M. Williams

- 19 Leaky Pipeline Myths: In Search of Gender Effects on the Job Market and Early Career Publishing in Philosophy**

Sean Allen-Hermanson

Exploring the Gender Gap Via National Datasets

- 29 Are Recent Cohorts of Women With Engineering Bachelors Less Likely to Stay in Engineering?**

Shulamit Kahn and Donna K. Ginther

- 44 All STEM Fields are Not Created Equal: People and Things Interests Explain Gender Disparities Across STEM Fields**

Rong Su and James Rounds

- 64 Math Achievement is Important, But Task Values are Critical, Too: Examining the Intellectual and Motivational Factors Leading to Gender Disparities in STEM Careers**

Ming-Te Wang, Jessica Degol and Feifei Ye

Stereotypes About Brilliance and Male-Oriented Fields

- 73 Women are Underrepresented in Fields Where Success is Believed to Require Brilliance**

Meredith Meyer, Andrei Cimpian and Sarah-Jane Leslie

- 85 Cultural Stereotypes as Gatekeepers: Increasing Girls' Interest in Computer Science and Engineering by Diversifying Stereotypes**

Sapna Cheryan, Allison Master and Andrew N. Meltzoff

- 93 On the Gender–Science Stereotypes Held by Scientists: Explicit Accord With Gender-Ratios, Implicit Accord With Scientific Identity**

Frederick L. Smyth and Brian A. Nosek

- 112 The Bachelor's to Ph.D. STEM Pipeline No Longer Leaks More Women Than Men: a 30-Year Analysis**

David I. Miller and Jonathan Wai

Importance of Competitive Schools and Perceived Math Ability

122 *The Role of School Performance in Narrowing Gender Gaps in the Formation of STEM Aspirations: a Cross-National Study*

Allison Mann, Joscha Legewie and Thomas A. DiPrete

133 *Perceived Mathematical Ability Under Challenge: A Longitudinal Perspective on Sex Segregation Among STEM Degree Fields*

Samantha Nix, Lara Perez-Felkner and Kirby Thomas

Wisdom from the Trenches of Academia

152 *Does Gender of Administrator Matter? National Study Explores U.S. University Administrators' Attitudes About Retaining Women Professors in STEM*

Wendy M. Williams, Agrima Mahajan, Felix Thoemmes, Susan M. Barnett,
Francoise Vermeylen, Brian M. Cash and Stephen J. Ceci



Editorial: Underrepresentation of Women in Science: International and Cross-Disciplinary Evidence and Debate

Wendy M. Williams*

Department of Human Development, Cornell University, Ithaca, NY, United States

Keywords: women in science, underrepresentation of women, women in STEM, STEM careers, work-life balance

Editorial on the Research Topic

Underrepresentation of Women in Science: International and Cross-Disciplinary Evidence and Debate

There is no shortage of articles and books exploring women's underrepresentation in science. Everyone is interested—academics, politicians, parents, high school girls (and boys), women in search of college majors, administrators working to accommodate women's educational interests; the list goes on. But one thing often missing is an evidence-based examination of the problem, uninfluenced by personal opinions, accounts of “lived experiences,” anecdotes, and the always-encroaching inputs of popular culture. This is why this special issue of *Frontiers in Psychology* can make a difference. In it, a diverse group of authors and researchers with even more diverse viewpoints find themselves united by their empirical, objective approaches to understanding women's underrepresentation in science today.

OPEN ACCESS

Edited and reviewed by:

Jessica S. Horst,
University of Sussex, United Kingdom

*Correspondence:

Wendy M. Williams
wendywilliams@cornell.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 October 2017

Accepted: 22 December 2017

Published: 22 January 2018

Citation:

Williams WM (2018) Editorial:
Underrepresentation of Women in
Science: International and
Cross-Disciplinary Evidence and
Debate. *Front. Psychol.* 8:2352.
doi: 10.3389/fpsyg.2017.02352

OVERVIEW OF ARTICLES IN SPECIAL ISSUE

The questions considered within this special issue span academic disciplines, methods, levels of analysis, and nature of analysis; what these articles share is their scholarly, evidence-based approach to understanding a key issue of our time.

Sexism in Professorial Hiring

Ceci and Williams re-visited the experimental paradigm from their 2015 *Proceedings of the National Academy of Sciences* article in which (in four of their five experiments) faculty were asked to rate three short-listed finalists for a tenure-track position. The 2015 study revealed a 2:1 preference for women when finalists were equivalently excellent. The new study contrasted a male finalist who was slightly superior to the female finalist. Women's advantage vanished when the male applicant was depicted as slightly stronger, suggesting that fears that affirmative action goals will undermine hiring of most-qualified applicants are unfounded.

Allen-Hermanson examined an overlooked aspect of the women's underrepresentation debate: Are philosophers prejudiced against hiring women applicants despite professing conscious, explicit egalitarian beliefs? Unlike other humanities departments, philosophy departments have far fewer women professors than might be expected. He reviewed several recent data sets demonstrating that female applicants are favored when it comes to tenure-track hiring in philosophy departments.

Exploring the Gender Gap via National Datasets

Using 1993–2010 nationally representative data, Kahn and Ginther examined whether the gender gap in engineering has narrowed recently. They discovered that the majority of the gender retention gap was due to women leaving the labor force coincident with child-bearing. There was no gender retention difference by 7–8 years post-bachelors for those full-time employed; single childless women were more likely than men to remain in engineering than were single childless men, and women who left engineering entirely were just as likely as men who left to remain in math-intensive fields. Their findings caution against past assertions that women do not persist in STEM fields as long as men.

In their latest meta-analysis, Su and Rounds examined data from 52 samples entailing over 430,000 respondents between 1964 and 2007. Gender differences in interests favoring males were largest in engineering-related fields, and favored women in allied health fields and social sciences. This adds to the large body of empirical findings that have revealed similar sex differences along the people-thing dimension.

Miller and Wai reported the results of their analysis of longitudinal data to examine the baccalaureate-to-PhD transition. In contrast to the traditional leaky pipeline metaphor, they found that over time, women have segued from the baccalaureate to PhD programs in increasing numbers. Their work suggests that researchers and policy makers need to look elsewhere for causes of women's underrepresentation.

Wang et al. studied factors predicting gender differences in selection of STEM occupations, and whether math task values and altruism mediate the pathway through which gender affects STEM career choice through math achievement. Based on longitudinal analyses, they found that the association between gender and working in a STEM career by one's early- to mid-thirties was mediated by math achievement scores in twelfth grade; females did more poorly on standardized math tests than did males.

Stereotypes about “Brilliance” and “Male-Oriented” Fields

Meyer et al. examined field-specific beliefs regarding the importance of brilliance. They provide support for the hypothesis that women are most likely to be underrepresented in fields that members believe require raw intellectual talent, which women are stereotyped to possess less of than do men. The beliefs of participants with exposure to a field predicted the magnitude of the field's gender gap, independent of their beliefs about the level of mathematical ability required. Their findings are consistent with female high school students taking fewer AP courses in all areas of science except biology (Ceci et al., 2014).

Cheryan et al. presented data and argument showing that modern American culture stereotypes as male-oriented those fields that involve social isolation, an intense focus on machinery,

and inborn brilliance. These stereotypes are compatible with qualities that are typically more valued in men than women in American culture. Their work continues their recent insights and is consistent with the findings of some of the other contributors, particularly Meyer et al.

Smyth and Nosek explored whether variation in female representation across scientific disciplines is associated with differences in the strength of gender-science stereotypes, explicit and implicit, held by men and women in these fields. For explicit stereotypes that associate science with “male,” the strength of stereotyping varied across scientific disciplines as a function of gender ratios in the disciplines; however, implicit stereotypes did not vary as a function of such ratios. Giving currency to their findings is recent evidence that children continue to associate science with being male (Miller et al., in press).

Importance of Competitive Schools and Perceived Math Ability

Mann et al. analyzed findings from PISA data from 55 countries. They placed schools along a continuum from most to least competitive, based on average math and science performance. Schools that are most competitive are often associated with stronger math-science environments. The authors found that the aspirations gender gap narrowed for high-performing students in stronger performance environments.

Nix et al. reported an analysis of longitudinal, nationally representative high school data. They found that perceived mathematics ability when under challenge predicted important outcomes such as taking advanced science courses in high school, and that high school men scored higher than women did in their perceived ability under mathematics challenge. Their findings are consistent with female high school students taking fewer AP courses in all areas of science except biology (Ceci et al., 2014).

Wisdom from the Trenches of Academia

Finally, Williams et al. collected and analyzed an original national empirical dataset in which provosts, deans, associate deans, and department chairs of STEM fields at 96 U.S. research-intensive universities rated the quality and feasibility of strategies for retaining women in STEM fields. For example, administrators agreed that gender quotas were a weak idea, and that campus childcare centers were an excellent idea. Women administrators were more supportive than were men of shared tenure lines, and saw it as more feasible for men to stop the tenure clock for 1 year for childrearing.

In sum, readers will find multiple perspectives in this special issue, and the editors hope it will stimulate new directions of thinking and scholarship on women and science.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

REFERENCES

- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Miller, D. I., Nolla, K. M., Eagly, A. H., and Uttal, D. H. (in press). The development of children's gender-science stereotypes: a meta-analysis of five decades of U.S. Draw-A-Scientist studies. *Child Dev.*

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Women have substantial advantage in STEM faculty hiring, except when competing against more-accomplished men

Stephen J. Ceci* and Wendy M. Williams

Department of Human Development, Cornell University, Ithaca, NY, USA

OPEN ACCESS

Edited by:

Steven E. Mock,
University of Waterloo, Canada

Reviewed by:

David I. Miller,
Northwestern University, USA
Amy Yeung,
University of Waterloo, Canada

*Correspondence:

Stephen J. Ceci,
Department of Human Development,
Cornell University, G80 MVR Hall,
Ithaca, NY 14853, USA
sjc9@cornell.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 14 April 2015

Accepted: 22 September 2015

Published: 20 October 2015

Citation:

Ceci SJ and Williams WM (2015)
Women have substantial advantage in
STEM faculty hiring, except when
competing against
more-accomplished men.
Front. Psychol. 6:1532.
doi: 10.3389/fpsyg.2015.01532

Audits of tenure-track hiring reveal faculty prefer to hire female applicants over males. However, audit data do not control for applicant quality, allowing some to argue women are hired at higher rates because they are more qualified. To test this, Williams and Ceci (2015) conducted an experiment demonstrating a preference for hiring women over identically-qualified men. While their findings are consistent with audits, they raise the specter that faculty may prefer women over even more-qualified men, a claim made recently. We evaluated this claim in the present study: 158 faculty ranked two men and one woman for a tenure-track-assistant professorship, and 94 faculty ranked two women and one man. In the former condition, the female applicant was slightly weaker than her two male competitors, although still strong; in the other condition the male applicant was slightly weaker than his two female competitors, although still strong. Faculty of both genders and in all fields preferred the more-qualified men over the slightly-less-qualified women, and they also preferred the stronger women over the slightly-less-qualified man. This suggests that preference for women among identically-qualified applicants found in experimental studies and in audits does not extend to women whose credentials are even slightly weaker than male counterparts. Thus these data give no support to the twin claims that weaker males are chosen over stronger females or weaker females are hired over stronger males.

Keywords: affirmative action, women in science, bias, sexism, academic hiring

Introduction

Much has been written about the campaign to diversify faculty at American colleges and universities, an effort that started in earnest during the 1980s and continues unabated. To this end, hundreds of analyses of faculty hiring for tenure-track positions have been reported, and the temporal changes in the fraction of female and minority applicants in the American professoriate have been charted (e.g., Smith et al., 2004; Kang and Banaji, 2006; Turner et al., 2008; Niederle et al., 2013). Despite substantial gains in diversity of faculty, the dominant view appears to be that racial and gender preferences continue to be needed to counter not just historical prejudice but also current biases held by faculty—most of which may be implicit, and which result in barriers against hiring women and minorities. It is alleged that such biases create, in the words of Kang and Banaji, “threats to fair treatment—threats that lie in every mind” and that affirmative-action hiring

programs should be continued until data are available to indicate such threats are over: “such data should be a crucial guide to ending affirmative action” (Kang and Banaji, 2006, p. 1063). With some notable exceptions demonstrating female-friendly hiring preferences by faculty (Williams and Ceci, 2015), there continues to be evidence of implicit and occasionally explicit biases directed at women and ethnic minorities. Although few of these demonstrations of bias concern hiring of academic science faculty, some of them are indirectly relevant. The present experiment was undertaken to determine whether gender differences trump applicant quality in tenure-track hiring decisions.

Stereotypes, Hiring Bias, and Gender Congruity

A growing literature reveals people are apt to explicitly associate science with men, including not only students but also scientists (Smyth and Nosek, 2015), and that such stereotypes are pervasive, as shown recently by Miller et al. (2015). In their transnational analysis, Miller et al. showed higher female enrollment in post-secondary course-taking in nations with weaker implicit and explicit gendered stereotypes regarding science. Such stereotypes can lead to biased evaluations against women in so-called gender-incongruous contexts, such as in STEM fields in which men have historically been dominant (engineering, physics, economics, computer science, geosciences, and mathematics). This form of bias is particularly likely to emerge when information about applicants' competence is unavailable or when the evaluators are not experienced professionals. For example, Ernesto Reuben and his colleagues (Reuben et al., 2014) asked nearly 150 men and women (mostly undergraduates) to add a string of four 2-digit numbers. They were given 4 min to do as many additions as possible. The authors then assigned the role of hiring manager to nearly 200 male and female students, who were asked to decide whom among these 150 students to hire. Afterward these managers were given an implicit bias test. The authors found that men were hired at twice the rate of women; most of the students playing the role of hiring manager believed men were better at math and science. Even when informed of superior arithmetic scores by women, some hypothetical managers continued to prefer to hire men. In their “cheap talk” condition (which had the largest gender bias), applicants selected the lower performing male over the higher performing female in 29% of the cases compared to selecting the lower performing female over the higher performing female in only 2% of the cases. Hence, in that study, the pro-male bias trumped even applicant quality. Taken together, these transnational and experimental studies indicate that implicit biases and sometimes explicit ones can lead to fewer women preparing for a career in STEM and ultimately being hired.

Studies of gender biases suggest that stereotypes are not always activated but rather are invoked when information about applicants is limited or ambiguous or when evaluators lack motivation to be careful. In such situations stereotypes can reduce cognitive load during decision-making. However, relying upon stereotypes may be unnecessary when information about applicants indicates unambiguously high competence, as in the

case with tenure-track hiring. In their recent metaanalysis, Amanda Koch and her associates found that gender-role congruity bias was largest when so-called “individuating information” that was informative of applicants' competence was ambiguous or not clearly diagnostic of success. They reported that sex bias shrinks in male-dominated fields when diagnostic information about applicants' competence is available (see Koch et al., 2015, pp. 130–131). The authors reported near-zero bias when female applicants were evaluated by experienced professionals in male-dominated fields if information regarding their competence was available ($d = 0.02$). This finding is relevant to the Reuben et al. hiring manager study because the evaluation was not done by experienced professionals, and in one version student managers were significantly less biased when they were supplied with the women's actual arithmetic scores, albeit some smaller number of student managers still exhibited a male hiring bias even in this condition. Typically, however, the studies in the metaanalysis examined applicants with equally strong records, thus telling us little about whether bias occurs for female applicants possessing inferior credentials, as some have alleged (e.g., Niederle et al., 2013). We directly address this lacuna in the current experiment.

Relatedly, Moss-Racusin et al. (2012) found that both male and female faculty preferred to hire, remunerate, and mentor a male applicant for a lab manager post than an identically-qualified female applicant. However, the lab manager post was baccalaureate-level and the lab manager applicant was depicted as ambiguously-competent rather than as unambiguously stellar. Thus, the Reuben et al. and Moss-Racusin et al. experiments leave unanswered the question of whether such bias would be found in hiring applicants for professorships under conditions in which experienced faculty have motivation to be careful and possess diagnostic information about applicants' competence—in other words, the real-world conditions under which faculty are usually hired.

Background for the Present Study

In the present study we report findings from an ongoing program of experimental research aimed at examining biases in the hiring of women scientists in male-dominated fields in the academy. The major focal question in the current study is: How much do gender-related biases trump preferences for the candidate with the highest quantitative competence index, based on publications, letters, interview, and job talk? Recent experimental evidence indicates that when evaluators are themselves experienced professionals, women applicants for professorships are preferred over equally-competent men when both are depicted realistically, as identically and unambiguously stellar (Williams and Ceci, 2015). Here we ask whether this preference for female applicants will extend to situations in which women are quantitatively slightly weaker than men.

Many blue-ribbon panels and national organizations argue for the continued use of preferential hiring programs because biased hiring is viewed as a cause of women's underrepresentation in academic science, by “inadvertently foreclosing consideration of the best-qualified persons by untested presuppositions which operate to exclude women and minorities” (AAUP, 2014).

Notwithstanding the recent pro-female hiring data of Williams and Ceci, there are recent empirical data implying that hiring is sexist and that it possibly forecloses the prospects of the best-qualified female applicants. However, none of these data concern the hiring of academic science faculty by professionals who possess diagnostic information, but they nevertheless are relevant. Below we describe a survey study and an experiment that are relevant to gender bias in academic hiring, even though neither actually involves hiring of professors in male-dominated fields.

Sheltzer and Smith (2014) surveyed biology department web pages and departmental directories to ascertain the numbers of graduate students and postdoctoral researchers employed by faculty members. They found that elite male faculty (winners of lifetime awards, members of the National Academy of Sciences, recipients of funding by the Howard Hughes Medical Institute) employed fewer female graduate students and postdoctoral researchers than did elite female faculty, who did not exhibit a gender asymmetry. New assistant professors in biology were disproportionately comprised of individuals who came from these elite laboratories, which had an overabundance of male grad students and postdocs, thus reflecting a seeming causal loop. However, two features of this study merit mention: first, biology is a field in which women are well represented among both PhD recipients and among the professoriate, so it is unlikely to be the ideal field in which to detect gender bias. Second, because this was not an experiment, it leaves open alternative explanations for the observed gender asymmetry, such as whether female postdocs self-selected (i.e., were more likely to apply to work with female faculty). Despite these concerns, the findings are suggestive of a male faculty bias in recruiting and appointing postdocs that can eventuate in more male professors being hired, despite the fact that biology is a field that appears to be female-friendly.

There is one experiment in the last 30 years that has addressed the question of sex bias in the hiring of professors; it was conducted by Steinpreiss and her associates 16 years ago (Steinpreiss et al., 1999). They found faculty of both genders preferred to hire the male applicant over the identically-qualified female applicant. However, there are two features of this experiment that limit its applicability: first, it examined bias in only one field, psychology, which is the field in which women are best represented—psychology has the largest fraction of women professors of all STEM fields, constituting the majority of faculty. Second, Steinpreiss et al. did not find a preference for hiring a man over a woman when the hypothetical applicants were depicted as unambiguously stellar senior faculty applicants (considered for early tenure). The reason these points are noteworthy is that Koch et al.'s metaanalysis found small-to-moderate sex bias in male-dominated jobs when applicants had average or ambiguous competence ($d = 0.29$) but, as noted above, no bias when applicants had high competence ($d = 0.02$) or when evaluators were motivated to be careful ($d = 0.01$), both conditions that characterize tenure-track hiring. For hiring tenure-track professors in male-dominated fields such as engineering, physics, and economics, experienced professionals might be expected to exhibit little or no sex bias when evaluating

applicants who are unambiguously competent. Finally, some evidence suggests that an implicit stereotypic association of race with violence in a videogame simulation did not lead to racist behavior when participants held relatively high implicit negative attitudes toward prejudice (Glaser and Knowles, 2008). This suggests that motivation against possessing or demonstrating bias influences behavior and attitudes of even those possessing implicit biases.

In contrast to experiments showing hiring bias, Williams and Ceci (2015) reviewed 8 large-scale audits of actual hiring that indicate women are preferred for tenure-track hiring in the real world. For example in a large National Research Council (NRC) (2009) analysis, women were hired at rates higher than their application numbers in every field assessed at the 89 research universities the NRC panel studied: in mathematics, women constituted 20% of applicants but 32% of hires; in electrical engineering women were 11% of applicants but 32% of hires; in chemistry women were 18% of applicants but 29% of hires; and in physics they were 12% of applicants but 20% of hires. Similar pro-female hiring data were reported in the National Computer Research Association hiring report for professorships in computer science: “as new PhDs, women submitted far fewer applications than men but received many more offers per application. Female new hires applied for only 6 positions (compared with 25 for men), obtained 0.77 interviews per application (vs. 0.37 for men), and received 0.55 offers per application (vs. 0.19 for men). Obviously women were much more selective in where they applied, and also much more successful in the application process.” Against this backdrop of actual hiring data showing a preference for female applicants, a goal of this program of research has been to determine whether this hiring advantage occurs because women applicants are more qualified than men. Williams and Ceci (2015) showed in their experiments this is not what is driving the female hiring preference because women applicants continue to be preferred over male applicants who are equally qualified. This is in contrast to frequent claims to the contrary.

Present Study

In a recent series of experiments, Williams and Ceci (2015) asked a nationally stratified sample of 873 faculty from four academic fields (economics, psychology, biology, and engineering) to rank two otherwise identically-qualified hypothetical finalists for a tenure-track assistant professorship in their department. These identically-qualified finalists were referred to as Dr. X and Dr. Z and they were presented to faculty with identical quantitative ratings of their candidacy based on their research, job talk, letters, and interview; the sole difference between them was their gender. Faculty were informed that Dr. X and Dr. Z were both rated 9.5 by their departmental colleagues on the basis of their publications, interview, letters, and meetings, where 10.0 = outstanding/exceptional and 1 = cannot support for tenure-track hiring. Thus, Drs. X and Z were depicted as unambiguously strong applicants, which is realistic for tenure-track applicants who have made it to the short list of finalists in searches that often generate hundreds of PhD

applicants.¹ Faculty preferred to hire the female 2-to-1 over her identically-qualified male counterpart. This strong pro-female bias was found in all four fields and by faculty of both genders with the exception of male economists who showed no preference between equivalently-qualified female and male applicants. Because of its stratified national sampling and use of sampling weights, the Williams and Ceci (2015) findings were representative of the size of the ratio at all types of institutions, from small teaching-intensive colleges to large, research-intensive ones.

There were two features in Williams and Ceci's experimental design that were implemented to obscure the true purpose of their experiment, one of which is relevant in the present context. To obscure the true nature of their hypothesis so that faculty would not realize they were being assessed to determine whether they harbored sexist biases in hiring, Williams and Ceci disguised the study to appear as a competition between different personalities. In actuality the personalities were counterbalanced with gender and varied in a between-subjects design. In addition to the use of this personality disguise, there was another ploy used to minimize faculty respondents' awareness; it was the addition of a third applicant, a foil. In addition to pitting an equally-qualified Dr. X against a Dr. Z, Williams and Ceci added a third short-listed competitor who was pretested to be slightly inferior to X and Z, labeled Dr. Y. Unlike Drs. X and Z who were both given quantitative scores of 9.5, Dr. Y was given 9.3, which although still very strong is slightly inferior. In the Methods section we describe this feature in more detail because it is a central aspect of the present study. Thus, the inclusion of these two features—a slightly lower-rated foil (Dr. Y) and the counterbalanced adjectives—served to disguise the true purpose of the experiment. And the misdirection appeared to work: A survey of 30 faculty in their study reported no suspicion that the experiment had to do with gender preference in hiring.

Summing across numerous analyses, Williams and Ceci reported the odds of preferring a woman over an identically-qualified man was roughly 2-to-1. Importantly for the purpose of the present experiment, only 2.53% of faculty preferred to hire Dr. Y over his slightly stronger competitors, Drs. X and Z. In a subsequent experiment that excluded the Dr. Y foil, these researchers asked faculty to rate only one applicant (either a female or male finalist), to avoid implicit competition between a woman and man. Faculty assigned their own quantitative scores to the applicant they were sent to evaluate. Again, there was

a preference for women, with faculty of both genders giving the female candidate a higher quantitative score than other faculty gave the identically-qualified male candidate. This latter finding suggests that faculty have internalized the norm of gender diversity and were not merely responding in a manner that is politically correct or to exhibit some other form of impression-management, because faculty had no knowledge that other faculty were evaluating the identical accomplishments in the form of an opposite-sex applicant.

These results raise an intriguing question regarding the pervasiveness of the preference for women: Would it still be observed if the Dr. Y foil was a woman instead of a man? If Dr. Y was a slightly less accomplished female finalist as compared to the two male finalists, would faculty still reject her—that is, would they still choose her only 2.53% of the time as was found when Dr. Y was a male? Or would the desire for gender diversity among faculty be sufficiently strong that they would prefer to hire a slightly less accomplished female Dr. Y over more accomplished male applicants? This is the question we attempt to answer in the current experiment. It will shed light on the extent of faculty's desire to diversify the academy: It is one thing to find that faculty of both genders prefer to hire a female applicant over her identically-qualified male counterpart by a ratio of 2-to-1, but it is another matter to ask whether this preference for female applicants extends to a preference to hire a slightly weaker female applicant, one described as 9.3 on a 10-point scale who is competing against two males who are described as 9.5.

Thus, the current experiment consisted of a comparison of a woman assigned a slightly lower quantitative score competing against two men assigned a slightly higher quantitative score, all of whom were competing for the same assistant professorship. We used the same 9.3 vs. 9.5 quantitative scores used by Williams and Ceci (2015) because their survey provides a national base-rate for faculty expectations for this contrast. If a preference is found for a female finalist depicted as 9.3 over men depicted 0.2 points higher, then subsequent contrasts between even lower-rated females would be in order. But first we sought evidence of preferential hiring of women who are only slightly weaker than their male competitors.

Methods

Participants

The pool of potential faculty participants was assembled by drawing a national stratified sample of 694 tenured/tenure-track professors (half female, across all ranks). This was done by randomly sampling from online directories for Carnegie Foundation's 3 Basic Classifications of: (a) Doctoral (combining all three levels of doctoral intensity), (b) Master's institutions (combining all three levels—small, medium, and large), and (c) Baccalaureate institutions (combining all three levels of such institutions). This sample of 694 professors was drawn equally from four popular fields, two math-intensive ones in which women faculty are greatly underrepresented— < 15% (engineering, economics)—and two non-math-intensive fields (biology, psychology) in which women faculty are well represented and are considered to have achieved what gender

¹For example, a faculty respondent in the field of biology in Williams and Ceci's experiment wrote: "In a typical search these days we will receive over 200 applications for one position. The search committee triages that down to a group of around 30 or 40, and then no more than around 6–8 are invited to come for a three-day visit and to give a seminar." Many similar comments were offered by others in their national survey, hence the finalists are usually unambiguously strong, as is true in our own department where a recent tenure-track search for an assistant professor generated 267 applicants in psychology. All applicants who survive to the short list are accomplished, having successfully completed doctorates, published papers, and garnered strong letters of recommendation. In a separate rating task we gave 35 faculty the CVs of actual short-listed candidates and asked them to rate these on a 10-point scale and, as expected, the mean rating was in the excellent range.

equity advocates regard as a critical mass, although even these fields still produce significantly more female PhDs than the female fraction of total professorships. There were two constraints in randomization. One was that for an institution to be included it had to have programs in at least three of the four fields. This was true of all doctoral institutions in the sampling frame, but it excluded many small colleges that lacked two or more of the four fields, and over half of the nation's combined master's programs. The second constraint was that only tenured or tenure-track faculty were included in the sample frame; off-line faculty (emeriti, adjuncts, lecturers, instructors, courtesy faculty members, and visiting professors) were excluded, as only faculty who actually vote on tenure-track hiring were desired as subjects.

Overall, out of the 694 faculty who were assigned to one of two conditions, 252 responded with full data (36.3%): 158 rated a male Dr. X who was pitted against a female Dr. Y and a male Dr. Z; and 94 rated a female Dr. X who was pitted against a male Dr. Y and a female Dr. Z.

Materials

Two sets of materials were used, the first containing profiles of two *male* applicants, Dr. X and Dr. Z, with identical scholarly qualitative scores but differing in gendered adjective descriptors ("kind, socially-skilled, creative" vs. "analytical, competitive, powerhouse"). As noted, these descriptors were used to disguise the actual hypothesis, leading raters to believe the research question was whether they preferred one type of individual over the other. These gendered descriptors were counterbalanced so that half the faculty received Dr. X portrayed as a male "analytical, competitive, powerhouse" competing against Dr. Z as a male "kind, socially-skilled, creative" colleague, and half received Dr. X and Z portrayed with the opposite terms. Dr. Y was described as "shy and reserved," which is more negative than "socially skilled" or a "real powerhouse," and in the chair's notes some concern was raised about Y's teaching performance, whereas no concern was raised for X or Z. Thus, the quantitative "pre-rankings" gave an explicit cue that Drs. X and Z were stronger than Dr. Y, albeit only slightly so (see Supplementary Material for one set of these materials). These different personae were the same as those used by Williams and Ceci (2015) and were based on gender congruity norms (Diekmann and Eagley, 2000; Cuddy et al., 2004). Notwithstanding this systematic variation of these descriptors between faculty raters, Drs. X and Z were otherwise identical: both were rated 9.5 out of 10.0 in quality on the basis of their scholarly accomplishments, job talk, and faculty meetings. This corresponded to "impressive."

In every contest between the male Drs. X and Z, a third candidate was added, a female Dr. Y, who was depicted as slightly lower in scholarly quality (9.3) than the male Drs. X and Z, and who was pretested with an independent group of faculty who did not participate in this experiment to ensure that raters perceived her quality as slightly lower. Dr. Y was always depicted in the same terms used by Williams and Ceci for their Dr. Y foil when he was a male, since it was established that under these conditions their male Dr. Y was chosen by only 2.53% of faculty in their large stratified national sample.

The second set of materials simply reversed the genders so that Drs. X and Z were depicted as women and Dr. Y as a man; everything else was identical.

Procedure

Thus, the contest presented to every faculty member was to choose between three finalists for a tenure-track position, in one condition with Drs. X and Z both being male candidates of equivalent quality (9.5) and Dr. Y being a slightly lower quality female candidate (9.3), and in the other condition with these genders reversed (see Supplementary Material for materials). Faculty members were sent personal emails containing one of the counterbalanced depictions, and were asked to rank these three finalists in order of their hiring preference: first, second, and third for a tenure-track assistant professorship in their own department. The question of interest is whether faculty exhibit preferential hiring for female applicants possessing slightly lower quantitative scores than their male counterparts.

Results

The main analysis examined which candidate was ranked first by faculty of each gender and at each type of college/university, and in each of four academic disciplines. In addition to the four disciplines (engineering, economics, psychology, biology) there were three types of colleges/universities based on the Carnegie classification (1 = doctoral, 2 = bachelors/masters, 3 = baccalaureate).

The response rates for every cell (university Carnegie type by discipline by gender, $3 \times 4 \times 2$) were evaluated in a logistic regression. Response rates for the 252 faculty across these 24 cells were unrelated to the findings. These data were analyzed with both unweighted and weighted logistic regression models to provide a stronger test on their representativeness. Here we report only the traditional unweighted analyses but the weighted results (weighted to account for differences in the numbers of men and women in the population and in the sampling frame) were highly similar, with no result changing.

Across the 158 contests between the equivalently strong male Drs. X and Z, only 7 faculty respondents preferred the slightly weaker female Dr. Y, and one faculty rater gave tied ranks for X and Y for first place. This resulted in an overall female Dr. Y-preference of 4.8%. In the condition in which 92 faculty were asked to choose between two slightly more accomplished women—Drs. X and Z—and a slightly less accomplished male Dr. Y, only 1 out of 92 respondents chose the latter (1.2%). There was no statistical difference between Y foils when depicted as male vs. female, chi square 2.136, $p = 0.144$ (The 95% CIs for the ratio of choosing Dr. Y 7 times out of 158 contests is between 2 and 9 percent, and the ratio of choosing Dr. Y 1 time out of 92 contests is between 0 and 6 percent; the CI of the difference in proportions covers 0, ranging from -1.59 to 7.3 percent).² There were no differences between the four disciplines in this male vs. female Y-preference, nor were there

²With low counts some approximations used to compute CIs will not work well, so we used a number of methods to compute the CIs in R. The results were similar: The CIs for the ratio of choosing 7 Ys out of 158 pairings is between 2 and 9 percent,

any differences between the three types of Carnegie institutions or between male and female faculty members, all $p > 0.20$. Finally, faculty gender did not interact with the gender of the Y foil. Basically, everyone preferred the more accomplished X and Z candidates over the less accomplished Y candidate, regardless of Y's gender. And this extended even to fields in which women are very underrepresented (engineering and economics).

and the ratio of choosing 1 Y out of 92 pairings is between 0 and 6 percent. The CI of the difference in proportions covers 0 and ranges from -1.59 to 7.3 percent. Exact numbers below using the R code:

```
(PropCIs)
library(binGroup)
binCI(158,7,.95)
95 percent CP confidence interval
[ 0.018, 0.08915 ]
Point estimate 0.0443
binCI(92,1,.95)
95 percent CP confidence interval
[ 0.0002752, 0.05908 ]
Point estimate 0.01087
scoreci(7,158,.95)
data:
95 percent confidence interval:
0.0216 0.0886
scoreci(1,92,.95)
data:
95 percent confidence interval:
0.0019 0.0590
binom.test(7,158)
Exact binomial test
data: 7 and 158
number of successes = 7, number of trials = 158,  $p < 2.2e^{-16}$ 
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
0.01799534 0.08915030
sample estimates:
probability of success
0.0443038
binom.test(1,92)
Exact binomial test
data: 1 and 92
number of successes = 1, number of trials = 92,  $p < 2.2e^{-16}$ 
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
0.0002751557 0.0590778511
sample estimates:
probability of success
0.01086957
add4ci(7,158,.95)
data:
95 percent confidence interval:
0.02028254 0.09082857
sample estimates:
[1] 0.05555556
add4ci(1,92,.95)
data:
95 percent confidence interval:
0.06605514
sample estimates:
[1] 0.03125
wald2ci(7,158,1,92,.95,adjust = T)
data:
95 percent confidence interval:
-0.01590209 0.07334890
sample estimates:
[1] 0.0287234
```

In view of this finding, there was no justification for conducting a follow-up experiment in which the female Y foil was depicted as less qualified than the 9.3 value used in this experiment, given that she was not preferred even at this slightly lower level, meaning that she would not be ranked higher if she were depicted as lower in quality than 9.3 out of 10.

Discussion

When, as in the present experiment, women candidates are depicted as slightly less accomplished than their male counterparts, they did not have a significant gender advantage in hiring, and were bypassed in favor of slightly superior male candidates 95.2% of the time, which is not significantly different from the 97.47% bypass rate of males depicted as slightly less accomplished (2.53% choosing male Y foil who had a 9.3 score) in Williams and Ceci's (2015) experiment. That is, this result is also similar to the situation in which Dr. Y is depicted as a less accomplished male competing against two stronger female Drs. X and Z. In this latter contest, the male Dr. Y is chosen only 1.2% of the time (all $p > 0.10$, n.s.). Even taking into account low power to detect differences between magnitudes this small, the hundreds of faculty in the Williams and Ceci (2015) study and the hundreds in the present study suggest that it is rare (<5%) to prefer any applicant who is depicted as even slightly weaker than her or his competitors. Apparently, academic faculty view quality as the most important determinant of hiring rankings, which suggests that when women scientists are hired in the academy it is because they are viewed as being equal or superior to male competitors.

Hence, the current findings should help dispel concerns that affirmative hiring practices result in inferior women being hired over superior men (e.g., Niederle et al., 2013). Even though the Dr. Y foil was described as only slightly less accomplished, faculty almost always preferred to hire a slightly more accomplished candidate, and this preference was independent of the gender of the candidates and the gender of faculty raters, and it was observed in both math-intensive and non-intensive fields.

The absence of preference for male Dr. Y does not necessarily imply that academic hiring is meritocratic under all conditions. It is possible that with different levels of candidate information (or if the candidates involved were at a somewhat lower level as opposed to being in the top tier), different results might have been found. For example, in the Steinpreis et al. (1999) study no gender preferences were found when the candidate's CV was highly competitive, but a male preference was found when the CV was less strong. The current study is consistent with these results at the highly competitive candidate level, and showed that slightly less exceptional female candidates were not preferred over exceptional male candidates. Relatedly, Dovidio and Gaertner (2004) findings on aversive racism and selection decisions found that white participants did not discriminate against an unambiguously strong black candidate (vs. a white candidate), but discrimination occurred when the candidates' qualifications were depicted as ambiguous. These findings suggest that discrimination may be a concern when candidate qualifications are ambiguous, but not when candidates

are exceptionally strong. Thus, the most prudent interpretation of the present results is that exceptionally strong candidates of both genders are unlikely to face gender discrimination. Given that the current study focused on top-tier candidates, any conclusions drawn should be confined to excellent tenure-track candidates.

The present findings may provoke concern of a different sort. If affirmative action is intended to not merely give a preference to hiring a woman over an identically-qualified man, but also to tilt the odds toward hiring a woman who may be slightly less accomplished but who is still rated very highly (recall that a 9.3 was in the “extremely impressive” range), gender diversity advocates may be disheartened by these findings. Those who have lobbied for more women to be hired in fields in which they are underrepresented, such as engineering and economics, may find the present findings dismaying and argue that, in the context of hiring in a field in which women are underrepresented, extremely well-qualified female candidates should be given preference over males rated a notch higher. Walton et al. (2013) argued on both empirical and theoretical grounds that hiring more members of devalued groups would actually promote meritocracy, diversity, and organizational performance, not undermine it. (Consideration of this argument entails complexities that are beyond the scope of this study.)

Notwithstanding differing views regarding affirmative hiring of impressive women in underrepresented fields, one claim finds no support in the present results. It is the allegation that the dearth of women in some fields is the result of superior women being bypassed in favor of less accomplished men—a claim made by numerous commentators.³ If academic hiring is

³ Many commentators have opined that female scientists are superior to their male counterparts, and therefore the fact that they are hired at the same rate as men obscures the fact that they should be hired at even higher rates, if merit was the basis for hiring. Consider:

“The studies [claiming gender neutrality] examined odds ratios rather than details of the proposals submitted. This does not rule out the possibility of gender bias. As Marie Vitulli and I said in 2011 [Kessel and Vitulli, 2011], “selection bias can also explain why, in the presence of gender discrimination, female scientists might still fare as well as their male colleagues in some respects if their work was better on average than that of their male peers.” (Kessel, 2012) “Given qualified women drop out of math-intensive fields at higher rates than their male peers. the women who remain are probably, on average, better than their male colleagues and should be having better (hiring) outcomes on average. If their salaries, resources, publication rates, etc. are similar, it then indicates gender discrimination still exists, not that this problem has been solved.” (<http://blogs.sciencemag.org/sciencecareers/2011/02/the-real-cause.html>; retrieved on June 22, 2014) “Female scientists were either not retained or not hired so that only a couple of super-brilliant female scientists were working in staff-scientist positions. On the other hand, several mediocre male scientists were hired and retained, many rising to staff-scientist positions or higher. If you compare these super-brilliant female scientists with their mediocre male counterparts, of course you will not see the difference in their treatment.” (Kali, 2011)

anti-meritocratic, then the weaker male Dr. Y should have been chosen over his stronger female competitors. But as seen, only 1.2% of males who were depicted as the slightly weaker candidate were preferred over slightly stronger female candidates. Thus, there is no support for the view that superior women are being bypassed in favor of inferior men when the contest is between highly accomplished candidates. Hence, these findings call into question claims of current biased tenure-track hiring that have been put forward and they suggest this is a propitious time for talented women to launch tenure-track careers in academic science, where their impressive credentials will be viewed favorably by hiring committees vis-à-vis identically-qualified men.

None of this means that women no longer face unique hurdles in navigating academic science careers. Evidence shows that female lecturers’ teaching ability is down-rated due to their gender (Bug, 2010; MacNell et al., 2015), letter writers for applicants for faculty posts in chemistry and biochemistry use more standout (ability) words when referring to male applicants (Schmader et al., 2007), faculty harbor beliefs about the importance of innate brilliance in fields in which women’s representation is lowest (Leslie et al., 2015), and newly-hired women in biomedical fields receive less than half the median start-up packages of their male colleagues, which could conceivably result in fewer publications down the line (Sege et al., 2015)—to mention a few areas where women continue to face hurdles. Nor do the present findings deny that historic sexist hiring prevented many deserving women from being hired. But these findings do call into question broad or unqualified claims of biased tenure-track hiring that have been put forward. The present findings are not incompatible with earlier studies that found anti-women bias at lower levels hiring a lab manager, (Moss-Racusin et al., 2012) or getting emails returned (Milkman et al., 2012), or hiring members of a math team (Reuben et al., 2014) if one assumes that bias may come into play when diagnostic information is missing (Koch et al., 2015) but not when such information is present as in the case of hiring a candidates who earned doctorates and garnered strong letters and ratings. This suggests that sex biases might reduce the number of women entering training for the STEM pipeline, but our results indicate that when a woman emerges as a strong candidate for a faculty position, she is no longer handicapped as far as being offered the job. Thus, these earlier findings of bias against less accomplished women (e.g., those applying to be lab managers) and the present findings are not mutually exclusive with the current results showing that top-tier female candidates are viewed favorably. This suggests that the gender gap in math-intensive fields might be best addressed by focusing on earlier experiences (encouraging more females to take high school AP physics, computer science, Calculus BC, recruiting more women into college STEM majors—areas identified by Ceci et al. (2014) as associated with the underrepresentation of women in these fields).

These new data will be of interest to academics struggling to increase the representation of women, because our data refute the claim that affirmative hiring policies are non-meritocratic

and lead to less competent women being preferred for jobs. At the same time, these data debunk the claim that less-qualified men are favored over more-qualified women. We found no support for this, either. For those who believe affirmative action means giving a boost to an underrepresented group when all else is equal, our data will be welcome news, since we show that academic hiring preferences are quality-based. However, for those who argue that affirmative action means choosing slightly less accomplished individuals over more accomplished ones for reasons of diversity, our data suggest that at least when it comes to gender, faculty may be reluctant to embrace this pathway to diversity.

Possible Reactions to These Findings

Our work on this topic has led to certain comments that we have heard repeatedly. We note some of these below along with a few reactions in response to them:

1. *"Thank goodness the academy is still a meritocracy in which competence determines who's hired."* Some readers will likely be pleased that the academy is still a domain in which competence as traditionally quantified matters more than social factors. Such individuals believe that the academy should continue to function exactly this way, and are heartened when presented with evidence that it does.
2. *"The 1–10 rating scale assumes that what is being rated is what matters—and women are often good at things not assessed by this scale, such as collaborative work, advising, and service."* Of course other attributes are important for a professorial career, attributes that are not measured by the scale we used which was based on research publications, teaching awards, job talk, and letters. However, the assumption of those who raise this point is that women are superior to men in these unmeasured skills—which is actually an empirical question. Men may be superior in them, or women may be, or both groups may be equivalent. Simply because an attribute is not assessed does not mean that women are superior at it, nor will inclusion of it necessarily boost women more than men, and close the gender gap (Ceci and Papierno, 2005).
3. *"The entire male-centered, Western notion of assigning a job applicant a 'quality' score on a 1–10 scale is misguided at its core."* Critics espousing this view are often in favor of a reinterpretation of everyday constructs. They eschew the notion that publications, awards, letters of reference, job talks, citations, and grants are the most important indicators of ability and predictors of success as a professor. Proposals for alternative, empirically-tested, valid and reliable indicators, and predictors of professorial success are most welcome, so that we can think more broadly.
4. *"Such ranking experiments have nothing to do with real-world hiring."* As we noted, women have significant advantages in actual, real-world hiring—they are hired at higher rates than men. Some of our critics seem reluctant to acknowledge this fact, which is shown clearly in multiple audit studies that analyze who is actually hired at universities in the U.S. and Canada (see cites in Williams and Ceci, 2015). To argue that our experiment has no relevance to real-world hiring seems unpersuasive in view of the fact that in the real world of academic hiring women also are chosen over men in disproportionate numbers. As one commentator noted in arguing for the relevance of the current experimental design: *"One would have to say both that women are, in fact, stronger candidates (which is one strong assumption for which there is no direct evidence), implying that faculty don't prefer them over equally qualified men in real hiring contexts, and that, nonetheless, faculty DO prefer them in hypothetical situations (another strong assumption for which there is no direct evidence). By far the most sensible explanation is the most economical one: faculty prefer women both in the hypothetical case and the real case; their preferences don't swing wildly from the actual to the hypothetical."*
5. *"The process of assigning a rating to a woman's dossier is inherently prone to sexist bias; thus, women are less likely to receive an equivalent rating to that of male competitors."* This is a popular view; however, we found that subjects evaluating a single dossier, presented as either female or male, assigned a significantly HIGHER rating to that dossier when it belonged to a woman than a man (8.20 vs. 7.14, $p < 0.01$). The translation of traditional indicia (publications, letters, etc.) into ratings seems to work *at least as well* when ranking a female applicant.

Limitations of Present Study

No experiment is perfect, and this one is no exception. It is possible that the faculty raters rarely chose the less competent candidate because they were supplied with "pre-ranked" quantitative ratings of the candidates (e.g., 9.3 or 9.5 on a 10-point scale). Hence, the present results may have been influenced by giving faculty "pre-ranked" ratings. Perhaps in the absence of being given quantitative ratings, faculty will shift criteria to justify their final decisions (e.g., be influenced by gender to give more credence to the eminence of an applicant's advisor/institution if a woman's list of publications is shorter than her male competitor's). Assigning pre-ranking scores will likely be variable in actual hiring; this variability in assigning scores could increase the rate of selecting someone who's rated a 9.3 on average. In other words, disagreements would likely be more common in actual hiring decisions due to the variable ways faculty translate their impressions. Since concerns about personality and teaching performance were raised for Dr. Y, but not for Drs. X or Y the primary reason someone might want to hire Dr. Y was gender when Y was a woman.

The present data provide no hint of the extent to which this occurs. However, in Experiment 5 of the Williams and Ceci (2015) paper, 127 faculty were given only one applicant to rank, either a man or woman who were identically accomplished. When the applicant was a man the faculty who were asked to rate his strength gave him a rating of 7.14 but when the identical portfolio belonged to a woman, the faculty who were asked to rate her gave her a rating of 8.2 ($p < 0.01$). So there is some suggestion that faculty shift their quantitative ratings to justify their preference for women, even when they are asked to generate the rating themselves, for what are actually identical accomplishments of both genders.

If true, the present findings suggest this shift is limited to conditions in which candidates are identically competent and very accomplished.

In future research it would be interesting to vary the CVs of the 9.3-rated female applicant and the 9.5-rated male applicants in terms of their number of publications, advisor eminence, teaching awards during graduate school, the prestige of their PhD-granting institutions, etc. to determine how much shift in faculty-assigned quantitative ratings is observed as a function of applicant gender. In this experiment we began with the smallest difference of 9.5 vs. 9.3, with a plan to widen this gap if it turned out that faculty preferred slightly weaker women; but since they did not, there was no reason to widen the gap.

The low base rate for choosing the Y foil presents statistical issues: The rate of selecting the female Dr. Y (4.8%) was slightly higher than rate of selecting the male Dr. Y (1.2%), although this difference was not statistically reliable. Statisticians have written about the challenges of comparing frequencies of rare events (e.g., Bradburn et al., 2007). This has ramifications if the null result is affected by low statistical power, and future research might enlarge the sample size to see whether weaker women may be preferred over stronger men. However, faculty preference for the less qualified Dr. Y candidate was always rare in this experiment and in the Williams and Ceci (2015) one (<5%), regardless of applicant gender, so even if a preference for the weaker female became significant, the magnitude of such an effect would likely be quite small.

Although the current study is well-suited to address the specific question it posed, it employed a very specific methodology and DV that may have limited the operation and detection of implicit bias. It is possible that the use of implicit measures may have revealed bias as has been observed to occur even among university professors. Measures of explicit bias may not always be collinear with implicit measures (see Smyth and Nosek, 2015). As was noted in the introduction, findings from real-world hiring audits (not experiments, but actual hiring of university professors) indicate female applicants are typically hired at higher rates than their male counterparts—for at least the last two decades (Williams and Ceci, 2015).

Many have argued that the pro-women hiring preference is because women are on average stronger applicants, by dint

of the winnowing process they have survived from college-to-graduate school-to applying for tenure track jobs: it is argued that the reason women are more likely to be hired than their male counterparts for tenure track jobs in the real world is because those women who end up applying for tenure-track jobs represent the “cream of the cream,” a higher mean quality than the typical male applicant. Williams and Ceci (2015) designed their experiments to test this claim and reported that even when applicant strength was equated (experiments 1–3), faculty still preferred female applicants over identical male applicants. And as noted above, in their fifth experiment 127 faculty were asked to assign their own strength ratings (on a 10-point scale) to either a man or woman applicant. Faculty rated the same applicant 8.2 when it had a woman’s name on it but only 7.14 when it had a man’s name on it. So Dr. Y did not receive lower scores when described as a woman, and higher scores when described as a man, as some would predict.

Finally, the experimental condition that involved two female finalists (out of three) might have seemed odd for a STEM faculty member in math-intensive fields where 70%-plus of applicants are often male. On the flip side, having the woman be lower-rated than two men might have also made gender more salient. To the extent that either of these is true, it is an important issue that future research should address (e.g., by conducting focus groups or using a shortlists of only two applicants, only one of whom is female—a situation we deliberately rejected because we felt it might make the gender contest overly salient and explicit). However, in view of the media and publicity surrounding findings from these type of experimental designs, follow-up research cannot be undertaken in the near future without compromising the experimental reactivity of participants.

Funding

This research was supported by NIH Grant 1R01NS069792-01.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01532>

References

- AAUP. (2014). *Affirmative-Action Plans: Recommended Procedures for Increasing the Number of Minority Persons and Women on College and University Faculty*. AAUP Reports and Publications. Available online at: <http://www.aaup.org/report/affirmative-action-plans-recommended-procedures-increasing-number-minority-persons-and-women> (Accessed March 31, 2015).
- Bradburn, M. J., Deeks, J. D., Berlin, J. A., and Russell Localio, A. (2007). Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat. Med.* 26, 53–77. doi: 10.1002/sim.2528
- Bug, A. (2010, December 7). Swimming against the unseen tide. *Physics World*. doi: 10.1088/2058-7058/23/08/27
- Ceci, S. J., Ginther, D., Kahn, S., and Williams, W. M. (2014). Women in science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., and Papierno, P. B. (2005). The rhetoric and reality of gap-closing: when the “have-nots” gain, but the “haves” gain even more. *Am. Psychol.* 60, 149–160. doi: 10.1037/0003-066X.60.2.149
- Cuddy, A., Glick, P., and Fiske, S. T. (2004). When professional become mothers, warmth doesn’t cut the ice. *J. Soc. Issues* 60, 701–718. doi: 10.1111/j.0022-4537.2004.00381.x
- Diekmann, A. B., and Eagley, A. (2000). Stereotypes as dynamic constructs: women and men of the past, present, and future. *Pers. Soc. Psychol. Bull.* 26, 1171–1188. doi: 10.1177/0146167200262001
- Dovidio, J. F., and Gaertner, S. L. (2004). “Aversive racism,” in *Advances in Experimental Social Psychology*, Vol. 36, eds J. M. Olson and M. P. Zanna (San Diego, CA: Academic Press), 1–52. doi: 10.1016/S0065-2601(04)36001-6
- Glaser, J., and Knowles, E. D. (2008). Implicit motivation to control prejudice. *J. Exp. Soc. Psychol.* 44, 164–172. doi: 10.1016/j.jesp.2007.01.002

- Kang, J., and Banaji, M. R. (2006). Fair measures: a behavioral realist revision of "affirmative action." *Calif. Law Rev.* 94, 1063–1118. doi: 10.2307/20439059
- Koch, A. J., D'Mello, S. D., and Sackett, P. R. (2015). A meta-analysis of gender stereotypes and bias in experimental simulations of employment decision making. *J. Appl. Psychol.* 100, 128–161. doi: 10.1037/a0036734
- Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375
- MacNeill, L., Driscoll, A., and Hunt, A. N. (2015). What's in a name? Exposing gender bias in student ratings of teaching. *Innovative High. Educ.* 40, 291–303. doi: 10.1007/s10755-014-9313-4
- Milkman, K. L., Akinola, M., and Chugh, D. (2012). Temporal distance and discrimination: an audit study in academia. *Psychol. Sci.* 23, 710–717. doi: 10.1177/0956797611434539
- Miller, D. I., Eagly, A. H., and Linn, M. G. (2015). Women's representation in science predicts national gender-science stereotypes: evidence from 66 nations. *J. Educ. Psychol.* 107, 631–644. doi: 10.1037/edu0000005
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16474–16479. doi: 10.1073/pnas.1211286109
- Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Manage. Sci.* 59, 1–16. doi: 10.1287/mnsc.1120.1602
- National Research Council (NRC). (2009). *Gender Differences at Critical Transitions in the Careers of Science, Engineering and Mathematics Faculty*. Washington, DC: National Academy Press.
- Reuben, E., Sapienza, P., and Zingales, L. (2014). How stereotypes impair women's careers in science. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4403–4408. doi: 10.1073/pnas.1314788111
- Schmader, T., Whitehead, J., and Wysocki, V. H. (2007). A linguistic comparison of letters of recommendation for male and female chemistry and biochemistry job applicants. *Sex Roles* 57, 509–514. doi: 10.1007/s11199-007-9291-4
- Sege, R., Nykiel-Bub, L., and Selk, S. (2015). Sex differences in institutional support for junior biomedical researchers. *JAMA* 314, 1175–1176. doi: 10.1001/jama.2015.8517
- Sheltzer, J. M., and Smith, J. C. (2014). Elite male faculty in the life sciences employ fewer women. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10107–10112. doi: 10.1073/pnas.1403334111
- Smith, D. G., Turner, C. S. V., Osei-Kofi, N., and Richards, S. (2004). Interrupting the usual: successful strategies for hiring diverse faculty. *J. High. Educ.* 75, 133–160. doi: 10.1353/jhe.2004.0006
- Smyth, F. L., and Nosek, B. A. (2015). On the gender-science stereotypes held by scientists: explicit accord with gender-ratios, implicit accord with scientific identity. *Front. Psychol.* 6:415. doi: 10.3389/fpsyg.2015.00415
- Steinpreis, R., Anders, R. K., and Ritzke, K. D. (1999). The impact of gender on the review of the CVs of job applicants and tenure candidates: a national empirical study. *Sex Roles* 41, 509–528. doi: 10.1023/A:1018839203698
- Turner, C. S. V., Gonzalez, J., and Wood, J. L. (2008). Faculty of color in the academy: what 20 years of literature tells us. *J. Divers. High. Educ.* 1, 139–168. doi: 10.1037/a0012837
- Walton, G. M., Spencer, S. J., and Erman, S. (2013). Affirmative meritocracy. *Soc. Issues Policy Rev.* 7, 1–35. doi: 10.1111/j.1751-2409.2012.01041.x
- Williams, W. M., and Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5360–5365. doi: 10.1073/pnas.1418878112

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Ceci and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Leaky Pipeline Myths: In Search of Gender Effects on the Job Market and Early Career Publishing in Philosophy

Sean Allen-Hermanson*

Department of Philosophy, Florida International University, Miami, FL, United States

That philosophy is an outlier in the humanities when it comes to the underrepresentation of women has been the occasion for much discussion about possible effects of subtle forms of prejudice, including implicit bias and stereotype threat. While these ideas have become familiar to the philosophical community, there has only recently been a surge of interest in acquiring field-specific data. This paper adds to quantitative findings bearing on hypotheses about the effects of unconscious prejudice on two important stages along career pathways: tenure-track hiring and early career publishing.

Keywords: underrepresentation, gender bias, sexism, hiring, philosophy

OPEN ACCESS

Edited by:

Stephen J. Ceci,
Cornell University, United States

Reviewed by:

Rafael De Clercq,
Lingnan University, China
Matt L. Drabek,
ACT, Inc., United States

*Correspondence:

Sean Allen-Hermanson
hermanso@fiu.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 27 March 2017

Accepted: 23 May 2017

Published: 14 June 2017

Citation:

Allen-Hermanson S (2017) Leaky Pipeline Myths: In Search of Gender Effects on the Job Market and Early Career Publishing in Philosophy. *Front. Psychol.* 8:953. doi: 10.3389/fpsyg.2017.00953

INTRODUCTION

That philosophy is an outlier in the humanities when it comes to the underrepresentation of women has been the occasion for a lot of discussion about possible effects of subtle forms of prejudice, including implicit bias and stereotype threat. Though real-world effects are not strongly evidenced (Forscher et al., 2017), there is widespread concern in philosophy that involuntary and unconscious implicit associations might diverge from a person's declared beliefs affecting our actions, judgments, and attitudes. Unconscious bias might influence how we treat junior colleagues from socially stigmatized groups when it comes to sharing opportunities for professional development, advancement, and in evaluating scholarly potential and credentials. For example, a departmental committee might implicitly prefer a male candidate to a female candidate with the same qualifications despite holding conscious and explicit attitudes about the equality of the sexes. Meanwhile, stereotype threat is when awareness of and identification with stereotypes (such as that philosophy is for white males) results in heightened anxiety, performance disparities, and reduced interest.

These ideas have become familiar to the philosophical community, which continues to debate policy initiatives and other measures for improving diversity, such as making syllabi and conference line-ups more inclusive, adjusting the management of professional organizations, and reforming journal and hiring practices. These ongoing discussions need to be informed by the best possible evidence, and there is a growing interest in acquiring field-specific data. The investigatory model informing this study is inspired by the hiring audits used in STEM disciplines. This paper contributes data pertinent to hypotheses about the effects of prejudice on two important stages of career pathways: tenure-track hiring and early career publishing.

If women are evaluated more harshly because of unconscious bias on the part of letter writers and hiring committees, or have weaker files and perform less well in interviews because of stereotype threat, or even face conscious and explicit discrimination, then they might be expected

to be less successful at finding tenure-track employment. Indeed, biases are often conjectured to be a major cause of the underrepresentation of women in philosophy.¹ Fortunately there have been several recent studies of employment trends, including Jennings' for a 2-year period (2012 and 2013) and a follow-up funded by the American Philosophical Association known as the APDA report.² These and other resources can help test hypotheses predicting effects of biases on hiring.³ Working from her original data, it was decided that post-doctorate appointments would be ignored in order to focus on more desirable tenure-track lines, leaving us with 229 men and 109 women in the pool. Since this manuscript was written, the data used in the APDA report has been corrected and revised and will also be taken into account. So, how are women doing on the job market?⁴

RESULTS

Analysis of Jennings' original data suggests women and men are hired at a rate roughly proportionate to their numbers for entry-level tenure-track jobs in philosophy.⁵ Concerning the follow-ups, the first APDA report in 2015 actually found women were hired significantly more often, increasing the odds of obtaining a permanent academic position by 85%.⁶ A 2016 update corroborated this finding notwithstanding Jennings' statement that "we did not find a significant effect of gender on placement..."⁷ While technically true, however, this was only because in 2016 they elected not to examine this low hanging fruit. Despite this flagging of interest in this aspect of the gender question, she did acknowledge there were about 10% more women than men among those obtaining a permanent position over the entire data set (chi-squared test $p < 0.01$) and (with prompting by an anonymous commenter) admitted the number was much higher (around 23%) concerning those graduating within the most recent hiring cycles (2012–2015). Another commentator also brought the high significance of this result to our attention (chi-square, $p = 0.0007$). The upshot is that with the help of various anonymous commentators, we can be confident that the APDA findings offer strong support that men are significantly less likely to obtain permanent positions.

¹"Implicit bias and stereotype threat...will make it harder for women to do well...to be recognized in graduate school, less likely to get strong letters of recommendation, and less likely to be hired. The women who, despite this, get hired at strong research departments are likely to be especially exceptional philosophers" (Saul, 2012).

²The Academic Placement and Data Analysis project's results can be found here: <http://webcache.googleusercontent.com/search?q=cache:szvEA0v8tSgJ:dailynous.com/wp-content/uploads/2016/04/apdareportupdate2015report.pdf+&cd=2&hl=en&ct=clnk&gl=us&client=safari>

³<http://www.newappsblog.com/2014/12/gender-and-publications.html#more>. See also Solomon and Clarke (2009).

⁴Here and there minor corrections were made for errors, such as duplicate entries.

⁵All of our raw data can be accessed at Genderandphilosophy.blogspot.com

⁶See p. 11 of the 2015 report: <http://dailynous.com/wp-content/uploads/2016/04/apdafinalreport2015.pdf>

⁷<http://dailynous.com/2016/04/15/philosophy-placement-data-and-analysis-an-update/>

Another noteworthy finding obtained from Jennings' earlier results is that female candidates had about half as many publications as their male counterparts. The average publication counts for candidates (having no prior academic appointment) were 1.37 for men and .77 for women (medians 1 and 0; $p = 0.000808$)⁸ and is also extremely significant. However, as the quantity of publications is only one very crude measure of a candidate's strengths, we can also look at how several other variables might depart from these aggregated results. For example, besides quantity, Jennings' data also contained information about the putative "quality" of a publication (defined as a "top-15" journal according to a poll at the Brooks Blog). Here, male tenure-track hires are about three times more likely to have published in a highly regarded venue (see **Appendix** and **Figure 1**).

It is natural to wonder how important publications are when it comes to assessing job candidates, and certainly we can agree publications are not the only relevant factor. Even for a research-oriented position the quality of writing samples, the reputation of doctoral institutions, and the weight assigned to letters of recommendation will also be taken into account. As a rough proxy for reputational factors rankings of degree-granting programs obtained from the 2006–2008 edition of the Gourmet Report were utilized.⁹ How publication records of new hires might differ depending on whether they had a prior position or accepted a tenure-track job straight out of graduate school was also considered.

Some have claimed that prestige interacts with gender in that women from highly regarded programs tend to publish the least, whereas men from less fancy programs publish the most.¹⁰ However, we find the relationship between gender and prestige to be somewhat murkier. Although the finding men in general tend to publish more and in more prestigious places was corroborated, program prestige correlated positively with the output of high-quality publications regardless of gender. Gendered differences also depended somewhat on whether we were looking at candidates who had held a prior academic appointment.

First, mean Gourmet Report scores were incorporated within Jennings' spreadsheet revealing a disparity in average home department rankings of 3.31 for men and 2.93 for women (medians were 3.6 and 3.2). High prestige "top-20" departments have a score of at least 3.4, and so next men and women were divided into elite ("top-20") and non-elite ("non-top-20") subgroups to see if there would be any interesting effects.¹¹ The rankings for male and female "top-20" hires turn out to be very similar with mean scores of 3.94 for men and 4.02 for women (with medians of 3.7 each). Given this, we expect to find a small gendered difference in prestige among the remainder, and

⁸Meaning a 99.9192% confidence that the result is not due to chance. P was obtained using a Fisher exact test.

⁹It was reasoned that the reputational ranking should reflect the fact that candidates take around 5 or 6 years to obtain their degrees. But the results do not differ much if a slightly more recent or less recent edition is used instead.

¹⁰See <http://genderandprestige.blogspot.com>

¹¹Note that candidates from top-20 schools appear to fill half of new tenure-line positions.

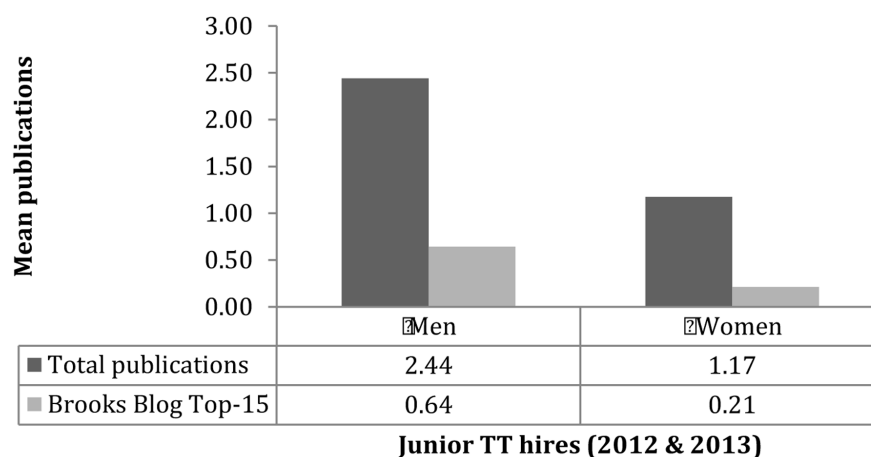


FIGURE 1 | Gender and publishing.

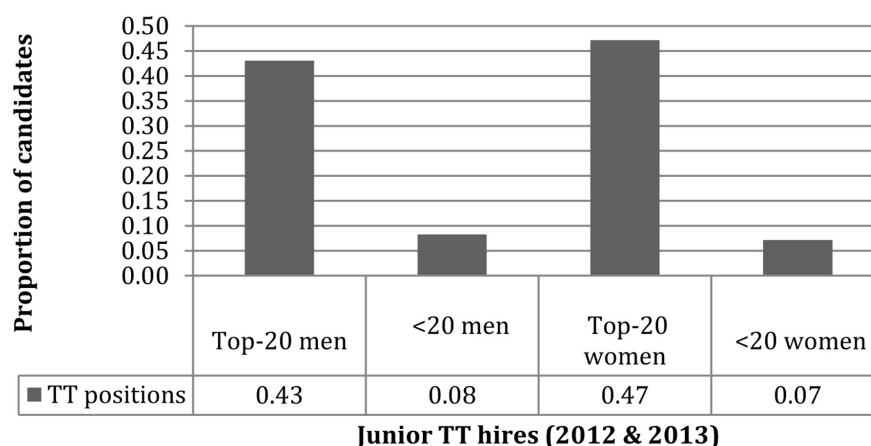


FIGURE 2 | Gourmet-ranked TT positions.

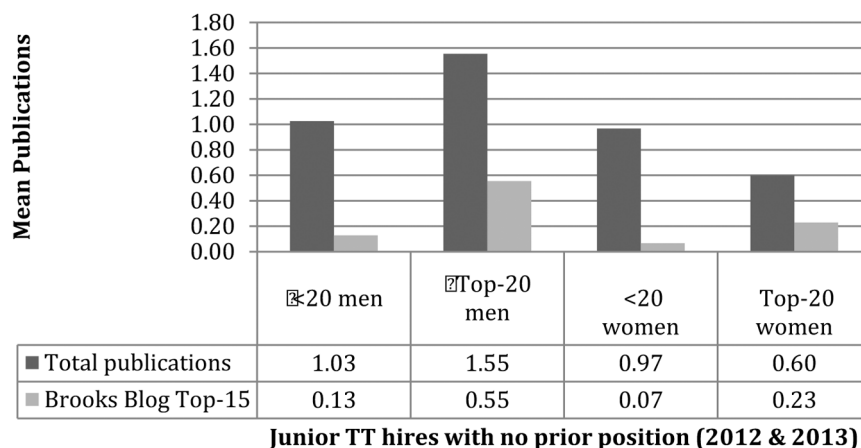
indeed the averages here are 2.26 for men and 1.9 for women (with medians of 2.6 and 2.3).¹² It was also noted that there was no significant interaction between gender and the prestige of hiring departments, though there was evidence candidates from relatively lower prestige institutions lack upward mobility: Whereas top candidates of either gender could expect to find a position at a Gourmet-ranked institution a bit less than half the time, this was true of other candidates only 7–8 percent of the time (**Figure 2**). This might indicate that there are, in effect, two semi-independent job markets. In terms of outcomes, there seems to be a top-20 market mostly closed to non-elite candidates and a non-top-20 market open to all.

Next, turn to consider how prestige might interact with gender when it comes to publishing. As mentioned earlier, differences depended on whether candidates had a prior appointment. In considering those with no prior appointment,

¹²Women seem to be slightly more likely to obtain degrees from unranked programs, which were scored as a “0.” Hence, if we ignore unranked programs these small differences in the overall rankings disappear.

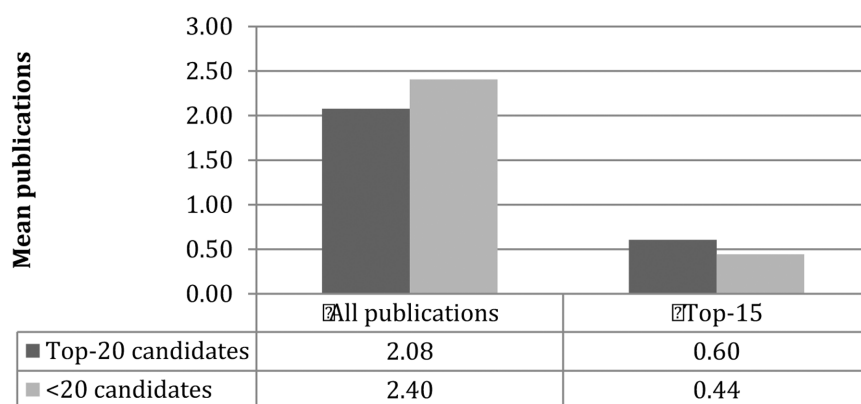
it was found that top-20 men stand out: they publish more, and in “better” places than the others. Meanwhile, top-20 women publish less often than non-top-20 men and women, nevertheless they tend to do better when it comes to quality (**Figure 3**). As it is unclear how to weight quantity versus “quality” in assessing candidate strength, no conclusions are drawn here about the advantages or disadvantages of the remaining subgroups. We can observe that top-20 women have much more access to top-20 jobs, which might suggest “quality” counts for more across the market. Alternatively, there might be different standards for the different “markets” proposed above: top-20 individuals appear to be a little stronger concerning “quality” and non-top-20 are stronger for quantity (**Figure 4**). Perhaps then publishing counts, but counts differently depending only whether one is competing on the “elite” market favoring “quality” or the “non-elite” market favoring raw output.¹³

¹³Against this the reader is asked to compare the data in **Figures 3, 5**, which might suggest that when it comes to re-entering the market with a prior position it may



Junior TT hires with no prior position (2012 & 2013)

FIGURE 3 | Publishing (no prior position).



Junior TT hires (2012 & 2013)

FIGURE 4 | Prestige and publishing.

Now we can consider the candidates who did have a prior position. Here, men had significantly higher averages for quantity and “quality” (Figure 5). For example, low-prestige men published almost three times as much as the average high-prestige women and were about two times more likely to have a top-15 publication. This might indicate that women, regardless of prestige, tend to submit to journals less frequently because they are less confident, as expected by hypotheses invoking stereotype threat¹⁴ or even disadvantages in the reviewing process.¹⁵ Then again, there are several other explanations for the publishing gap. Notwithstanding this uncertainty, men and women appear to be held to different standards.

be more important for non-elite candidates of either gender to improve “quality” and elite candidates to improve quantity.

¹⁴Though we note that enthusiasm for stereotype threat theory is in rapid decline thanks to concerns about ecological validity, experimental design and interpretation, replicability, and evidence of publication bias, see Sesardic and De Clercq (2014).

¹⁵This finding of a gender gap in publishing is in step with Krishnamurthy et al. (forthcoming).

Returning to an earlier suggestion, might it be the case that publications are not that important in hiring? This is hard to accept given that productivity is so often tied to securing research-intensive positions in the competitive academic environment and critical to determinations about prospects for earning tenure. This seems clear when we consider lateral hires, which constitute much of the data, and as just mentioned indicate upward trends in output. That candidates from less fancy programs publish more regardless of gender also suggests a widespread presumption that publishing compensates for other deficiencies. We can also note that previous research indicates that publication records are a critical indicator of candidate strength (Steinpreis et al., 1999).

In a blog comment, Jennings¹⁶ attempts to explain away the publishing gap by proposing that enhanced opportunities are more often offered to males. Others have also worried that “at the graduate level, supervisors may be more likely to encourage men to publish their work” (Saul, 2013). Jennings wonders if most of

¹⁶<http://www.newappsblog.com/2014/12/gender-and-publications.html#more>

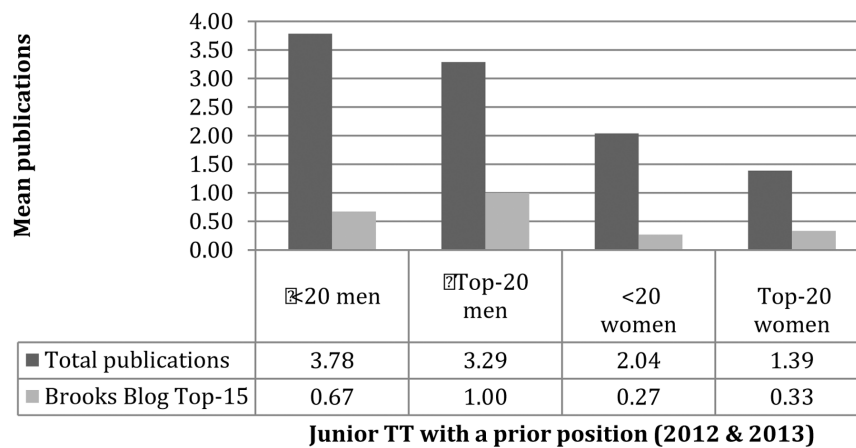


FIGURE 5 | Publishing (with prior position).

the difference between genders is attributable to a handful (15%) of high productivity (HP) men, defined as those with at least five publications at the time of hiring (5% of women are HP by the same standard).¹⁷ Yet these numbers are derived from looking at all hires, including those candidates who had a prior academic appointment, and therefore more opportunities to publish. When we consider those with no prior position, none of women and only 3% of the men ($n = 7$) are HP. Since there are so few, the gap cannot be explained by those who are highly productive.

Nevertheless, all HP hires were examined in order to see what proportion of their work might be attributable to enhanced opportunities.¹⁸ Using Google searches of cvs the number of such publications for each of the 61 HP candidates was obtained by counting works that were co-authored with a senior figure, articles or chapters in edited volumes, conference proceedings, and publications which otherwise appeared to be by invitation, such as introductions to special issues.¹⁹ Although the pool of HP women is small²⁰ it was found that 37.5% of their publications fell into this category. Turning to the men, first those who were exceptionally productive (having at least 10 publications) were examined. The rationale here is that if the HP men are favored with extra opportunities, this will likely be reflected in the output of those who publish the most. Yet for this group, only 34.3% of their work was attributable to enhanced opportunities²¹ while the result was 37.7% for all HP men.²² In addition, the means, medians, and modes for those in the HP group did not significantly vary by gender. While not exhaustive, as there could

be other kinds of special opportunities, favoritism in a non-blind review process, as well as differential barriers to obtaining prior positions, the data offered here suggests HP men and women are treated about equally. It was found that highly productive men were about twice as likely to publish in well-regarded (“top-15”) journals.²³

Returning to market outcomes, the previous results were augmented by placement data obtained from two additional sources: the American Philosophical Association’s Guide to Graduate Programs²⁴ and the Philjobs website²⁵ and cohere with similar findings from hiring audits in STEM fields (Williams and Ceci, 2015). In addition, individual notices of new appointments from Philjobs for the 2014 hiring season were monitored.²⁶ Next the analysis of the APA data is presented followed by a consideration of the findings obtained from Philjobs.

Data was transcribed about gender and hiring found in the APA’s 2013 and 2014 Guides to Graduate Programs for two 5-year periods: 2008–2013 and 2009–2014. Only programs that allowed for a comparison between hiring outcomes and how many men and women went to market were included in the calculations. The Guides provided data for 64 schools in the 2013 edition and 65 for 2014 (37 schools provided data twice, so there is placement information available for 92 distinct programs for these mostly overlapping timespans). For 2008–2013 it was found that 40% of men who went on the market eventually landed a tenure-track job compared to 50.6% of the women, meaning a woman’s probability of obtaining tenure-track employment was about 25% better ($p = 0.037$).²⁷ The corresponding probabilities of obtaining any kind of academic position (including much less desirable temporary appointments) were a lot closer at 86 and

¹⁷HP candidates were very likely to have had at least one prior position (median for both genders = 1; means were 0.88 for men and 0.86 for women).

¹⁸In this case, we also included 10 highly productive individuals with post-doctorates in order to increase the population. $M = 54$; $F = 7$. Post-docs are shaded in blue in the spreadsheet titled “high performers” and tenure-track hires are in yellow.

¹⁹We didn’t count book reviews as peer-reviewed publications. If we do, the proportion of work attributable to enhanced opportunities falls to 31%.

²⁰There are about a third as many as one would expect given that women account for 32.2% of the hires.

²¹Falling to 31% with book reviews.

²²Falling to 32.8% with book reviews.

²³Using Jennings’ data, we found that the highly productive men (with at least five publications) had an average of 1.98 publications in top journals. Women had an average of 1.14. The median was 1 in both cases.

²⁴<http://www.apaonline.org/?page=gradguide>

²⁵<http://philjobs.org/appointments/dataFeed>

²⁶Philjobs <http://philjobs.org/appointments>

²⁷Calculated using a Fisher Test.

89%. Women also made up 26.3% of the market and 31.1% of the tenure-track placements. The 2014 Guide reinforces this pattern, with 35.3% of men and 46.7% of women finding tenure-track employment from 2009 to 2014 meaning the probabilities were about one-third better for women ($p = 0.016$).²⁸ Similar to before, 83.6% of men and 87.8% of women found any kind of academic job. Women made up 25.2% of the market and 31% of junior tenure-track hires.

One might wonder if schools with good placement records, and, especially, good records for placing women might be overrepresented in the APA data. However, this concern is not realistic. Consider what it would take for the schools where we didn't have data to close the gender gap. According to the APA, the 2008–2013 period comprised 530 junior tenure-track placements, and yet we would have to suppose an additional 200 competitions went unreported in which a man won every single time—in that case the chances equalize to 50%. There would have to be more than 500 unreported tenure-earning jobs going solely to men for the disproportion to be reversed (i.e., for men to have a 25% greater chance). In such a small profession, there are probably not enough unreported jobs for this to be the case: while Philjobs reported 816 junior tenure-track placements for the same period, many of these are lateral moves that placement officers would not normally pass on to the APA.

Adding to uncertainty about possible unreported hires, one might also wonder if these results would hold up for periods other than 2009–2014, and what the year-to-year results look like. With these concerns in mind we can turn to data provided by

Philjobs. This process began by examining the 2014 hiring season, which was arbitrarily defined as spanning July 1, 2014 to June 30, 2015. Over the course of the year information was gathered about individual tenure-track hires, including those who had a previous academic appointment as well as those going to market straight from graduate school. For 2014 it was found that 56 out of 148 hires (37.8%) went to women. While the number of doctorates awarded to women as a percentage of the total doctorates in philosophy fluctuates somewhat from 1 year to the next, it was assumed that the year immediately prior would give a reasonable approximation of how gender is distributed on the job market; in 2013, for example, 27% were awarded to women.²⁹

To add more depth to the investigation hiring outcomes for nine further years (2005–2013) were also examined using data provided by the Philjobs website. In order to make this information useable certain corrections and additions to their spreadsheet were necessary, including the elimination of duplicated entries, filtering out senior appointments and non-tenure-track hires, spot-checking for accuracy, and using Google searches to ascertain gender where it was missing or in doubt. Next, year-by-year comparisons were made between placement and the distribution of philosophy PhDs using the NSF's Survey of Earned Doctorates. The relationship between awarded PhDs and junior hiring from 2005 to 2014 is depicted in **Figure 6** and gives a general sense of the market. The distribution of philosophy doctorates by gender for the same period is found in **Figure 7**. For some years (2005, 2006, 2008, 2009, 2012) there was

²⁸Fisher Test.

²⁹See the 2013 edition of the NSF's Survey of Earned Doctorates.

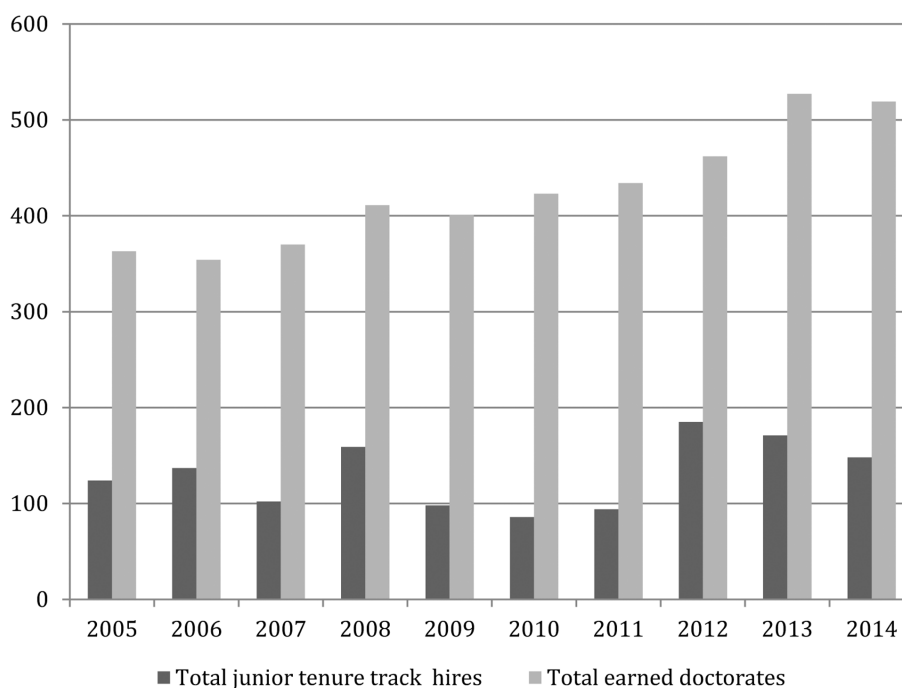


FIGURE 6 | Earned PhDs and TT hiring.

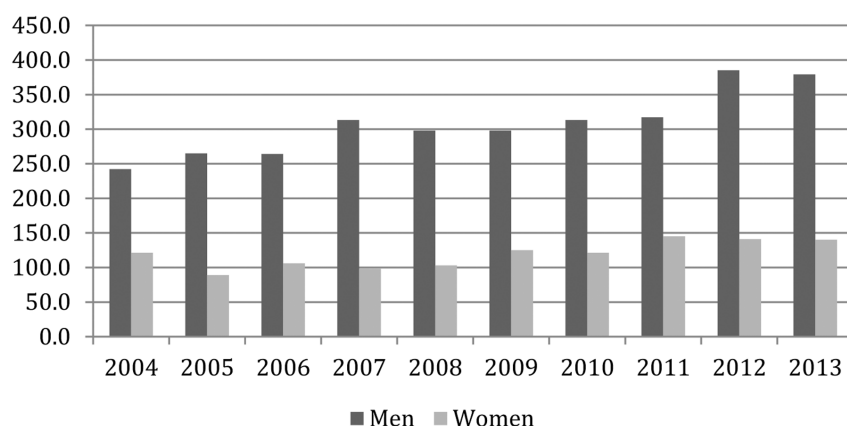


FIGURE 7 | Earned Doctorates by gender.

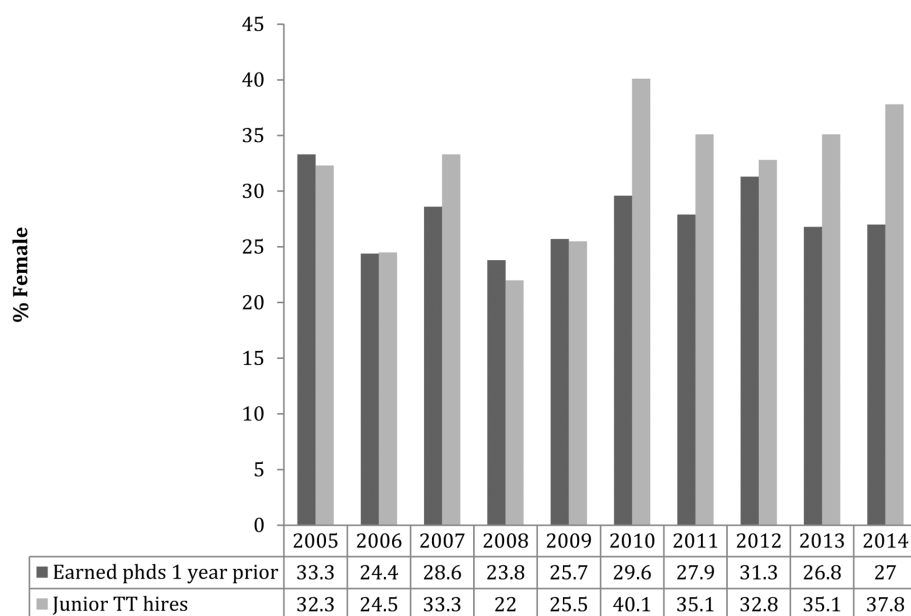


FIGURE 8 | Gender and tenure-track hiring.

a rough correspondence between the percentage of women who were hired and their share of philosophy doctorates awarded in the year immediately prior. As a further check the rankings for the 2005 market (using the 2001 edition of the Gourmet Report) were consulted, though there were no significant differences in the means (3.00 for men and 3.01 for women) or medians (men: 3.3; women: 3.35) of candidates. For the remaining years (including most recent hiring seasons) women appear to be overrepresented, accounting for 28.4% of the earned doctorates but 35.73% of tenure-track hires (**Figure 8**).³⁰

Finally, these results were compared to updates found in the 2016 APDA report. First, my list of successful job candidates for the 2012 season was merged with the APDA's. Although

these mostly overlapped, there were some differences. In order to seek greater accuracy every candidate was re-checked, one-by-one, in attempts to verify gender and success in a tenure-track competition in 2012 (e.g., by consulting cvs, locating welcome messages at hiring Departments, etc.). Both data sets contained errors resulting in 56 changes to my list (37 additions and 19 deletions) and 36 changes to the APDA's (29 additions and 7 deletions), thus bringing the two into harmony.³¹ Though this process was tedious and time-consuming, it was hoped it would maximize the accuracy of the data for at least 1 year and so allow us to see if this additional scrutiny would alter the results

³⁰ $p = 0.000372$.

³¹The final tally for 2012 was 140 men and 61 women. Note that there was only one instance where a gender was assigned incorrectly due to a limitation of the software used by the APDA project.

in any significant way. With this revised data it was then a simple task to recalculate the hiring figures. According to my original survey 32.7% of tenure-track hires went to women in 2012 whereas the APDA's 2016 report puts this a little lower at 30.7%. The outcome for the revised and re-verified data is just shy of their result at 30.3%. To place this in context, note that in the previous year 31.3% of doctorates in philosophy went to women. Hence, it can be reaffirmed that the 2012 market outcomes do not attest to a significant gender effect in hiring. However, 2012 was also unusual in light of the pattern for the years 2010, 2011, 2013, and 2014, which might indicate significant bias in favor of female candidates (**Figure 8**). Would this pattern also stand up to further scrutiny? This time instead of more forensic checking of merged data sets, the APDA's numbers were taken at face value with a result in keeping with my original findings provided in **Figure 8**. Going by the APDA's data women obtained 32.5% of the tenure-track jobs in 2013 and 39% in 2014 whereas my results were 35.1% (with 26.7% earning doctorates) and 37.8% (27% earning doctorates). Instead of quibbling about a percentage point here or there, it can be agreed there is no evidence women are underrepresented among those obtaining tenure-track jobs for at least a decade. To the contrary, recent years seem to attest to a reverse gender effect.

DISCUSSION

Market outcomes starting in 2014 and going back 10 years offer no evidence women are at a disadvantage in tenure-track competitions. The same can be said for the other objective measures that were examined including publishing and the reputations of home and hiring departments. No statistically significant evidence that pervasive dysfunction in departmental cultures is harming early career market outcomes of budding women philosophers could be found. Meanwhile, the biggest drop in women's participation appears to occur almost immediately, right after first exposure to philosophy's themes, methods, and traditions (Adelberg et al., 2013; Dougherty et al., 2015). Although evidence that the gender gap in philosophy is attributable to pre-university influences has been available since at least 2012 (Paxton et al., 2012) the present study adds to the case against the hypothesis that sexist attitudes (whether conscious or unconscious) held by philosophers are a major cause of disproportion according to gender.

All the same, we can be somewhat reticent to draw strong conclusions about the extent of philosophy's climate problems, and it might be premature to say that there is no systemic anti-female prejudice. Bias that was present but somehow neutralized

by measures departments have taken or coping strategies adopted by women might have been overlooked. Then again it seems doubtful that explicit policy changes and coping strategies were adopted more than 10 years ago, long before there was wider awareness of the issue of unconscious bias. It is also conceivable that bias shows up elsewhere, affecting outcomes for tenure and promotion, though keep in mind this conjecture is not supported even by mainstays of the implicit bias literature, such as Steinpreis et al. (1999) whose name-swapping experiments found "no main effects" for tenurability. The present findings are a better fit with the strong preference for women in STEM found by experimental manipulations (Williams and Ceci, 2015).

While counter thoughts are not to be dismissed lightly, the hypothesis that unconscious bias works against women in hiring and early career publishing is not well supported. Although it is conceivable implicit bias initially reduces perception of a woman's cv. and then "affirmative actors" reverse its impact, this proposal strikes one as overly complicated: why not just assume people are not downgrading the accomplishments of talented women?

The suggestion that there is a shyness effect making bias hard to detect is also hard to square with the evidence about pre-market publishing opportunities. Why doesn't bias reveal itself in disparities for special invitations to publish where there are no equity policies or structures, little to no collegial oversight, and it is hard to conceive of coping strategies? We should also worry that efforts to improve the representation of women could even backfire, e.g., if committees adopted blind review of candidates under the dubious assumption that more accomplished women are systematically undervalued.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

The author is indebted to the referees, Calvin K. Lai, Shen-yi Liao, Paul Draper, Neb Kujundzic, Matthew Mosdell, Monika Piotrowska, and Louis-Philippe Hodgson, as well as audiences at the Atlantic Region Philosophers Association, the University of Louisville, and the University of Miami for their insightful feedback and suggestions.

REFERENCES

- Adelberg, T., Thompson, M., and Nahmias, E. (2013). Women and philosophy: why is it 'Goodbye' at 'Hello?' *Paper Presented at the Society for Philosophy and Psychology*, Providence, RI.
- Dougherty, T., Baron, S., and Miller, K. (2015). *Female Under-Representation Among Philosophy Majors: A Map of the Hypotheses and a Survey of the Evidence*. Available at: <http://ir.lib.uwo.ca/fpq/vol1/iss1/4>
- Forscher, P. S., Lai, C., Axt, J., Ebersole, C. R., Herman, M., Nosek, B. A. et al. (2017). *A Meta-Analysis of Change in Implicit Bias*. Available at: <https://osf.io/awz2p/>
- Krishnamurthy, M., Liao, S., Deveaux, M., and Dalecki, M. (forthcoming). The underrepresentation of women in prestigious ethics journals. *Hypatia*.
- Paxton, M., Figdor, C., and Tiberius, V. (2012). Quantifying the gender gap: an empirical study of the underrepresentation of women in philosophy. *Hypatia* 27, 949–957. doi: 10.1111/j.1527-2001.2012.01306.x

- Saul, J. (2012). Ranking exercises in philosophy and implicit bias. *J. Soc. Philos.* 43, 256–273, 268, 276. doi: 10.1111/j.1467-9833.2012.01564.x
- Saul, J. (2013). "Implicit bias, stereotype threat and women in philosophy," in *Women in Philosophy: What Needs to Change?* eds F. Jenkins and K. Hutchison (Oxford: Oxford University Press), 39–60. doi: 10.1093/acprof:oso/9780199325603.003.0003
- Sesardic, N., and De Clercq, R. (2014). Women in philosophy: problems with the discrimination hypothesis. *Acad. Quest.* 27, 461–473. doi: 10.1007/s12129-014-9464-x
- Solomon, M., and Clarke, J. (2009). CSW jobs for philosophers employment study. *Proc. Address. Am. Philos. Assoc.* 8, 3–6.
- Steinpreis, R., Anders, K., and Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: a national empirical study. *Sex Roles* 41, 509–528. doi: 10.1023/A:1018839203698
- Williams, W. M., and Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure-track. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5360–5365. doi: 10.1073/pnas.1418878112
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Allen-Hermanson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

Raw data is available at genderandphilosophy.blogspot.com

Are recent cohorts of women with engineering bachelors less likely to stay in engineering?

Shulamit Kahn^{1*} and Donna K. Ginther²

¹ Markets, Public Policy and Law Department, Questrom School of Business, Boston University, Boston, MA, USA,

² Department of Economics and Center for Science, Technology and Economic Policy, University of Kansas, Lawrence, KS, USA

OPEN ACCESS

Edited by:

Jessica S. Horst,
University of Sussex, UK

Reviewed by:

Beth Livingston,
Cornell University, USA
Katherine Micheltore,
University of Michigan, USA

*Correspondence:

Shulamit Kahn,
Markets, Public Policy and Law
Department, Questrom School of
Business, Boston University, 595
Commonwealth Avenue, Boston,
MA 02215, USA
skahn@bu.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 14 April 2015

Accepted: 23 July 2015

Published: 19 August 2015

Citation:

Kahn S and Ginther DK (2015) Are
recent cohorts of women with
engineering bachelors less likely to
stay in engineering?
Front. Psychol. 6:1144.
doi: 10.3389/fpsyg.2015.01144

Women are an increasing percentage of Bachelors in Engineering (BSEs) graduates—rising from 1% in 1970 to 20% in the 2000s—encouraged by increasing K-12 emphasis on attracting girls to STEM and efforts to incorporate engineering and technology into K-12 curricula. Retention of women in STEM and in engineering in particular has been a concern historically. In this paper, we investigate whether this gap has increased because a larger proportion of females entering engineering find themselves ill-matched to this field, or whether the gap has decreased as engineering becomes more accommodating to women. Using 1993–2010 nationally representative NSF SESTAT surveys, we compare cohorts of BSEs at the same early-career stages (from 1–2 to 7–8 years post-bachelors). We find no evidence of a time trend in the gender gap in retention in engineering and a slightly decreasing gender gap in leaving the labor force. We find, as others have, that the majority of the gender retention gap is due to women leaving the labor force entirely and that this exit is highly correlated with child-bearing; yet women with engineering majors are half as likely as all college-educated women to leave the labor market. There are no clear time trends in female BSEs leaving the labor market. Single childless women are actually more likely than men to remain in engineering jobs. Some of the gender differences in retention we find are caused by differences in race and engineering subfield. With controls for these, there is no gender retention difference by 7–8 years post-bachelors for those full-time employed. There were two unusual cohorts—women with 1991–1994 BSEs were particularly likely to remain in engineering and women with 1998–2001 BSEs were particularly likely to leave engineering, compared to men. Cohorts before and after these revert toward the mean, indicating no time trend. Also, women who leave engineering are just as likely as men to stay in math-intensive STEM jobs.

Keywords: engineering careers, gender, leaving STEM, women engineers, retention

Introduction

Engineering has been and continues to be a field dominated by men. However, the percentage of women getting bachelors in engineering (BSE) has grown dramatically over the decades, from approximately 1% in 1970 to 10% in 1980, 15% in 1990 and stabilizing near 20% in the 2000s (NSF WebCASPARE). This has been a period of consciousness-raising

about the paucity of women in STEM fields, of rising math test scores among K-12 girls, of more girls taking high school math and science courses, and of women increasing their general college attendance relative to boys (Ceci et al., 2014). **Figure 1** illustrates the growth in representation of females among BSE and in other STEM fields.

There has been considerable interest and research on women's retention in STEM in general, and in engineering in particular. Most recently, the National Academy of Engineering and National Research Council (National Academy of Engineering and National Research Council, 2014) convened a workshop on this topic. Women working in the engineering profession, represented by the Society of Women Engineers (SWE), have been very active in surveying women in their field to understand women's greater exit rates. In addition to the Society of Women Engineers (2009) and the National Academy of Engineering and National Research Council (2014) studies, Morgan (2000); Hunt (2010), and Singh et al. (2013) addressed women's exit from engineering particularly, while work such as Preston (1994, 2004), Xie and Shauman (2003), and Xu (2008) addressed exit more generally in all STEM fields.

Previous work on women's retention in engineering was primarily based on cross-sectional data which combines people from many different cohorts (which we identify here by the year of their bachelor's degree in engineering). Measuring retention at different career stages in a cross-section actually combines differences across career stages and differences across cohorts. For instance, in a 2010 cross-section, the only people who would be observed 1 year from their bachelor's degree are the millennials who graduated in 2009, and the only ones who would be observed

at 20 years from their bachelors are the Gen X'ers who graduated in 1990.

There are reasons to believe that recent cohorts of engineering majors may behave differently than earlier cohorts did when they were the same age. On the one hand, we might expect later cohorts of women to be more likely to remain in the field because of women's increasing representation among engineering graduates. Hunt (2010) has shown that scientific fields of study with lower female percentages tend to have higher exit rates of women from the field.

On the other hand, we might expect that later cohorts of women engineering majors will be less likely to remain in the field than earlier cohorts. Women might be majoring in engineering in greater numbers because high school curricula have increasingly included engineering and computer education and educators have been encouraged to attract women to engineering (Carr et al., 2012). It may be that some of the women choosing engineering majors today might be less well-matched to the occupation and not find working in engineering satisfying. Therefore, a larger proportion of later cohorts of engineering BSE women may leave engineering after they have spent a few years working in non-engineering fields. In a similar vein, the recent National Academies conference report indicates that excessive workloads, unclear expectations, lack of work-life balance, and a "chilly climate" were associated with women leaving engineering (National Academy of Engineering and National Research Council, 2014). If it is the case that recent cohorts of BSE women are less well-matched to the engineering occupation, these climate issues may increase

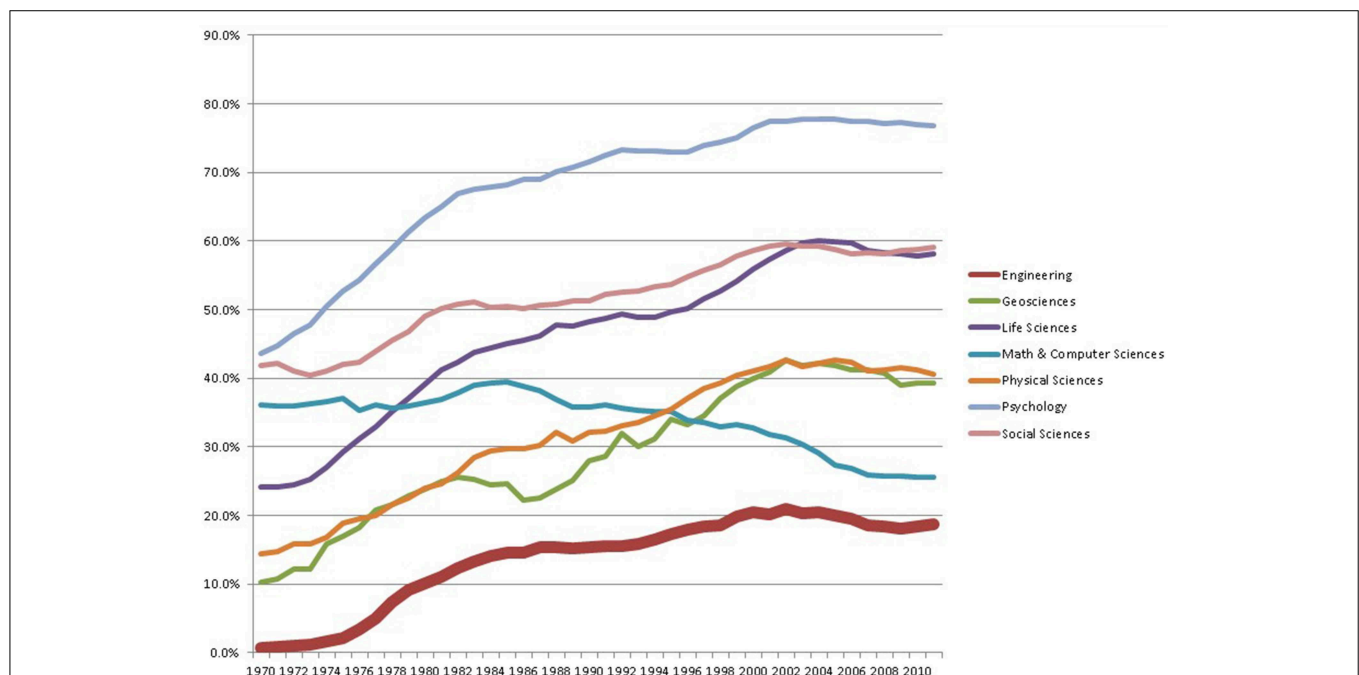


FIGURE 1 | Percent female among bachelors in engineering compared to other STEM fields over time. Data Source: NSF WebCASPAR data base (<https://ncesdata.nsf.gov/webcaspar/>).

the propensity of women in more recent cohorts to leave engineering.

These possibilities suggest that we should compare different cohorts of BSE, particularly during the first years after they graduate. We compare whether recent cohorts of women with a BSE leave the engineering field with greater or lesser frequency than previous cohorts. We also test whether there is a general time trend in gender retention differences over the last few decades. Along with this, we might expect that those women who leave engineering might move to non-math-intensive occupations with greater proportions of women.

We test these hypotheses using NSF longitudinal SESTAT data that allow us to study cohorts as recent as 2009 bachelors in engineering (BSE) and as early as those with BSEs in 1985. We use data from eight different waves of the same survey spread over 18 years (1993–2010), allowing us to tease apart differences across cohorts from differences in retention that occur as careers develop, and to further to identify whether the career pattern is different across the cohorts. Moreover, given the panel nature of these surveys, we can follow specific individuals longitudinally for periods as long as 8 years which gives us a better sense of the timing of exit.

Previous Research

Preston (1994, 2004), Xie and Shauman (2003), Xu (2008), and Glass et al. (2013) have studied women's exit from science and engineering as a whole using a variety of national data sources. Preston found large differences in the 1980s and early 1990s. Using data from the 70s, 80s, and early 90s, Xie and Shauman found that women with bachelors in STEM (excluding social sciences) are about one quarter less likely than men to work in STEM occupations and that married women with children are the most affected. Xu (2008), using the 1999 National Survey of Postsecondary Faculty, found that women and men were equally likely to seek to leave STEM academic careers but that women had greater intentions to seek another position within academia. Glass et al. (2013) followed female college graduates from the National Longitudinal Survey of Youth 1979 and found that women in STEM occupations were more likely to leave their field early in their career compared with women in other professional occupations. They find that women in STEM occupations move to non-STEM occupations at very high rates and attribute women's departure from STEM careers to climate issues or job matching.

Research on gender differences in retention in engineering specifically are most germane to this paper. The Society of Women Engineers (2009) surveyed engineering alumni of 21 colleges from 1985 and later. In their 2005 cross section of graduates from these 21 schools whose BSE was their highest degree, there was an average 10% gender gap in the likelihood of working in engineering. Further, they found that 90% of this gender gap was a result of women leaving the labor force entirely. These gender differences were similar to those from the more nationally representative 2003 NSF SESTAT, although overall their retention rates were higher than those in SESTAT.

Morgan (2000) used the 1993 National Survey of College Graduates (NSCG) and captured employment of those who received BSEs between 1965 and 1989 but measured the gap only for those with highest degrees in engineering (i.e., only those who did not choose immediately post-bachelors to enter into a different field via a degree). As such, her estimate of exit is likely to be lower than ours. She found a 3 percentage point (*ppt.*) gender gap in the likelihood that full-time workers with highest degrees in engineering were employed in engineering jobs, defined using a survey question asking whether respondents were working in a field closely or somewhat related to their field of highest degree. In contrast, women in other fields were 6 *ppt.* more likely than men to remain in the field of their highest degree. She also found these women were 9 *ppt.* more likely than men to be out of the labor force and 7 *ppt.* more likely to be working part-time.

Hunt (2010) also uses the NSCG, but from both the 1993 and 2003 surveys. Like Morgan, she studied those with *highest degrees* in engineering and based her analysis on the question of how closely their job related to the field of highest degree. Hunt found about a 10% average gender difference in overall retention¹, of which 70% could be accounted for by women leaving the labor force (similar to Morgan's 3% gender gap among full-time workers). Also like Morgan (2000), Hunt found that the gender differences in engineering were slightly larger than gender differences in other sciences or in non-STEM fields. Unlike Morgan (2000) and Society of Women Engineers (2009), Hunt estimated gender differences with regression models allowing her to control for field, age, degree level, and race among other factors. Holding these constant, women who studied engineering were slightly more likely than women in other fields to be working (about 1 *ppt.*) but considerably less likely than women in other fields to have a job related to her highest degree (on the order of 5 *ppt.* of those working or about 4 *ppt.* of those irrespective of whether they worked). Finally, Hunt finds that including the male share of the field in the regression model that estimates female exit more-than-explains the lower female retention of women in engineering compared to other non-STEM fields.

The only research using longitudinal data to examine retention in engineering was Greenfield's presentation in National Academy of Engineering and National Research Council (2014), which used data from the Department of Education's Baccalaureate and Beyond. She primarily analyzed the 1992–1993 BSE cohort whose sample was small (560, with 80 women). She measured retention as working in an engineering or architecture job. She found that retention rates for employed women engineering BSEs were *higher than* those of men's in engineering—13.7 *ppt.* higher after 1 year, 14.8 *ppt.* higher than men after 4 years, and 6.8 *ppt.* higher after 10 years. The retention rate of women in engineering was not lower than other fields at 4 years, but was substantially lower at 10 years. She also looked at 1-year retention rates across several later cohorts and found

¹This number is computed from Hunt's figures although she herself did not make this calculation.

that later cohorts of women were less likely to stay in engineering immediately after receiving their bachelors.

A final relevant finding in Hunt (2010) is that the share of men in the specific sub-field of STEM study was positively highly correlated with women's exit from science ($r = 0.51$). She finds that including the male share of the field in a regression of female exit more-than-explains the lower female retention of women in engineering compared to other non-STEM fields.

Thus, all of these studies find gender differences in retention in engineering that are small relative to the percentages who stay in engineering, contradicting the general impression of much higher exit rates from engineering (e.g., see Singh et al., 2013). One (National Academy of Engineering and National Research Council, 2014) even found women more likely to remain in engineering.

Materials and Methods

SESTAT is collected by the National Science Foundation (NSF) and is the most comprehensive database on the employment, educational, and demographic characteristics of U.S. scientists and engineers available. SESTAT actually includes observations from three NSF surveys: the National Survey of Recent College Graduates (NSRCG), the National Survey of College Graduates (NSCG) and the Survey of Doctorate Recipients (SDR). From the NSCG respondents, SESTAT includes only those who received a degree in STEM or had ever worked in a STEM occupation. From the NSRCG, SESTAT includes recent bachelor's and master's degree recipients in STEM fields. The SDR samples US-awarded PhDs in STEM disciplines. SESTAT oversamples women and under-represented minorities (URMs) in order to allow more accurate measures of gender and racial differences.

Within each decade, SESTAT followed individuals through the different waves, adding new people to represent more recent graduates (from the NSRCG). The 1990s panel includes 4 waves: 1993, 1995, 1997, and 1999. The 2000s panel includes 4 waves: 2003, 2006, 2008, and 2010. SESTAT thus includes as many as four observations on a single individual over a 7 or 8 year span in each decade (although for various reasons many people are seen for fewer than four surveys²). Note that there are primarily 2-year gaps between survey waves, although there is one 4-year and one 3-year gap.

SESTAT collects information on education, employment including labor force status, occupation, employer characteristics, work activities, and comprehensive demographic information on gender, race/ethnicity, marital status, children, citizenship, and immigration status.

The measurement of who is working as an engineer is not straightforward. In the majority of this analysis, we define people as working in engineering if their primary occupation is categorized by the NSF as engineering. This excludes all jobs categorized as computer and information scientists, such as computer system analysts. Moreover, it excludes jobs categorized

as being engineering-related, such as "electrical, electronic, industrial, and mechanical technologists and technicians" or architects. Based on the 2010 SESTAT, we calculate that 1.73 million people were employed full-time in engineering jobs, 2.31 million in computer jobs, and 0.46 million in engineering-related jobs. Beginning in 2003, SESTAT began including low to mid-level "engineering managers" within engineering occupations, but not "top level managers, executives, and administrators." "Engineering managers" (or managers, a term we have coined) represented 15.6% of the 1.73 million full-time engineering jobs in 2010. Because we want to compare cohorts working in the 1990s as well as the 2000s, we exclude engineering managers in our analysis of engineering retention across cohorts. That said, we also analyze whether BSEs moved into management jobs and if so, whether the job was required technical STEM education.

We use the SESTAT data to examine gender differences in remaining in engineering by cohort and years since degree. Our cohort analysis is based on the 28,117 individuals in SESTAT surveyed who received their first bachelor's degree in engineering (BSE)³ between 1985 and 2009. For ease of presentation, we divide cohorts into approximately 3- to 5-year BSE groupings starting with the 1985–1990 cohort and ending with the 2006–2009 cohort, choosing endpoints so each cohort has enough observations to create reasonably accurate statistics. Individuals in the analysis were observed in a SESTAT survey at either 1–2 years, 3–4 years, and/or 7–8 years post-BSE. We also examine outcomes for people working 15–16 years after the degree, but the number of women in this older cohort is small.

We begin our cohort analysis using descriptive statistics to examine gender differences in remaining in engineering by years since PhD for the outcomes of (1) being "engaged in engineering," defined as working in an engineering occupation or enrolled in an advanced engineering degree program⁴; (2) working full-time in an engineering occupation for the subsample that is employed 35 or more hours per week; and (3) being out of the labor force—defined as not working and not looking for work. We then use linear probability regressions to estimate gender differences in these same outcomes, controlling for things that might be responsible for gender differences but that are not directly attributable to gender *per se*, including engineering subfield, survey year, immigrant status, race, and one measure of socioeconomic class, whether the parent had graduated college. We present the coefficient on gender from these models in order to examine differences in remaining in engineering across cohorts. We then take a closer look at factors associated with leaving the labor force by adding interaction terms to our linear probability models, specifically interaction terms for female X cohort X family-status. Finally, for those who leave engineering, we examine where they go—to engineering related, other mathematically intensive STEM, non-mathematical STEM, or non-STEM occupations.

³We limit this analysis to first bachelors because we are interested in those who originally chose engineering as a field in college, not those who came to it later. Also, those for whom the engineering BS is not their first bachelors degree may be at a different career stage. The vast majority of BSEs are first bachelors.

⁴After a handful of years from the BSE when some complete a masters, there is no distinction between engaged in and working in engineering.

²This includes: adding new cohorts; adding others if needed to balance others who had dropped out of the survey; non-response for one wave (people were dropped if they did not respond for 2 waves); aging out at age 75 etc.

Stata 13 was used for all statistical analysis including the linear probability multiple regression models. The paper only includes those results related to gender differences. Full regression results for all regression tables are available in the Supplementary Material.

Results

Average Gender Differences in Retention Post-bachelors 2010 Averages

Figure 2 shows the proportion of women and men, respectively with BSEs who in 2010 are “engaged in engineering” graphed by years since the BSE. We use 3-year moving averages because of the erratic periodicity of SESTAT surveys and the small number of females at each point. **Figure 2** demonstrates the starting point of this paper, that in the 2010 cross-sectional data, after a few years post-BSE a gap appears and women with BSEs become less likely to be working in engineering jobs than men. The average gender difference in remaining in engineering (for those within 30 years of the BSE) is 7.8 percentage points (or *ppt.*) At 10 years post-bachelors, the gender difference is 8.2 *ppt.*; at 20 years, it is 15.5 *ppt.* and at 30 years, it is 10.4 *ppt.* We note, however, that the sample size of women engineers who in 2010 were more than 18 years post-BSE is very small (<100 individuals per year), so the right-hand side of the graph must be considered only suggestive.

Some of the gender difference in engineering retention may simply be due to the fact that more women than men are not working at all (either unemployed or out of the labor force) or working part-time. Among those in the 2010 SESTAT within 30 years of their BSE, 19.2% of women but only 5.6% of men were not working, a difference of 13.6 *ppt.* The 2010 percentage of women not working among BSEs is similar to the 20.0% not working in 2010 among all US women with a bachelors or higher⁵.

Moreover, rather than leave the labor force, some people instead choose to work part-time. In 2010, 5.7% of those with

BSEs in engineering (within the past 30 years) worked part-time. There is a large gender difference in the likelihood of working part time (as would be expected if women are the primary child-caregivers): 12.7% of women with BSEs but only 4.1% of men were working part-time.

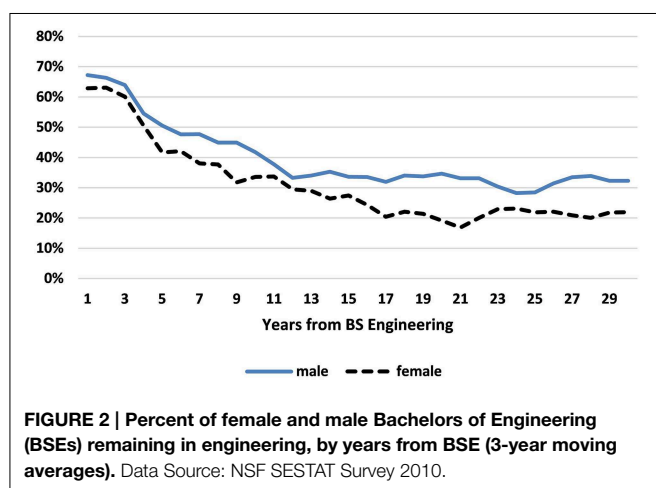
Two facts suggest that there are fewer part-time jobs available within engineering than are desired by BSEs. First, 32.4% of women with BSEs who worked part-time were in engineering jobs compared to 38.5% of women with BSEs who worked full-time. Second, only 5.7% of all those with a BSE work part-time, much less than the 14.4% working part-time of those with non-engineering STEM bachelors. This suggests that if a person with a BSE wants to work part-time, she/he is much more likely to be forced to work outside of engineering. This paucity of part-time jobs within engineering may be due to choices made by employers insensitive to women’s flexibility needs, a point we discuss in the conclusion.

Including only those BSE’s working full-time eliminates 32.4% of female BSEs compared to 10.3% of male BSEs. The average gender difference in remaining in engineering among full-time-working BSEs (2010, first 30 years) is 1.6 *ppt.*, much less than the 7.8 *ppt.* average for the entire population.

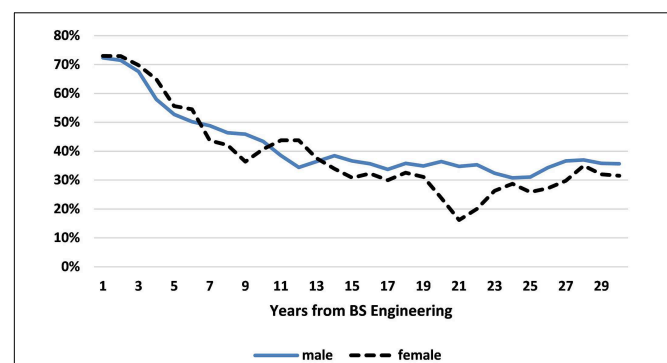
Figure 3 includes only those BSEs who are working *full-time* and graphs the percent in engineering for men and women separately. We see that in the 15 years after their undergraduate diploma, on average men and women are *equally likely* to remain in engineering, with periods when women are more likely than men to do so. Beyond 15 years post-BSE, however, men are consistently more likely to remain in engineering, with the gap fluctuating considerably due to even smaller sample sizes of full-time working women than in **Figure 2**.

1993–2010 Averages

As noted earlier, using a single SESTAT year (2010) confuses cohort and career stage differences. Instead, we use the data from all 8 SESTAT waves from 1993 to 2010 to measure the gender retention gap at three different early career stages (measured by years from BSE): 1–2 years after their bachelors, 3–4 years after their bachelors, and 7–8 years after their bachelors. We use 2-year career-stage spans because in most cases, SESTAT surveys were



⁵ Also within 30 years of their bachelors. Calculated by the authors from the Bureau of the Census’s American Community Survey.



administered every 2 years (We also do limited analyses for the stage 9–16 years post-BSE).

Table 1 gives the average probability that men and women remain in engineering (either working or getting higher degrees) at the three different career stages averaging over individuals in the sample observed at this career stage. Before we discuss cohort-specific gender retention, we first describe this average retention at each career stage using both descriptive statistics (**Table 1**) and regression analysis (**Table 2**).

The first row of **Table 1** tells us that 61% of both male and female BSEs enter an engineering job (or schooling) in the 1–2 years immediately after graduating with a BSE, 39% do not. There is no (significant) gender difference. By 3–4 years post-BSE, a gender difference had appeared, where women were 3.6 percentage points (*ppt.*) less likely than men to remain in engineering; and by 7–8 years, this gender difference had widened to 8.3 *ppt.* Columns 4 through 6 include only those working full time. Since women are more likely than men to leave the labor force as well as more likely to work part-time, excluding these two groups from the population (as well as the unemployed⁶) changes

the gender difference considerably at all career stages. At 1–2 years, those women working full-time were significantly *more* likely than men (3.1 *ppt.*) to remain in engineering on average; at 3–4 years men and women were insignificantly different; and only by 7–8 years were women less likely to remain in engineering, with a significant gender difference of 3.0 *ppt.*

The last three columns confirm that at each career stage, on average females are more likely than men to be out of the labor force completely, but that the main movement out of the labor force occurs between 4 and 8 years of the BSE.

Regression Analyses of Average Retention

Table 2 uses linear probability regressions to calculate these same measures at the same three career stages, controlling for engineering subfield, survey year, immigrant status, race, and one measure of socioeconomic class, whether the parent had graduated college.

We highlight only those **Table 2** results that are qualitatively different from what was found in the simple descriptive statistics of **Table 1**. Compared to **Table 1**, at 3–4 years post-BSE, the addition of controls erased the gender difference for the population as a whole (Neither table finds a gender difference

⁶Unemployment rates of BSE engineers are similar for men and women.

TABLE 1 | Average probability of remaining in engineering (working or studying) or out of the labor force: all cohorts combined.

	% of all BSE grads engaged in engineering			% of BSE grads working FT in engineering			% Out of the Labor Force			# Observations	
	Male (%)	Female (%)	Female-male difference (%)	Male (%)	Female (%)	Female-male difference (%)	Male (%)	Female (%)	Female-male difference (%)	Male	Female
1–2 years post-BSE	61.38	60.54	–0.84	67.75	70.82	3.07***	5.55	8.41	2.86***	12162	4695
3–4 years post-BSE	61.35	57.79	–3.56***	65.95	66.45	0.50	4.39	6.76	2.37***	10733	3773
7–8 years post-BSE	53.58	45.33	–8.26***	56.04	53.00	–3.04**	1.77	10.30	8.53***	9205	2607

Gender difference *t*-test ****p* < 0.01, ***p* < 0.05. 15–16 averages cannot be given because the #observations in some cases are too small to report.

TABLE 2 | Coefficient on female from linear probability models of remaining in engineering: all cohorts combined.

	Probability of remaining in engineering		Probability of leaving the labor force
	Population: all	Population: working FT	Population: all
1–2 years post-BSE	0.0127 (0.0094)	0.0362*** (0.0100)	0.0073 (0.0048)
3–4 years post-BSE	–0.0163 (0.0102)	0.0185* (0.0108)	0.0102** (0.0047)
7–8 years post-BSE	–0.0620*** (0.0119)	–0.0092 (0.0131)	0.0834*** (0.0044)
15–16 Years post-BSE if still in Eng at 7–9 years	–0.0072 (0.0474)	0.0905* (0.0508)	0.1053*** (0.0159)

Coefficient significance ****p* < 0.01, ***p* < 0.05, **p* < 0.1.

Standard errors in parentheses.

Controls include dummies for engineering subfield, survey year, BSE year, if parent had ≥BA/BS, immigrant status, race.

#obs: All population: 1–2 years: 16,857; 3–4 years: 14,506; 7–8 years: 11,812; 15–16 years: 884.

#obs: FT only: 1–2 years: 13,382; 3–4 years: 12,501; 7–8 years: 10,585; 15–16 years: 848.

retention disadvantage for full-time workers at this stage). At 7–8 years, for the whole population, what was an 8.3 *ppt.* gender difference in **Table 1** becomes 6.2 *ppt.* with controls (**Table 2**); in contrast, among those working full time, there is no longer a significant gender difference. Finally, with controls, gender differences in being *out of the labor force* (**Table 2**) are somewhat smaller than without controls (**Table 1**) and no longer significant at 1–2 years. Overall, then, the control variables do explain some of the gender differences observed in the descriptive statistics. In work not shown, we investigated which of the controls variables were the major mediating factors. We found that subfield was one important factor but that race/ethnicity was the most important control variable responsible for some of the average gender gap⁷. Women in engineering are less likely than men to be white (non-Hispanics)—the race with the highest retention rates—and more likely to be Asian or black, both groups with lower retention rates. This result suggests that racial retention rates are important to study in future research.

The last row models retention at an even later career stages by asking, “Of those who remain working in engineering 7–8 after their degree, what is the gender difference in the likelihood

of remaining in engineering approximately 8 years later?”⁸ This allows us to incorporate BSEs as early as 1984, even though the earliest BSEs we can observe at their careers’ beginning are from 1991⁹. This row indicates that there was no significant gender retention difference during years 8–16 among those people who were still in engineering at the beginning of this stage. When we look only at those who are still full-time employed at year 15–16 post-BSE, on average women are *more* likely than men to remain in engineering.

Differences across Cohorts

Tables 3,4 present gender differences for cohorts defined by narrow ranges of BSE years. **Table 3** gives averages per cohort/gender, while each panel of **Table 4** gives coefficients from a linear probability regression run with interaction terms between the female dummy variables and a dummy variable for each cohort, as well as on other control variables.

We cannot compare exactly the same cohorts across all career stages, for two reasons. First, the latest BSE years are only observed in their first career stages, while the earliest BSE years are only seen in their later career stages. Second, we lose

⁷Thus estimating the gender gap at 7–8 years from BSE, controlling for race variables alone made the gender coefficient fall. Our race variables are defined as follows: We separated out non-black Hispanics and we combined black with other under-represented races such as Native American. Asians were a separate category. There were no gender differences in the percentage of men and women who were Hispanic.

⁸We use a range for beginning and end points because of the spacing of SESTAT surveys. To further increase our sample size, if someone was not observed in years 7 or 8 but was observed in year 9 still in engineering, we also include them in this panel.

⁹Analysis for 1984 BSEs uses SESTAT 1993 for the 9-year point and SESTAT 1999 for the 15-year point. Analysis of 1995 BSEs uses SESTAT 2003 and 2010 for the 8 and 15 year points, respectively. Those with 1985, 1986, 1989, and 1993 BSEs could not be observed at both career points so are not included in the Panel D analysis.

TABLE 3 | Average probability of remaining in engineering (working or studying) or out of the labor force by BSE year cohort.

Cohort (BSE years)	% of all BSE grads engaged in engineering			% of BSE grads working FT in engineering			% Out of the Labor Force			# Observations	
	Male (%)	Female (%)	Female-male difference (%)	Male (%)	Female (%)	Female-male difference (%)	Male (%)	Female (%)	Female-male difference (%)	Male	Female
(A) 1–2 YEARS POST-BSE											
1991–1994	57.34	65.94	8.59***	63.99	73.64	9.65***	4.20	4.28	0.07	4601	1077
1995–1997	62.89	60.48	–2.41	68.47	68.77	0.31	4.53	6.45	1.92*	2237	663
1998–2001	62.04	57.08	–4.96*	69.59	69.95	0.35	5.18	11.79	6.61***	1362	546
2002–2005	59.47	55.95	–3.51	66.87	66.46	–0.42	6.79	9.78	2.98**	1957	886
2006–2009	64.86	62.06	–2.80	70.45	74.18	3.73	6.82	10.06	3.24***	2005	1523
(B) 3–4 YEARS POST-BSE											
1989–1990	62.04	58.22	–3.82	68.44	68.75	0.31	4.74	7.66	2.92***	2526	561
1991–1994	61.94	67.55	5.61**	66.76	75.38	8.62***	4.65	5.18	0.52	2575	598
1995–1997	60.20	57.31	–2.89	63.78	62.26	–1.52	3.66	3.94	0.27	1104	328
1998–2001	60.39	53.18	–7.21**	64.86	63.34	–1.52	2.75	6.08	3.34***	933	366
2002–2005	61.51	53.45	–8.05***	66.65	62.04	–4.61**	5.07	8.93	3.85***	2510	1336
2006–2007	61.13	57.91	–3.22	64.12	68.10	3.97	4.36	6.77	2.42*	1085	584
(C) 7–8 YEARS POST-BSE											
1985–1990	56.75	49.14	–7.61***	59.13	58.14	–0.99	1.69	12.40	10.71***	4607	957
1991–1994	54.56	57.90	3.33	56.64	65.84	9.20**	1.88	9.94	8.06***	996	253
1995–1997	49.66	42.52	–7.15*	52.97	55.38	2.42	1.81	11.90	10.09***	919	234
1998–2001	56.20	43.19	–13.01***	59.44	49.25	–10.19***	1.84	8.92	7.07***	1763	789
2002–2003	44.93	38.80	–6.13*	45.57	43.67	–1.90	1.74	8.14	6.40***	920	374

Gender difference *t*-test ****p* < 0.01, ***p* < 0.05, **p* < 0.1.

TABLE 4 | Gender differences in remaining in engineering or leaving the labor force by cohort (calculated as the coefficient on female–cohort interaction from a linear probability regression at each stage).

Cohort (BSE years)	Probability of Remaining in Engineering		Probability of Leaving the Labor Force
	Population: All	Population: Working FT	Population: All
(A) 1–2 YEARS POST-BSE			
1991–1994	0.1049*** (0.0203)	0.1140*** (0.0210)	−0.0102 (0.0103)
1995–1997	−0.0074 (0.0206)	0.0026 (0.0213)	0.0036 (0.0105)
1998–2001	−0.0303 (0.0242)	0.0143 (0.0260)	0.0414*** (0.0123)
2002–2005	−0.0191 (0.0198)	−0.0097 (0.0214)	0.0125 (0.0101)
2006–2009	0.0012 (0.0178)	0.0430** (0.0194)	0.0032 (0.0090)
(B) 3–4 YEARS POST-BSE			
1989–1990	−0.0327 (0.0308)	0.0074 (0.0334)	0.0249* (0.0140)
1991–1994	0.0717*** (0.0222)	0.1014*** (0.0233)	−0.0017 (0.0101)
1995–1997	−0.0184 (0.0292)	−0.0044 (0.0303)	−0.0039 (0.0133)
1998–2001	−0.0570** (0.0277)	−0.0066 (0.0295)	0.0230* (0.0126)
2002–2005	−0.0552*** (0.0188)	−0.0299 (0.0198)	0.0183** (0.0086)
2006–2007	−0.0135 (0.0252)	0.0434 (0.0272)	0.0012 (0.0115)
(C) 7–8 YEARS POST-BSE			
1985–1990	−0.0574*** (0.0206)	0.0141 (0.0231)	0.1057*** (0.0075)
1991–1994	0.0696* (0.0396)	0.1213*** (0.0429)	0.0781*** (0.0145)
1995–1997	−0.0682** (0.0323)	0.0226 (0.0369)	0.0987*** (0.0118)
1998–2001	−0.1201*** (0.0221)	−0.0926*** (0.0241)	0.0695*** (0.0081)
2002–2003	−0.0390 (0.0275)	−0.0035 (0.0299)	0.0568*** (0.0101)
(D) FROM 9–16 YEARS POST-BSE IF STILL IN ENGINEERING AT 7–9 YEARS			
1984	−0.1313** (0.0654)	0.0310 (0.0731)	0.1833*** (0.0218)
1987–1994	−0.0623 (0.0978)	−0.0058 (0.0997)	0.0521 (0.0326)
1995	0.3289** (0.1287)	0.2708** (0.1324)	−0.0504 (0.0429)

Controls include dummies for engineering subfield, survey year, BE year, if parent had \geq BA/BS, immigrant status, race.

Because of the irregular SESTAT periodicity, the following intermediate BE years are not in the data.

(A) 1999, 2000, 2003; (B) 1997, 1998, 2001; (C) 1993, 1994, 1997; (D) 1989, 1993.

#obs: All population: (A) 16,857; (B) 14,506; (C) 11,812; (D) 884.

#obs: FT only: (A) 13,382; (B) 12,501; (C) 10,585; (D) 848.

some BSE years when SESTAT did not have the standard 2-year periodicity¹⁰. Specifically, we do not observe those with BSEs in 1999, 2000, or 2003 at the 1–2 year mark, we do not observe those with BSEs in 1997, 1998, and 2001 at the 3–4 year mark, and we do not observe those with BSEs in 1993, 1994, and 1997 at the

7–8 year mark. In the analysis of the 8 to 16 year career stage, we have information about even fewer cohorts since the cohorts need to be observed both at the 7–9 year point (to see if they start the stage in engineering) and again at the 15–16 year point, meaning the last observed cohort have 1995 BSEs.

In addition, we have estimated linear probability models with single-year cohorts (Table A1 in Supplementary Material). Since

¹⁰Recall that SESTAT skips from 1999 to 2003 and then to 2006.

each annual cohort sample is small, the majority of single-year-cohort gender gaps are not significantly different from zero. Nevertheless, this analysis does help us to analyze whether our arbitrary cohort definitions hid large variation within multi-year cohorts. The Supplementary Table A1 gender gap coefficients for the whole population are graphed as **Figure 4**. Our discussion below will primarily be based on the multi-year cohorts of **Tables 3, 4**; however, we refer to Table A1 in Supplementary Material analysis when results on gender differences in single years adds to our understanding.

Cohort Differences at 1–2 Years

In our earlier discussion of the averages across all cohorts, we found no differences in the retention of women and men in engineering in the first 2 years post-BSE receipt, with or without controls. There was a significant but modest difference in women leaving the labor force that seemed to be due to race and subfields. Among who were working full time, however, women were actually significantly more likely to remain in engineering than men at this stage (with and without controls).

This same pattern is not shared by all cohorts. For four out of the five cohorts—all those with 1995 to 2009 BSEs—the estimated average differences (**Table 3** first columns) suggest that women were less likely than men to remain in engineering at this early career stage. While this difference was only significant for one cohort (those with BSEs 1998–2001), if we combined the four cohorts 1995–2009, the overall gender difference is highly significant ($p = 0.001$). Adding controls (**Table 4** first column) lowers numerical estimates of the gender difference for these 4 cohorts. Moreover, not only are none of the gender differences in these four cohorts significant in **Table 4** (not even 1998–2001), but the combined 1995–2006 effect is small and insignificant as well. The year-by-year results in the Supplementary Material Table A1 (graphed in **Figure 4**) show only a single year—2006—with a significant and negative gender difference at the 1–2 year stage between 1995 and 2009.

Returning to **Table 3**, the four cohorts (1995–2009) where women were less or equally likely to remain in engineering in the 2 years post-BSE are balanced by a single cohort where

women are much *more* likely to remain, leading to a zero average gender difference. Women in the 1991–1994 cohort were 8.6 *ppt.* *more likely* than men to remain in engineering; adding controls (**Table 4**) increases the gender difference to a positive 10.5 *ppt.* (Table A1 in Supplementary Material demonstrates that significantly higher women's retention was observed for 1991, 1992, and 1993 BSEs). Comparing the 1991–1994 cohort to the one immediately after, **Table 3** suggests that both a higher engagement of women in engineering and a lower engagement of men contributed to the gender difference.

Gender differences in leaving the labor force were significant for all four cohorts, although smaller in **Table 4** with controls and not significant except for the 1998–2001 cohort. The more noisy year-by-year analysis of Table A1 in Supplementary Material indicates 4 years with significantly higher female labor force exit (1996, 1998, 2001, 2007) and 2 years with significantly lower female labor force exit (1995, 2009), scattered throughout the period.

Limiting the analysis to those who worked full-time, there were *no cohorts where women were significantly less likely than men to remain* within 2 years of their BSEs in either **Table 3** or **Table 4**. Full-time working BSE women in the cohort of 1991–1994 were much more likely to remain in engineering than men, with full-time women 9.6 *ppt.* more likely to remain without controls and 11.4 *ppt.* more likely with them¹¹. In addition, the most recent cohort of full-time working women, those who received their BSEs in 2006–2009, were also more likely than comparable men to remain in engineering in years 1–2, with the difference more significant with controls ($p = 0.027$) than without ($p = 0.106$). In the year-by-year analysis, this is reflected in positive coefficients in 2006–2008, significant (and large) in 2006.

Cohort Differences at 3–4 Years

In the averages discussed earlier, women were less likely than men to remain in engineering at 3–4 years post-BSE, although this was mostly explained by controls. Women were also more likely to leave the labor force. Limiting to those working full time, not only did the average gender difference in retention disappear, but with controls it seemed that FT working women were 1.8 *ppt.* *more* likely than men to stay in engineering at this career point.

When we divide this into cohorts, we find that this pattern was generally accurate for five of the six cohorts observed at this stage, with the exception again being those with BSEs 1991–1994. Thus, for each of the other five cohorts, women were less likely to remain in engineering than men at the 3–4 year point; these differences were significant for only two of the five cohorts: 1998–2001 and 2002–2005. This was true without (**Table 3**) or with (**Table 4**) controls. The year-by-year effects (Table A1 in Supplementary Material) corroborate these results.

In terms of exit from the labor force, significant gender differences are present for these two cohorts as well as for the earliest cohort (BSEs 1989–1990). As a consequence, limiting the analysis to full-time workers shrinks the gender retention

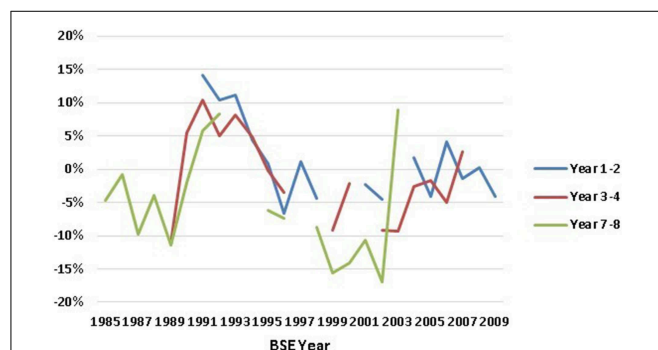


FIGURE 4 | Gender gap in retention in engineering, by BSE-year, calculated as coefficients on Female X BSE-year interaction terms in regression results of Table A1 in Supplementary Material. Data Source: NSF SESTAT Surveys 1993–2010.

¹¹As above, the three individual years 1991, 1992, and 1993 were separately significant in Table A1.

differences for these 5 cohorts: without controls only the average gender gap for the 2002–2005 full-time cohort remained significantly negative; with controls, *none* of these five cohorts had significantly lower full-time female retention rates.

As we saw at 1–2 years, the exceptional cohort at 3–4 years was those with BSEs in 1991–1994. These women were 5.6 *ppt.* more likely to remain in engineering than men on average (Table 3), 7.2 *ppt.* more likely with controls (Table 4). Full-time working women were 10.1 *ppt.* more likely than full-time men to remain in engineering with controls, and there was no gender difference in exit from the labor force. The year-by-year results of Table A1 corroborate this unusual pattern for each year of this cohort, including 1990. Men's participation in engineering at this stage was not particularly low for this cohort; instead, women's participation was particularly high.

Based only on the 1–2 year career stage, we might have concluded that women in later cohorts were more likely than men to leave engineering, because the earliest cohort observed (1991–1994 BSEs) were so different than those after it. At the 3–4 year career stage, we can now observe earlier cohorts than 1991–1994 BSEs. We see that 1991–1994 BSE was *not* representative of earlier cohorts. Instead, it was only the 1991–1994 cohort that was exceptional in its staying power.

Cohort Differences at 7–8 Years

Seven to eight years post-BSE, averaging across cohorts women were less likely to remain in engineering with or without controls, with larger differences (8.3 *ppt.*) than seen at earlier stages. This had been primarily due to 8.5% more women than men leaving the full time labor force. Among those who worked full-time, the average gender difference in retention dropped to 3.0 *ppt.* and with controls became less than 1 *ppt.* and insignificant.

Again, the cohort analysis indicates that a higher retention of women compared to men in the 1991–1994 cohort had been balancing out negative gender differences among the other cohorts. Women from all other cohorts (1985–1990, 1995–2003) were significantly less likely than men to remain in engineering by year 7–8, with gender differences in cohorts ranging from 6.1 *ppt.* to 13.0 *ppt.* (Table 3). Adding controls (Table 4) makes these gender differences only modestly smaller and still significant, with the exception of the 2002–2003 cohort—the latest one—whose significance falls to $p = 0.15$.

Women were much more likely than men to have left the labor force at year 7–8 across *all* cohorts including the 1991–1994 cohort and the 1995–1997 cohort (with 8.1 *ppt.* and 10.1 *ppt.* gender differences), two cohorts that previously had not left in greater numbers than men.

Despite this, women in the 1991–1994 cohort who remained working full-time continued to be *much more likely* than men in this cohort to remain in engineering with and without controls (9.2 *ppt.* and 12.1 *ppt.*, respectively), and also much more likely to remain in engineering than women in the previous or subsequent cohorts.

Only women in the 1998–2001 cohort continued to have a significant and large gender disadvantage in retention among those working full-time, 10.2 *ppt.* without controls and 9.3 *ppt.* with. This gender difference was equally due to men's

high likelihood of remaining in engineering and women's low likelihood of remaining.

The year-by-year effects from Table A1 in the Supplementary Material and Figure 4 add interesting nuances. Every one of the separate year effects 1998–2002 showed significantly lower female retention for both the whole and the full-time sample, and significantly higher female rates of leaving the labor force. Among other things, this suggests that the cohort should have been defined as 1998–2002. BSEs from 1991 and 1992 (the only years between 1991 and 1994 observed by SESTAT at the 7–8 year point¹²) had significantly positive gender differences for full-time women.

Cohort Differences at Later Career Stages

We only observe a limited number of BSE years at later career stages. The cohort analysis of Table 4 Panel D follows those who were observed working in engineering at approximately 7–9 years post-BSE through year 15–16. It includes only 884 observations, 152 of whom were female. The earliest observable cohort year of 1984 had large gender-differences (13.1 *ppt.*) in engineering retention by the 15th–16th year. This was due to an extremely high rate of women's leaving the labor force: no gender difference remained among those working full-time. Those with 1995 BSEs who had remained working in engineering through year 7–9 were *more* likely than men to remain in engineering at year 15–16 and equally likely as men to remain in the labor force. Given the SESTAT timing, we observe few people who received BSEs between 1985 and 1994 so results completely lacked power and significance. Because Panel D analysis is based on so few observations, we consider these results only suggestive.

Estimating Cohort Gender Differences as Careers Unfold

A final way we illustrate the differences between cohorts over careers is via six regressions, one for each of the six cohorts, each one on all years of data that we observe that cohort. In each regression, we estimated the likelihood of a cohort remaining in engineering as a function of the regular covariates (race, field dummies, year dummy, citizenship dummy) as well as on two polynomials functions (quartics) for year from BSE, one for male and one for female. This allows us to predict the gender differences in mobility as careers develop separately for men and women. These gender differences for the whole population are illustrated as Figure 5.

The average for each cohort illustrates similar differences to those found earlier, i.e., the cohorts of BSE < 1991, BSE 1995–1997 and particularly BSE 1998–2001 have negative gender differences and the cohort of 1991–1994 has the most positive gender difference.

However, this figure adds interesting information on patterns as careers develop, although we are reluctant to base too much of our analysis on this figure because the size of some cohorts at some post-BSE years is quite small. The earliest two cohorts have gender differences that start with women being more likely to be in engineering, but these differences becomes increasingly negative as they age and many have children. Interestingly, for

¹²Recall that there were no SESTAT surveys 2000–2002.

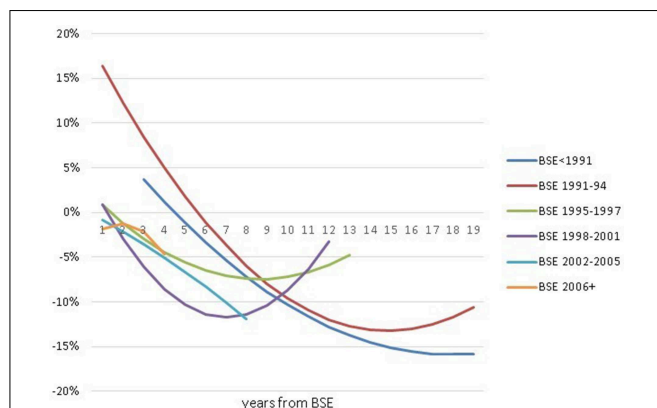


FIGURE 5 | Cohort-specific estimated time-paths of gender gaps in retention in engineering, calculated as the difference of the female and male retention rates by year-from-BSE predicted from regression. Data Source: NSF SESTAT Surveys 1993–2010.

the cohort of 1991–1994 this trend reverses and the gender gap begins narrowing at 16 years post-BSE, presumably when children's caregiving needs fall.

All later cohorts start at zero gender difference but immediately after, a gender gap appears and widens at careers develop, particularly due to women dropping out of the full-time labor force. The most enigmatic pattern is shown by the 1998–2001 cohort, with a strong U-shaped pattern bottoming out at year 7–8¹³. This reflects a reverse pattern in women's tendency to leave the labor force (also evident in the **Table 3** averages), where women's probability of being out of the labor force first decreases and then increases¹⁴, a pattern that may reflect macroeconomic conditions during the 2000s.

Alternative Measures of Retention

It is possible that our definition of “engineering” jobs based on the NSF engineering occupations classifications is too narrow, since engineering is a field that may be used in a variety of other jobs. If we are allowed to use a more expansive definition of an “engineering job”—including jobs that are “engineering-related” (e.g., engineering technicians, architects) and management jobs “requiring technical expertise in engineering or the natural sciences”—we find generally the same qualitative gender differences in retention, although the broader measure leads to somewhat more negative gender gaps. The few qualitative differences from **Table 4** are in later cohorts: 2006–2009 BSEs working full-time with controls no longer have a significantly positive coefficient at 1–2 years; at 3–4 years, 2006–2007 BSEs—but not its full-time subset—now have significantly negative coefficients; and the 2002–2003 cohort now has significantly negative retention gender differences at 7–8 years, but again not for its full-time subset.

¹³The cohort of 1995–1997 BSEs also has a U-shape, but this nonlinearity is insignificant ($p = 0.45$) in sharp contrast to the 1998–2001 BSE cohort where the nonlinearity has a p -value of 0.02.

¹⁴This remains the case even if we exclude people who are currently in school. The same pattern of labor force participation is seen to a much smaller extent among men.

Synthesis of Cohort Differences

Our main research question was to investigate whether the latest cohorts are unusual in terms of gender differences in retention, or more generally whether we observe a time trend across cohorts. We find no evidence that the gender differences in the cohort of the last half of the 2000s were consistently and significantly different than cohorts of the preceding decade. We tested and rejected that the gender gap was significantly different between the 2006+ cohort and the preceding one (2002–2005) at both the 1–2 and the 3–4 year stages (We do not observe BSEs from the last half of the 2000s at the 7–8 year stage).

Moreover, **Figure 4** and Table A1 in the Supplementary Material show individual cohort-year gender retention gaps with variations from 1998 BSE and later that look more like noise than trend. We have statistically tested for general time trends in cohort-year gender retention gaps in any of the 9 time-series of Table A1 in the Supplementary Material (corresponding to retention by the whole population, by the full-time working population, and leaving the labor force, at each of the 3 career-stages)¹⁵. The only significant time-trend we find (at $p \leq 0.10$ levels) is a trend toward larger negative gender differences in retention over time at the 1–2 years post-BSE stage (for both the whole population and the FT population). However, this estimated time trend is entirely due to the fact at the 1–2 year point, the 1991–1994 cohort—where women remain more than men—is the earliest cohort observed. This trend disappears at the career stages that include pre-1991 cohorts. Moreover, excluding the 1991–1994 cohort, there are no significant time trends in any of 9 time-series (as evident in **Figure 4**). The one other time-series that approaches being significant (with or without the 1991–1994 cohort) is a slightly decreasing tendency to leave the labor force at the 7–8 year career stage (both $p = 0.11$).

Our results highlight two cohorts as being unusual: (1) the cohort who received their BSEs 1991–1994, where in regression results women were more likely than men to stay in engineering at each career stage, for the whole population as well as for full-time workers only; and (2) the cohort of 1998–2002, where in women were substantially less likely than men to remain in engineering at the 7–8 year stages, even among those women working full-time.

We analyzed whether these two cohorts were unlikely to have occurred randomly. If we assume that all of annual coefficients on the gender retention differences at the three different career stages from Table A1 in the Supplementary Material were generated randomly from a normal distribution, we can examine whether the coefficients for these cohorts were sufficiently different from the mean coefficient such that they were less than 10% likely to have been generated randomly so that the coefficients appear in the normal distribution's top or bottom 5% tail. We found coefficients in the top 5% of the distribution at various career stages in the years 1991, 1992, and 1993; we found coefficients in the bottom 5% in 1999 and 2000 only at 7–8 year stage; and finally we found coefficients for 2002 in the bottom tail, again at the 7–8 year stages. In an alternative test to distinguish

¹⁵To do this, we run regressions of the coefficients on a time trend variable. Each regression has 12–14 observations depending on the career stage.

outliers (looking at distributions within each column in Table A1 in the Supplementary Material separately), the early 90s remained as outliers. However, neither 1999, 2000 nor 2002 were in the hypothetical bottom tail. We conclude that the finding that women with early 1990s BSEs were less likely than men to leave engineering at all three career points is quite robust, but that we are less certain that women with 1998–2002 BSEs were unusually likely to leave engineering at the 7–8 year point.

Where Do They Go?

Leaving the Labor Force for Family Reasons

Women leaving the labor force are responsible for a good portion of the gender retention differences observed, and variation in the rate of leaving the labor force over the career cycle and across cohorts propel some of our findings. As **Table 1** showed, an average 10.3% of female BSEs did so by 7–8 years. For many women, leaving the labor force coincides with having children. For instance, of the women out of the labor force at 7–8 years post-BSE, 72% had children compared to only 29% of BSE women working full time. In this section, we investigate whether cohort differences observed were a result of changing fertility decisions over cohorts such as postponing child-bearing or marriage. To do this, we add family terms, specifically interaction terms for female X cohort X family-status to our regressions of **Table 4** (first columns). We combine males of all family types into a single category because few men leave the labor force irrespective of family status. We divide women into three categories: single women without children, married women without children, and women with children¹⁶. The coefficients of the three family-status terms by cohort are graphed in **Figure 6**, where a value of 0 means that the women were similar to men. **Figure 6** shows that single women without children are more likely than men to remain in engineering at the 7–8 year point for every cohort except for the unusual 1998–2001 cohort. This is true both overall and among the subset working full time. For the cohort with 1991–1994 BSEs, single childless women's

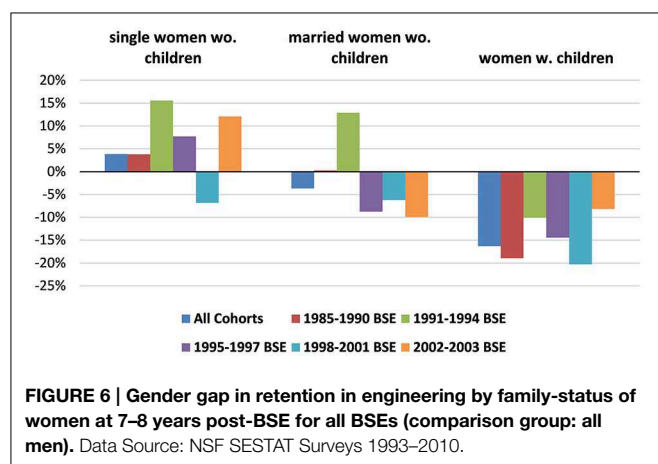
advantage over men in staying (15.6 *ppt.*) is more than twice the average for all women in **Table 4**. This indicates that the 1991–1994 BSEs were not outliers because they tended to be single or childless: instead, they were outliers within the group of single (or married) women without children. For the remaining three cohorts of single childless women, the gender advantage is not always significant, but positive and jointly significant.

In contrast, women with children (right-hand set of histogram bars) are much less likely than men to remain in engineering at the 7–8 year point for all cohorts. For these women, the magnitudes of the gender differences for the four cohorts with female retention disadvantages are between 70 and 230% greater than in **Table 4** gender differences, with 1998–2001 being largest and the earliest 1985–1990 cohort second largest. Gender differences for the fifth cohort—1991–1994 BSEs—switch from significantly positive to insignificantly negative.

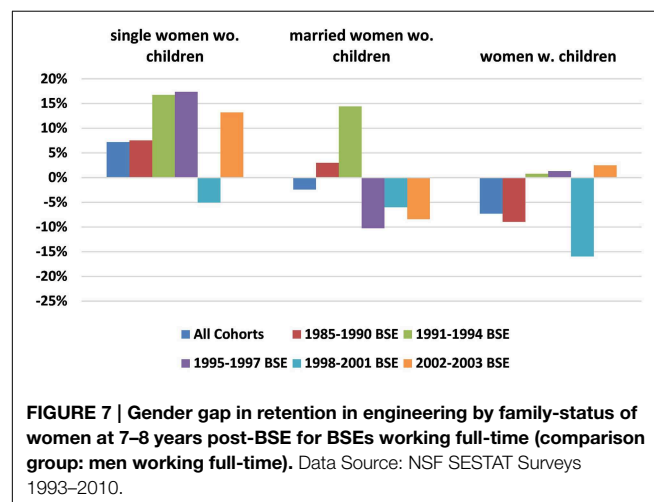
Finally, marriage alone—even in the absence of children—seems to affect women in some cohorts. Thus, in the 1995–1997 and 2002–2003 cohorts, childless married women are significantly less likely to continue in engineering than single childless women.

We have also re-estimated our regression of the likelihood of *leaving the labor force* including gender-family status interactions and found that women of all family situations are significantly more likely than men to leave the labor force, although by far the largest differences are for those women with children. Specifically, married women without children are least likely to leave (gender difference 1.9 *ppt.*), single women without children are slightly (but significantly) more likely to leave (gender difference 3.3 *ppt.*), but women with children are a huge 18.4 *ppt.* more likely than men to leave the labor force by the 7–8 year career stage. Dividing into cohorts, the impact of children on remaining in the labor force has no time trend, with gender differences ranging from 13.8 *ppt.*–22.6 *ppt.*

Even for those who remain working full-time, children may lead women to leave the engineering occupation if engineering is particularly demanding in terms of hours or hours-inflexibility (Goldin, 2014). **Figure 7** illustrates the gender engineering retention differences of those working full time, by family status.



¹⁶There are too few single women with children to separate them from married women with children. We have tried dropping them and results are similar, not surprising in light of the fact that children rather than marital status dominates the results for married women.



For women without children—both single and married—the gender differences for those working full time are similar to the ones in **Figure 6**, with one difference in scale: single childless women with 1995–1997 BSEs who work full time are now *much* more likely (17.4 *ppt.*) to remain in engineering than comparable men.

For women with children working full time (right-hand set of bars), however, there are basically zero gender differences for 3 of the 5 cohorts (including the 1991–1994 cohort). Children did not deter these cohorts of women from remaining in engineering.

Among women with children working full-time, both the exceptional cohort of 1998–2001 BSEs and the earliest cohort (1985–1990) continue to have large and significant female disadvantages. But while the 1998–2001 cohort of women is less likely than men to remain in engineering irrespective of their family status, it takes marriage and/or children to deter the earliest 1985–1990 cohort. This may be representative of the period before 1985 as well, where marriage and children have a large impact not just on whether a woman works, but on whether she works in engineering jobs.

To summarize, single women without children are actually more likely than men to remain in engineering. Children have the greatest effect pulling women out of the labor force and thus out of engineering jobs. Among women and men working full-time, women with children in three cohorts behave like men. Children and marriage deter even full-time working women from remaining in engineering for the earliest cohort. The cohort of women with 1998–2001 BSEs has the least attachment to engineering irrespective of family situation. The cohort of women with 1991–1994 BSEs only has a higher likelihood than men of staying in engineering if they have no children.

Leaving for Other Occupations

Even though children explain much of the gender differences in remaining in engineering in most cohorts, we are interested in knowing whether more recent cohorts of women who work full-time are more likely than previous cohorts not just to leave

engineering, but to leave all technical or math-intensive fields (chemistry, physics, math, geology, economics) STEM jobs. This may occur if they were overly encouraged to enter fields that did not particularly interest them.

For those who have left engineering but remain working full-time at the 7–8 year post-BSE point, **Figure 8** shows the gender difference in the percent of full-time working BSEs working in various types of occupations. The largest gender difference across all cohorts is that women are more likely than men to move to non-intensive STEM occupations, in which we include biology, psychology, and social science jobs. In fact, women are on average more than four times as likely as men to move from engineering BSEs to being in these non-mathematical STEM occupations, a sector that grew considerably over the study period and that increasingly attracted women majors (**Figure 1**). Women are also significantly more likely than men to move to health jobs (which included health management). We note that women in the latest cohort observed at the 7–8 year point (2002–2003) are more likely to move to both health and non-math STEM jobs.

While women are more likely to move to non-math-intensive STEM jobs, men are more likely to move to non-STEM jobs.

On average, women and men are equally likely to move out of the more technical, math-intensive jobs shown in the first, second and fourth sets of bars of **Figure 8**. Isolating cohorts, the 2002–2003 cohort does not demonstrate a consistent tendency to move from these jobs, suggesting that recent cohorts of women are not running away from technical/math fields. The only cohort with consistent behavior across these sectors is that of 1991–1994: although women in this cohort were more likely to stay in engineering than men, they were less likely to go into other technical, math-intensive jobs, perhaps because the more technical-focused women of this cohort remained in engineering.

The third set of bars represents technically-oriented managerial jobs. Men are clearly more likely to move to these jobs. However, women have a small advantage in moving to non-STEM management jobs. We presume that this difference is

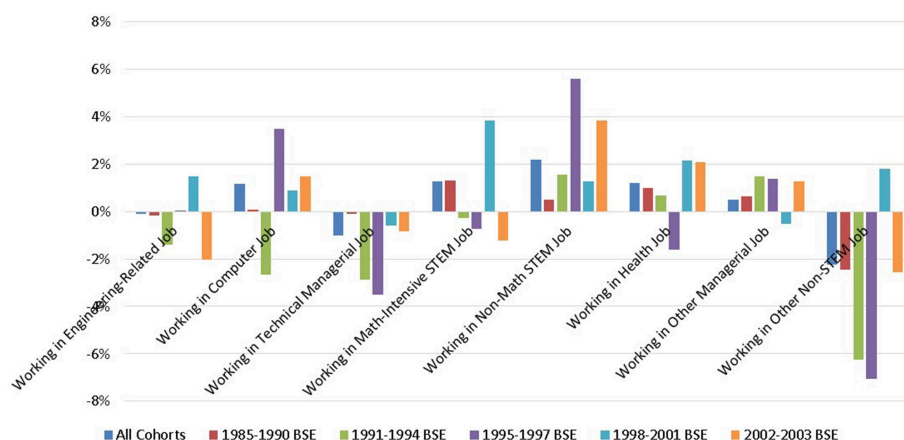


FIGURE 8 | Cohort-specific gender differences in the probability of being in occupations other than engineering among those working full-time (at 7–8 years post-BSE). Data Source: NSF SESTAT Surveys 1993–2010.

likely to be dominated by opportunities for advancement rather than choice.

Summary and Discussion

This paper uses NSF longitudinal SESTAT data to study recent cohort differences in gender-specific careers of people who received BSE. It concentrates on the first 8 years of people's post-bachelors career because we cannot observe many cohorts for longer periods. Our analysis misses data for certain cohorts due to the irregular periodicity of the SESTAT surveys. Nevertheless, the sample is large and complete enough to find significant results related to changes in gender differences over cohorts.

The paper's major contribution is to consider whether there are time patterns in the gender differences in leaving engineering for other jobs within the first 8 years after receipt of a Bachelors in Engineering (BSE). This is of particular interest if recent cohorts of female BSEs are opting out of engineering because they feel it is a bad match.

We find that overall, women are more likely than men to leave engineering by 3–4 years post-BSE for some cohorts and by 7–8 years post-BSE for all but one cohort. However, there are no clear time trends in this gender difference. Particularly, retention of women in the most recent cohorts is neither particularly high nor low.

We find that much of this gender difference is attributable to women leaving the labor force, similar to the findings of several others (Society of Women Engineers, 2009; Hunt, 2010). Thus, at 7–8 years post-BSE, the gender difference in leaving the labor force completely is 8.5 *ppt.*, more than enough to account for the overall gender difference. Gender differences in leaving the labor force for BSEs was shown to be similar to that among all college graduates (calculated from American Community Survey). There is a small time trend toward women in later cohorts being less likely to leave the labor force at the 7–8 year career point.

Family status is of key importance. Women with children are most likely to leave the labor force and therefore engineering. Single women without children are actually less likely than men to leave engineering (by the 7–8 year point) for 4 of the 5 cohorts.

Similarly, women who remain working full-time on average are somewhat *more* likely than full-time men to remain in engineering jobs through the 3–4 year post-BSE point, and equally or more likely 7–8 years post-BSE for four of the five cohorts. Dividing by family status, single women without children who work full-time are more likely to remain for four of the five cohorts at the 7–8 year point and even women with children are equally likely to remain for 3 of the 5 cohorts.

Two cohorts stand out. The first is the cohort with BSEs in the early 90s (1991–1994) where women were 8–12% *more* likely than men to remain in engineering jobs through the 7–8 year point. Having children did discourage even these women

to leave the labor force and thus engineering, but those with children who remained working full-time were equally likely as men to remain in engineering. Moreover, unlike the previous cohort (BSE < 1991), **Figure 5** indicates that this cohort's gender gap in retention (not limited to full-time workers) bottoms out at 16 years post-BSE, again reflecting the unusual aspect of the 1991–1994 cohort in that they returned to engineering once their child-rearing responsibilities lightened.

On the other hand, the cohort of women with 1998–2001 BSEs seems more likely than any of those studied to leave engineering jobs for other jobs, particularly by the 7–8 year point, irrespective of family status. The unusual pattern of this cohort of women's labor force commitment (with more out of the labor force in the years immediately post-BSE than some years later, later followed by increased exit) suggested the possibility of macroeconomic factors' influencing this cohort.

The earliest cohort picked up by SESTAT at the 7–8 year point are 1985–1990 BSEs. Children and marriage lead this cohort of women to be more likely to leave engineering even if they remain working full-time. This suggests an improvement in the environment of engineering jobs since 1990 making it easier for mothers to remain in their jobs, perhaps the result of the 1993 Family and Medical Leave Act.

Full-time working women who left engineering were equally likely as full-time men to remain in technical, math-intensive jobs, with no clear time trend, again suggesting that recent cohorts of women BSEs are not more ill-suited to mathematical/technical work than previous ones.

In sum, women who get BSE behave similarly to other college-educated women in terms of their likelihood to leave the labor force for family reasons. There has been a slight decrease over time in this likelihood. Of those who remain working full-time, women and men are equally likely to stay connected to engineering and, if they do leave engineering, to use their technical skills. There is no evidence that later cohorts of women who work full-time are different than previous cohorts of women. With the large growth in female engineering majors and an unchanging rate of retention, we can expect future growth of women in engineering careers.

Acknowledgments

The use of NSF data does not imply NSF endorsement of the research, research methods, or conclusions contained in this report.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01144>

References

- Carr, R., Bennett, L., and Strobel, J. (2012). Engineering in the K-12 STEM standards of the 50 US States: an analysis of presence and extent. *J. Eng. Educ.* 101, 539–564. doi: 10.1002/j.2168-9830.2012.tb00061.x
- Ceci, S. J., Ginther, D., Kahn, S., and Williams, W. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Glass, J. L., Sassler, S., Levitte, Y., and Micheltore, K. (2013). What's so special about STEM? A comparison of women's retention in STEM and professional occupations. *Social Forces* 92, 723–756. doi: 10.1093/sf/sot092
- Goldin, C. (2014). A grand gender convergence: its last chapter. *Am. Econ. Rev.* 4, 1091–1119. doi: 10.1257/aer.104.4.1091
- Hunt, J. (2010). "Why do women leave science and engineering?" *National Bureau of Economic Research Working Paper No. 15853*. Available online at: <http://www.nber.org/papers/w15853>, (Accessed March, 2010).
- Morgan, L. (2000). Is engineering hostile to women? An analysis of data from the 1993 National Survey of College Graduates. *Am. Sociol. Rev.* 65, 316–321. doi: 10.2307/2657444
- National Academy of Engineering and National Research Council. (NAE/NRC). (2014). Sara Frueh, rapporteur. *Career Choices of Female Engineers: A Summary of a Workshop*. Washington, DC: National Academies Press.
- Preston, A. E. (1994). Why have all the women gone? A study of exit of women from the science and engineering professions. *Am. Econ. Rev.* 84, 1446–1462.
- Preston, A. E. (2004). *Leaving Science: Occupational Exit from Scientific Careers*. New York, NY: Russell Sage Foundation.
- Singh, R., Fouad, N., Fitzpatrick, M., Liu, J., Cappaert, K., and Figuereido, C. (2013). Stemming the tide: predicting women engineer's intentions to leave. *J. Vocat. Behav.* 83, 281–294. doi: 10.1016/j.jvb.2013.05.007
- Society of Women Engineers (SWE). (2009). *National Survey About Engineering. (Excerpts from Society of Women Engineers Magazine, 2007, 2008 and 2009)*. Available online at: <http://www.nxtbook.com/nxtbooks/swe/nationalsurveyengineering/>
- Xie, Y., and Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.
- Xu, Y. J. (2008). Gender disparity in STEM disciplines: a study of faculty attrition and turnover intentions. *Res. High. Educ.* 49, 607–624. doi: 10.1007/s11162-008-9097-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kahn and Ginther. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields

Rong Su^{1*} and James Rounds²

¹ Department of Psychological Sciences, Purdue University, West Lafayette, IN, USA

² Department of Educational Psychology and Psychology, University of Illinois at Urbana-Champaign, Champaign, IL, USA

Edited by:

Stephen J. Ceci, Cornell University, USA

Reviewed by:

Teresa Wilcox, Texas A&M University, USA
Richard Allen Lippa, California State University, Fullerton, USA

*Correspondence:

Rong Su, Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907-2081, USA
e-mail: rsu@purdue.edu

The degree of women's underrepresentation varies by STEM fields. Women are now overrepresented in social sciences, yet only constitute a fraction of the engineering workforce. In the current study, we investigated the gender differences in interests as an explanation for the differential distribution of women across sub-disciplines of STEM as well as the overall underrepresentation of women in STEM fields. Specifically, we meta-analytically reviewed norm data on basic interests from 52 samples in 33 interest inventories published between 1964 and 2007, with a total of 209,810 male and 223,268 female respondents. We found gender differences in interests to vary largely by STEM field, with the largest gender differences in interests favoring men observed in engineering disciplines ($d = 0.83$ – 1.21), and in contrast, gender differences in interests favoring women in social sciences and medical services ($d = -0.33$ and -0.40 , respectively). Importantly, the gender composition (percentages of women) in STEM fields reflects these gender differences in interests. The patterns of gender differences in interests and the actual gender composition in STEM fields were explained by the people-orientation and things-orientation of work environments, and were not associated with the level of quantitative ability required. These findings suggest potential interventions targeting interests in STEM education to facilitate individuals' ability and career development and strategies to reform work environments to better attract and retain women in STEM occupations.

Keywords: interests, gender differences, people-orientation, things-orientation, gender disparities in STEM fields

INTRODUCTION

Despite major advancement of women's participation and status in the workforce over the past decades, women overall remain the minority in science, technology, engineering, and mathematics (STEM) disciplines. The underrepresentation of women in STEM fields keeps our society from fully utilizing human capital and is of great concern to researchers, educators, and the general public. However, past research on this topic typically treated all STEM fields as a whole and ignored the differences among sub-disciplines of STEM. It is important to note that all STEM fields are not identical. Sub-disciplines of STEM vary in their culture and climate, training and preparation required, and the type of work activities involved. The percentages of women across sub-fields of STEM also vary vastly. For example, women have made immense progress in biomedical and social sciences, now earning over 50% of bachelor's and master's degrees, whereas the percentage of women obtaining any level of engineering degree lingers below 20% (National Science Foundation, 2013). To build a more balanced and competitive workforce, we need to gain a better understanding about the psychological and socio-cultural factors that contribute to the differential participation of women across STEM sub-disciplines. Investigating why women are scarce in some STEM fields but not in others may offer us insight into how to increase women's overall representation in STEM.

The current study focuses on the differential interests of men and women that may drive career choices within STEM fields just as they influence the selection between STEM and other careers¹. Interests have been consistently shown as a critical predictor for career choice and career attainment. Existing studies have suggested that the differential interests of men and women are one of the most important psychological mechanisms that underlie gendered career choices and gender disparities in the STEM fields (e.g., Lubinski and Benbow, 1992; Ceci et al., 2009; Su et al., 2009). For example, Su et al. (2009) examined gender differences in vocational interests and two work-task dimensions (namely, *Things–People* and *Data–Ideas*; Prediger, 1982) and found substantial gender differences in the *Things–People* dimension ($d = 0.93$), with men preferring working with things and women preferring working with people. The effect size of this

¹ We acknowledge that many other factors may underlie women's underrepresentation in STEM fields, including (1) cognition and learning, particularly factors related to mathematical preparation and achievement, (2) developmental environments, such as influences from home, school, and peers, and (3) institutional and organizational biases in the hiring, training, and promotion processes. We focus on interests to offer the evidence for one important, yet less emphasized perspective that explains the differential representation of women across STEM sub-disciplines as well as the overall underrepresentation of women in STEM fields.

gender difference in interests was close to one standard deviation, and among the largest reported in the literature of individual differences (Lubinski, 2000). Interests in people-oriented careers may explain women's underrepresentation in some STEM fields, which are typically things-oriented.

Despite these findings suggesting the role of interests in gendered career choices, several gaps exist in this research. First, no study has looked within STEM fields and investigated men and women's interests in each sub-discipline of STEM. Second, although many studies reported statistics on the percentages of women in STEM occupations, no research has compared the trend in labor statistics with gender differences in interests to examine whether actual percentages of women in each STEM sub-discipline match or mismatch their interests. This information is critical, as it will help identify areas where interventions could be fruitful for increasing the participation of women. Third, past research typically studied the determinants of STEM career choices using individuals as the unit of analysis and rarely incorporated indicators of occupational characteristics to study their effects on men and women's interests at the occupational level. Understanding the interaction between individuals' interests and the characteristics of STEM occupations is essential for explaining why women remain severely underrepresented in some STEM fields and yet are growing in numbers in other STEM fields that are equally demanding intellectually and temporally.

In this article, we seek to advance the literature in the following ways. First, we highlight the differential interests of men and women within STEM fields and offer it as one explanation for the uneven distribution of women across the STEM disciplines. We extended Su et al. (2009) meta-analysis and examined gender differences in basic interests (i.e., specific and homogeneous interests in activities and objects with shared properties, such as *Mathematics* or *Biological Science*). Specifically, we examined men's and women's basic interests in the full range of STEM fields, from Engineering in which the number of women are the sparsest, to Social Sciences in which women are over-represented. Further, we investigated the extent to which gender differences in basic interests contributed to the gender composition (percentages of men and women) in corresponding occupational fields, and the degree to which these gender differences in basic interests mediated the effects of occupational characteristics, such as people- and things-orientation and job requirement in quantitative ability, using a person-environment (P-E) fit approach.

THE OTHER SIDE OF THE COIN: THE ROLE OF INTERESTS IN STEM CAREERS

Person-environment (P-E) fit theories (e.g., Holland, 1959, 1997; Pervin, 1968; Dawis and Lofquist, 1984; Schneider, 1987) maintain that individuals and environments can be described using a commensurate set of characteristics. For example, an individual can be described in terms of his/her interests in social or people-related activities, and an environment can be described in terms of its likelihood to fulfill such interests. An environment may be conceptualized at a variety of different levels, such as an academic major, an occupational field, organizational culture and climate, or the relationship with supervisor and work team (Su et al., 2014). Further, the degree of compatibility between

individual and environmental characteristics is associated with career choice, satisfaction, and performance. Individuals seek out and thrive in environments that provide a good fit with their traits and motives; they are likely to stay in environments that are compatible, and will leave those environments that are incompatible. As such, people's interests in work environments channel their career decision-making and career advancement.

It has been consistently shown that, compared to men, women have stronger preference for work environments that provide more opportunities and activities to work with people. Such preference has been explained under different theoretical frameworks, such as people-orientation (e.g., Thorndike, 1911; Woodcock et al., 2013), social interests (e.g., Su et al., 2009; Robertson et al., 2010), subjective task values (e.g., Meece et al., 1982; Eccles, 2007), and communal goals (e.g., Diekmann et al., 2010; McCarty et al., 2014). Regardless of the theoretical framework used, research in this area has shown that differential preferences of men and women are associated with the gender disparities in STEM fields. For example, in a series of 15 studies, Woodcock et al. (2013) examined the people-orientation and things-orientation of 7450 participants and found that females consistently scored higher than males in people-orientation (mean $d = 0.49$, range from 0.11 to 0.86), whereas males consistently scored higher than females in things-orientation (mean $d = 0.99$, range from 0.58 to 1.33). Moreover, Woodcock et al. (2013) showed that people- and things-orientations predicted the choice of a STEM major in college, with things-orientation positively associated with STEM major choice and people-orientation moderating this relationship (that is, a particularly strong relationship between things-orientation and STEM major choice when people-orientation is low). The effects of people- and things-orientations on STEM major choice fully accounted for the effect of sex.

Similarly, Su et al. (2009) conducted a meta-analysis that quantitatively synthesized data from 47 interest inventories with 503,188 respondents, and reported substantial gender differences in interests. Specifically, males on average scored higher on the Realistic scale that measured interests in working with things and gadgets or working outdoors ($d = 0.84$); in contrast, females on average scored higher on the Social scale that measured interests in helping people ($d = -0.68$). Su et al. (2009) argued that gender disparities in STEM fields occurred for two reasons: first, from an inter-individual perspective, men outnumber women in the upper tail of the Realistic interest distribution, which predicts entry into things-oriented careers including STEM fields; second, from an intra-individual perspective, given the same level of Realistic interests, women are more likely than men to have a competing level of Social interests, which orient them toward people-oriented careers, or, within STEM fields, those sub-disciplines that are more likely to fulfill their interests in helping people, such as medical science and services.

Eccles and her colleagues (Meece et al., 1982; Eccles, 1994, 2007, 2009; Jacobs et al., 2005) argued that the perceived task values of various occupational options (e.g., "Can I directly relate to people and help people in this occupation?") is one of the most important mechanisms underlying educational and occupational choices, including the decision to enter STEM fields and the

choice within various STEM sub-disciplines. Because females are socialized to possess higher social values in interacting and helping people, they are more likely to be drawn to occupational fields with work tasks that are perceived to fulfill these values, such as teaching, nursing, or medical science, rather than fields that are perceived to be low in these values, such as physical science and engineering.

Lastly, through two experimental studies, McCarty et al. (2014) demonstrated that participants who highly valued communal goals, regardless of gender, had aversive and avoidant reactions to work environment that is low in communion. Specifically, Diekmann et al. (2010, 2011) showed that the endorsement of communal goals significantly impeded intention to pursue STEM careers, even when controlling for past experience and self-efficacy in science and mathematics. Consistent with the literature, women on average scored higher on communion than men, suggesting that women were less likely to favor work environments that are perceived less compatible with communal goals, including some STEM fields.

Based on the above evidence, we argue that men and women's differential interests for work environments provide the other side of the coin—an equally, if not more, important psychological mechanism for understanding the gender disparities in STEM fields—in addition to cognition and learning pertaining to math preparation and achievement. We propose that the interests that underlie women's overall underrepresentation in STEM fields also underlie the differential distribution of women across STEM disciplines and explain why women tend to choose some STEM disciplines over others.

THE CONCEPTUALIZATION, FUNCTION, AND MEASUREMENT OF INTERESTS

Interests are defined as “trait-like preferences for activities, contexts in which activities occur, or outcomes associated with preferred activities” that orient individuals toward certain environments and motivate goal-oriented behaviors within environments (Rounds and Su, 2014). Such intrinsic preferences construe an essential part of individuals' identity and serve as an impetus for individuals to navigate through and function effectively in their environments (Hogan and Blake, 1999; Su et al., 2009).

Based on P-E fit theories (e.g., Holland, 1997), interests directly influence educational and career choices as people gravitate toward academic or work environments that are congruent with their interests. It has been reliably shown that interests predict academic major and occupational membership (e.g., Strong, 1943; Campbell, 1971; Kuder, 1977; Savickas and Spokane, 1999). In addition, interests also impact career trajectory and attainment through its indirect effects on learning and knowledge acquisition, which prepares as well as constrains one's pursuit in certain educational and occupational fields. Interests in an activity act as a source of intrinsic motivation that drives individuals to learn more about it. An accruing volume of research has linked interest with persisted learning, deeper engagement, and better knowledge acquisition (e.g., Hidi, 2001; Silvia, 2006) and has shown the increasing coupling of interests and domain-specific knowledge/ability over time (Ackerman, 1996; Denissen et al., 2007). Thus, an individual with strong interests in mathematics, for

example, is more likely than his/her uninterested peers to aspire to education and a career in mathematics; in the meanwhile, this individual is more likely to engage in activities to learn math that leads to increased math knowledge and ability, which, in turn, prepares him/her for entry into a math major or a math-related career as well as persistence and attainment in that field.

This dynamic relationship between interests and knowledge/ability development is critically important for understanding the significance of interests for educational and career attainment in STEM. Interests do not only serve as a self-selection mechanism for a few binary choices in life such as choosing a college major or entering an occupation; rather, interests contribute to individuals' preparedness for STEM fields by promoting learning in these fields and provide the foundation for individuals' educational and career development throughout the lifespan.

Interests can be conceptualized and measured at different levels of specificity. The most commonly studied interest typology is Holland's (1959, 1997) RIASEC model (abbreviation for *Realistic*, *Investigative*, *Artistic*, *Social*, *Enterprising*, and *Conventional*), which is used to categorize both individual interests and corresponding characteristics in work environments. The RIASEC model captures the broadest level of interests and work environments. Each of the six broad categories encompasses a heterogeneous group of occupations and activities that share a common “theme.” Therefore, the RIASEC types are sometimes also referred to as the general occupational themes. For example, the *Realistic* (R) theme captures interests in working with things and gadgets, working with hands, or working outdoors. Typical occupations and work activities represented in the *Realistic* theme include carpenters, automotive engineers, farming, or putting out forest fires. The *Social* (S) theme captures interests in working with people and helping people. Typical occupations and work activities included in the *Social* theme are teachers, social workers, volunteering at a charity, or helping people solve their emotional problems. Realistic and social interests are closely associated with the constructs of things- and people-orientations (Woodcock et al., 2013). The *Investigative* (I) theme, as its name suggests, captures interests in science and research. It is the best indicator for the interests in pursuing education or careers in STEM fields. However, STEM is a broad term with heterogeneous sub-disciplines. Many disciplines in natural sciences, such as physical science, astronomy, and chemistry, also involve a heavy *Realistic* component; the most quintessential is the field of engineering, with a strong focus on working with things, in addition to its emphasis on research and investigation; in comparison, other disciplines in health and human sciences, such as psychology, medicine, or nutrition science, also involve a *Social* component. Therefore, although most STEM fields fall within the *Investigative* theme, they may be arranged on a continuum from the most things-oriented to the least things-oriented field, and from the most people-oriented to the least people-oriented. The broad occupational themes are not sufficient to capture the nuances among various sectors in the world of work and the heterogeneous interests represented in these environments. More specific measures of interests are needed.

Basic interest scales characterize shared properties of homogeneous sets of work activities and environments (Liao et al., 2008). For example, instead of broad *Social* interests, a basic interest scale may measure interest in *Teaching*, *Counseling*, or *Professional Advising* activities; similarly, instead of broad *Investigative* interests, a basic interest scale may measure interest in *Mathematics*, *Physical Science*, or *Medical Science* activities. What is unique about basic interest scales is that the interest measured by a basic interest scale is often implied in the object of interest. A *Medical Science* basic interest scale may include items like “work in a lab,” “study blood samples,” and “develop a new medicine to cure a disease.” Taken together, responses to these items reflect an individual’s level of interest in the field of medical science. In other words, the specificity of basic interests corresponds precisely with the targeted environments. As a result, basic interests provide an excellent measure of individuals’ preferences for specific work environments; gender differences in basic interest scales represent differential preferences of men and women for these work environments, such as sub-disciplines of STEM.

OVERVIEW AND HYPOTHESES OF THE CURRENT STUDY

The purposes of the current study were three-fold: first, we examined gender differences in basic interests by STEM field, including physical sciences, biological science, medical science, medical services, social sciences, mathematics, applied mathematics, computer science, engineering, and mechanics and electronics. Because the definition of STEM disciplines varies by organization and a unified list is not available, we adopted the definition from two federal agencies: (1) *STEM-Designated Degree Programs List* from the U.S. Immigration and Customs Enforcement (2012), and (2) *STEM Workforce Sectors* from the U.S. Department of Labor (2007). We expected gender differences in basic interests to vary largely across these different STEM disciplines. Second, we demonstrated that the gender composition (percentages of men and women) in these STEM fields closely mirrored the pattern of gender differences in basic interests. Third, we sought to understand the occupational characteristics that were associated with the gender differences in basic interests.

To answer these research questions, we meta-analytically reviewed technical manuals of interest inventories that included a relevant basic interest scale. Because traditional meta-analysis is subject to sampling errors from individual studies reviewed, we selected norm groups from technical manuals to be our data source as they are typically large and well sampled (cf. Hedges and Nowell, 1995). Data from these technical manuals provide relatively accurate estimation of the differential interests of men and women for each sub-discipline of STEM. In addition, we obtained occupational characteristics from the Occupational Information Network (O*NET; National Center for O*NET Development, 2014) on the things-orientation, people-orientation, and level (i.e., amount) of quantitative ability required for each sub-discipline of STEM.

Based on P-E fit theories and existing studies showing that women had higher people interests and lower things interests compared to men, we expected to find greater gender differences in basic interests favoring men in STEM fields that are high in things-orientation and low in people-orientation, such

as engineering; we expected to find smaller gender differences in basic interests favoring men or gender differences in the opposite direction in STEM fields that are low in things-orientation and high in people-orientation, such as medical and social sciences. Consistent with the continuum of STEM sub-disciplines ordered by their things- and people-orientations, we expected the size of gender differences in basic interests in these fields to form a continuum as well. Given previous research reporting that people-orientation and things-orientation are two separate dimensions instead of opposite ends of one bipolar dimension (e.g., Graziano et al., 2011; Tay et al., 2011), we propose the following two hypotheses:

Hypothesis 1a: The gender difference in interests in a STEM field (favoring men) is positively associated with the level of things-orientation of that field.

Hypothesis 1b: The gender difference in interests in a STEM field (favoring men) is negatively associated with the level of people-orientation of that field.

With research evidence showing that gender differences in math ability and achievement are negligible (e.g., Hyde et al., 1990; Hyde and Linn, 2006), we reason that quantitative ability is not a factor that affects men and women’s differential career preferences. Thus, we expected the gender difference in interests in a STEM field to be unrelated to the level of quantitative ability required for that field once the things- and people-orientations of the field are accounted for. In other words, women’s lower interest in some STEM fields is not the result of their avoidance of work environments that require higher levels of quantitative ability, but rather, the result of their aversion to work environments that are high in things-orientation and low in people-orientation.

Hypothesis 2. Controlling for the things- and people-orientations of a STEM field, the gender difference in interests in that field is unrelated to the level of quantitative ability required.

More importantly, given the strong relationship found between interests and career choices, we expected the gender composition of various STEM fields to reflect observed gender differences in interests. Further, we expected the gender difference in interests in a STEM field to fully mediate the effects of occupational characteristics (things- and people-orientations) on the gender composition of that field. Similar to Hypothesis 2, we expected the gender composition of a STEM field to be unrelated to the level of quantitative ability required for that field once the things- and people-orientations of the field are accounted for.

Hypothesis 3. The percentage of women in a STEM field is negatively associated with the gender difference in interests in that field (favoring men).

Hypothesis 4a. The percentage of women in a STEM field is positively associated with the people-orientation of that field; this relationship is full mediated through gender differences in interests.

Hypothesis 4b. The percentage of women in a STEM field is negatively associated with the things-orientation of that field; this relationship is full mediated through gender differences in interests.

Hypothesis 5. Controlling for the things- and people-orientations of a STEM field, the percentage of women in that field is unrelated to the level of quantitative ability required.

Besides the above hypotheses, we examined several additional moderators for the gender differences in basic interests in STEM fields, including (1) job complexity, (2) the age group of a sample, (3) the year of data collection, and (4) the degree to which an interest inventory was developed to be gender-balanced (i.e., using item development strategies to remove items that displayed large gender differences and to increase the overlap between male and female interest score distributions). More details are provided for these moderators in the Methods section.

METHODS

META-ANALYTIC DATABASE

Database for the current study was composed of norm samples from vocational interest inventory technical manuals, published from 1964 till the current date. Procedures to identify and select the interest inventories were described in detail in Su et al. (2009). Because we were interested in the gender differences in basic interests in this study, the following criteria were applied to select interest inventories to form the current meta-analytic database: first, the interest inventory had one or more scales that measured *basic interests* in any fields related to physical sciences, biological science, medical science, social sciences, mathematics, computer science, and engineering. We included scales that measured interests in professional-level activities in these fields (i.e., activities performed by scientists, engineers, and mathematicians), as well as scales that measured interests in technical-level activities (i.e., activities performed by science technicians, engineering technicians, workers in applied mathematics, mechanics and electronics, and those in medical services). By including interest scales at both levels in our database, we were able to examine whether job complexity had an effect on the gender differences in basic interests and on the gender composition of various STEM fields. Second, the inventories used the same form for male and female respondents and reported means and standard deviations for both males and females in the technical manuals, allowing effect sizes of gender differences to be calculated. Third, because it was possible for an interest inventory to have multiple editions, we included data from a new edition only when it used an entirely new sample. Application of these inclusion criteria resulted in 52 samples from 33 inventories, with a total of 209,810 men and 223,268 women. The mean ages of the samples ranged from 12.50 to 42.55 years. The samples were surveyed between 1963 and 2007.

CLASSIFICATION OF BASIC INTEREST SCALES BY STEM FIELD

To identify relevant basic interest scales for each sub-discipline of STEM, we perused every interest inventory and classified each basic interest scale into corresponding STEM field based on (1) the items on the scale and (2) the correlates of the scale

score. Most basic interest scales measured interests as suggested by their titles, such as the *Social Science* scale in the Jackson Vocational Interest Survey (JVIS; Jackson, 2000). A few exceptions were classified differently than their title would suggest. For example, the *Engineering and Physical Sciences* scale in the Ohio Vocational Interest Survey II (OVIS-II; Winefordner, 1983) was classified as a scale measuring interests in engineering, rather than physical sciences, because the majority of its items were occupational titles in engineering, such as “Electronics Engineer” and “Nuclear Engineer.” Similarly, the *Mathematics and Science* scale in the Career Interest Inventory (CII; Psychological Corporation, 1991) had mostly engineering and computer science related items and was classified as a scale measuring engineering interests. The *Science* scale in the Vocational Interest Inventory-Revised (VII-R; Lunneborg, 1993) primarily measured interests in medical science, and the *Mechanical* scale in the Guilford-Zimmerman Interest Inventory (GZII; Guilford and Zimmerman, 1989) had items that measured interests in the professional-level of engineering activities, rather than the technical-level of mechanical activities. These scales were classified accordingly.

Further, some basic interest scales measured interests broader than one STEM field. For example, several scales, including the *Science* scale in the Career Assessment Inventory-Vocational edition (CAI-V; Johansson, 1984), measured interests in physical sciences and biological science. A separate category, *Natural Sciences*, was hence created to classify these scales, rather than forcing them into either the *Physical Sciences* category or the *Biological Science* category. Finally, scales that were designed to measure basic interests but rather measured the full range of interests in all disciplines of sciences and research, such as the *Research* scale in the Strong Interest Inventory (SII; Donnay et al., 2005), were excluded from the current study.

As a result, the basic interest scales from all the interest inventories were classified into 13 fields. Eight of these fields were at the professional-level, including *Physical Sciences*, *Natural Sciences*, *Biological Science*, *Medical Science*, *Social Sciences*, *Mathematics*, *Computer Science*, and *Engineering*; the other five fields were at the technical-level, including *Science Technicians*, *Engineering Technicians*, *Applied Mathematics*, *Mechanics and Electronics*, and *Medical Services*. **Table 1** lists all the basic interest scales classified under each STEM field by sample.

IDENTIFICATION OF OCCUPATIONAL CHARACTERISTICS AND STATISTICS

We obtained occupational-level information from two sources: information about the people-orientation, things-orientation, and level of required quantitative ability was acquired through the O*NET production database 18.1 (National Center for O*NET Development, 2014), and information about the percentages of women in STEM fields was obtained from the U.S. Bureau of Labor Statistics (2014) latest report *Women in the Labor Force: A Databook*. Both sources used the 2010 Standard Occupational Classification (SOC; U.S. Bureau of Labor Statistics, 2010) system, allowing us to combine two sources of information using matching occupational codes.

The O*NET database provides comprehensive and regularly updated information on various aspects of worker attributes and

Table 1 | Overview of the meta-analysis database: basic interest scale, moderator variables, and effect size by STEM field and sample.

Inventory	References	Scale	Moderator				Sample size			Effect size
			Complexity	Gender-balanced	Sample year	Age group	N	MN	FN	d
PHYSICAL SCIENCES										
SII	Donnay et al., 2005	Science	2	1	2002	5	2250	1125	1125	0.42
SII	Harmon et al., 1994	Science	2	1	1994	5	18,951	9484	9467	0.24
SII	Hansen and Campbell, 1985	Science	2	1	1985	5	600	300	300	0.28
SII	Campbell, 1974	Science	2	1	1974	5	600	300	300	0.32
CAI-E	Johansson, 2003	Scientific Research and Development	2	2	1986	5	900	450	450	0.23
KGIS-E	Kuder, 1964	Scientific	2	1	1963	1	4109	2080	2029	0.91
KGIS-E	Kuder, 1964	Scientific	2	1	1963	2	5704	2766	2938	0.87
KGIS-E	Kuder and Zytowski, 1988	Scientific	2	1	1987	1	5894	2714	3180	0.96
KGIS-E	Kuder and Zytowski, 1988	Scientific	2	1	1987	2	7113	3402	3711	0.80
KOIS	Kuder and Zytowski, 1991	Scientific	2	1	1985		3214	1583	1631	0.23
KCS	Zytowski, 2007	Scientific	2	1	2007		3619	1663	1956	0.33
JVIS	Jackson, 2000	Physical Science	2	1	1999	2	2380	1190	1190	0.71
JVIS	Jackson, 2000	Physical Science	2	1	1999	5	1120	560	560	0.52
JVIS	Jackson, 1977	Physical Science	2	1	1977	3	1000	500	500	0.76
GZII	Guilford and Zimmerman, 1989	Scientific	2	1	1989		215	97	118	0.88
NATURAL SCIENCES										
CAI-V	Johansson, 1984	Science	2	2	1976	5	1500	750	750	0.24
CISS	Campbell et al., 1992	Science	2	1	1992	5	5241	3442	1799	0.43
IDEAS	Johansson, 1996	Science	2	2	1989	1	1770	820	950	0.65
IDEAS	Johansson, 1996	Science	2	2	1989	2	2891	1208	1683	0.50
IDEAS	Johansson, 1978	Science	2	1	1977	1	598	292	306	0.30
IDEAS	Johansson, 1978	Science	2	1	1977	2	3436	1755	1681	0.35
BIOLOGICAL SCIENCES										
JVIS	Jackson, 2000	Life Science	2	1	1999	2	2380	1190	1190	0.24
JVIS	Jackson, 2000	Life Science	2	1	1999	5	1120	560	560	0.14
JVIS	Jackson, 1977	Life Science	2	1	1977	3	1000	500	500	0.18
MEDICAL SCIENCE										
SII	Donnay et al., 2005	Medical Science	2	1	2002	5	2250	1125	1125	0.06
SII	Harmon et al., 1994	Medical Science	2	1	1994	5	18,951	9484	9467	0.09
SII	Hansen and Campbell, 1985	Medical Science	2	1	1985	5	600	300	300	0.18
SII	Campbell, 1974	Medical Science	2	1	1974	5	600	300	300	0.01
CAI-E	Johansson, 2003	Medical Science	2	2	1986	5	900	450	450	−0.20
OVIS-II	Winefordner, 1983	Medical Service	2	2	1980	1	9242	4604	4638	−0.23
OVIS-II	Winefordner, 1983	Medical Service	2	2	1980	2	6416	3157	3259	−0.21
OVIS-II	Winefordner, 1983	Medical Service	2	2	1980	3	2792	1055	1737	−0.06
OVIS	D'Costa et al., 1970	Medical	2	1	1969	2	45,845	23,062	22,783	−0.13
IDEAS	Johansson, 1996	Medical	2	2	1989	1	1770	820	950	−0.15
IDEAS	Johansson, 1996	Medical	2	2	1989	2	2891	1208	1683	−0.20
JVIS	Jackson, 2000	Medical Service	2	1	1999	2	2380	1190	1190	0.15
JVIS	Jackson, 2000	Medical Service	2	1	1999	5	1120	560	560	0.11
JVIS	Jackson, 1977	Medical Service	2	1	1977	3	1000	500	500	0.12
VII	Lunneborg, 1981	Science	2	3	1976	2	600	300	300	−0.17
VII-R	Lunneborg, 1993	Science	2	3	1985	2	1562	748	814	−0.03
MEDICAL SERVICES										
SII	Donnay et al., 2005	Healthcare Services	1	1	2002	5	2250	1125	1125	−0.24
SII	Harmon et al., 1994	Medical Service	1	1	1994	5	18,951	9484	9467	−0.17

(Continued)

Table 1 | Continued

Inventory	References	Scale	Moderator				Sample size			Effect size
			Complexity	Gender-balanced	Sample year	Age group	N	MN	FN	d
SII	Hansen and Campbell, 1985	Medical Service	1	1	1985	5	600	300	300	−0.04
SII	Campbell, 1974	Medical Service	1	1	1974	5	600	300	300	−0.52
CAI-E	Johansson, 2003	Medical Service	1	2	1986	5	900	450	450	−0.12
CAI-V	Johansson, 1984	Medical Service	1	2	1976	5	1500	750	750	−0.47
CISS	Campbell et al., 1992	Medical Practice	1	1	1992	5	5241	3442	1799	−0.13
OVIS-II	Winefordner, 1983	Health Services	1	2	1980	1	9800	4479	4552	−0.58
OVIS-II	Winefordner, 1983	Health Services	1	2	1980	2	6672	3120	3217	−0.55
OVIS-II	Winefordner, 1983	Health Services	1	2	1980	3	2800	1052	1730	−0.32
OVIS	D'Costa et al., 1970	Nursing and Related Technical Services	1	1	1969	2	46,065	23,203	22,862	−0.90
IDEAS	Johansson, 1978	Medical Service	1	1	1977	1	598	292	306	−0.75
IDEAS	Johansson, 1978	Medical Service	1	1	1977	2	3436	1755	1681	−0.90
CII-1	Psychological Corporation, 1991	Health Services	1	1	1989	1	13,280	6545	6735	−0.42
CII-1	Psychological Corporation, 1991	Health Services	1	1	1989	1	19,780	9825	9955	−0.46
CII-1	Psychological Corporation, 1991	Health Services	1	1	1989	1	26,082	13,123	12,959	−0.50
CII-2	Psychological Corporation, 1991	Health Services	1	1	1989	2	14,300	6987	7313	−0.38
CII-2	Psychological Corporation, 1991	Health Services	1	1	1989	2	8556	4208	4348	−0.35
CII-2	Psychological Corporation, 1991	Health Services	1	1	1989	2	8399	4051	4348	−0.36
CDI	Jackson, 1986	Health Services	1	1	1986	3	1000	500	500	−0.37
CDI	Jackson, 2003	Health Services	1	1	2003	1	212	114	98	−0.43
CDI	Jackson, 2003	Health Services	1	1	2003	2	737	385	352	−0.34
CDI	Jackson, 2003	Health Services	1	1	2003	3	386	206	180	−0.31
CDI	Jackson, 2003	Health Services	1	1	2003	4	392	171	221	−0.30
CDI	Jackson, 2003	Health Services	1	1	2003	5	317	148	169	0.00
CDI	Jackson, 2003	Health Services	1	1	2003	5	276	145	131	−0.32
SOCIAL SCIENCES										
SII	Donnay et al., 2005	Social Sciences	2	1	2002	5	2250	1125	1125	−0.08
CII-1	Psychological Corporation, 1991	Social Science	2	1	1989	1	13,190	6491	6699	−0.42
CII-1	Psychological Corporation, 1991	Social Science	2	1	1989	1	19,674	9757	9917	−0.50
CII-1	Psychological Corporation, 1991	Social Science	2	1	1989	1	26,009	13,056	12,953	−0.59
CII-2	Psychological Corporation, 1991	Social Science	2	1	1989	2	14,871	7346	7525	−0.44
CII-2	Psychological Corporation, 1991	Social Science	2	1	1989	2	8849	4392	4457	−0.37
CII-2	Psychological Corporation, 1991	Social Science	2	1	1989	2	8679	4241	4438	−0.40
JVIS	Jackson, 2000	Social Science	2	1	1999	2	2380	1190	1190	−0.08
JVIS	Jackson, 2000	Social Science	2	1	1999	5	1120	560	560	−0.10
JVIS	Jackson, 1977	Social Science	2	1	1977	3	1000	500	500	−0.33
SCIENCE TECHNICIANS										
COPS-R	Knapp and Knapp, 1979	Science-Skilled	1	3	1979	2	400	200	200	0.02
COPS	Knapp et al., 1990	Science-Skilled	1	1	1988	2	14,619	7565	7054	0.13

(Continued)

Table 1 | Continued

Inventory	References	Scale	Moderator				Sample size			Effect size
			Complexity	Gender-balanced	Sample year	Age group	N	MN	FN	d
COPS	Knapp et al., 1990	Science-Skilled	1	1	1988	3	3237	1379	1858	0.12
COPS	Knapp and Knapp, 1984	Science-Skilled	1	1	1982	2	4145	2034	2111	0.24
COPS	Knapp and Knapp, 1984	Science-Skilled	1	1	1982	3	1445	773	672	0.17
ENGINEERING										
OVIS-II	Winefordner, 1983	Engineering and Physical Science	2	2	1980	1	9119	4530	4589	0.86
OVIS-II	Winefordner, 1983	Engineering and Physical Science	2	2	1980	2	6368	3132	3236	0.81
OVIS-II	Winefordner, 1983	Engineering and Physical Science	2	2	1980	3	2781	1052	1729	0.90
OVIS	D'Costa et al., 1970	Applied Technology	2	1	1969	2	45,832	23,058	22,774	1.55
CII-1	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	1	13,190	6491	6699	0.48
CII-1	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	1	19,674	9757	9917	0.61
CII-1	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	1	26,009	13,056	12,953	0.68
CII-2	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	2	14,871	7346	7525	0.48
CII-2	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	2	8849	4392	4457	0.46
CII-2	Psychological Corporation, 1991	Mathematics and Science	2	1	1989	2	8679	4241	4438	0.43
JVIS	Jackson, 2000	Engineering	2	1	1999	2	2380	1190	1190	1.24
JVIS	Jackson, 2000	Engineering	2	1	1999	5	1120	560	560	0.85
JVIS	Jackson, 1977	Engineering	2	1	1977	3	1000	500	500	1.13
VII	Lunneborg, 1981	Technical	2	3	1976	2	600	300	300	0.95
VII-R	Lunneborg, 1993	Technical	2	3	1985	2	1562	748	814	0.52
COPS-R	Knapp and Knapp, 1979	Technology-Professional	2	3	1979	2	400	200	200	0.19
COPS	Knapp et al., 1990	Technology-Professional	2	1	1988	2	14,619	7565	7054	1.01
COPS	Knapp et al., 1990	Technology-Professional	2	1	1988	3	3237	1379	1858	0.89
COPS	Knapp and Knapp, 1984	Technology-Professional	2	1	1982	2	4145	2034	2111	1.05
COPS	Knapp and Knapp, 1984	Technology-Professional	2	1	1982	3	1445	773	672	0.96
GZII	Guilford and Zimmerman, 1989	Mechanical	2	1	1989		215	97	118	1.28
WOWI	Ripley et al., 2001	Engineering and Related	2	1	1997		169,436	78,564	90,872	0.95
ENGINEERING TECHNICIAN										
OVIS	D'Costa et al., 1970	Appraisal	1	1	1969	2	46,002	23,151	22,851	1.19
CDI	Jackson, 1986	Science and Technology	1	1	1986	3	1000	500	500	0.98
CDI	Jackson, 2003	Science	1	1	2003	1	212	114	98	0.96
CDI	Jackson, 2003	Science	1	1	2003	2	737	385	352	0.74
CDI	Jackson, 2003	Science	1	1	2003	3	386	206	180	0.71
CDI	Jackson, 2003	Science	1	1	2003	4	392	171	221	0.87
CDI	Jackson, 2003	Science	1	1	2003	5	317	148	169	0.93
CDI	Jackson, 2003	Science	1	1	2003	5	276	145	131	0.75

(Continued)

Table 1 | Continued

Inventory	References	Scale	Moderator				Sample size			Effect size
			Complexity	Gender-balanced	Sample year	Age group	N	MN	FN	d
MECHANICS AND ELECTRONICS										
SII	Donnay et al., 2005	Mechanics and Construction	1	1	2002	5	2250	1125	1125	1.02
SII	Donnay et al., 2005	Computer Hardware and Electronics	1	1	2002	5	2250	1125	1125	0.77
SII	Harmon et al., 1994	Mechanical Activities	1	1	1994	5	18,951	9484	9467	0.68
SII	Hansen and Campbell, 1985	Mechanical Activities	1	1	1985	5	600	300	300	0.70
SII	Campbell, 1974	Mechanical	1	1	1974	5	600	300	300	0.91
CAI-E	Johansson, 2003	Mechanical/Fixing	1	2	1986	5	900	450	450	0.54
CAI-E	Johansson, 2003	Electronics	1	2	1986	5	900	450	450	0.63
CAI-V	Johansson, 1984	Mechanical/Fixing	1	2	1976	5	1500	750	750	1.15
CAI-V	Johansson, 1984	Electronics	1	2	1976	5	1500	750	750	1.15
CISS	Campbell et al., 1992	Mechanical Crafts	1	1	1992	5	5241	3442	1799	1.02
IDEAS	Johansson, 1996	Mechanical/Fixing	1	2	1989	1	1770	820	950	1.15
IDEAS	Johansson, 1996	Mechanical/Fixing	1	2	1989	2	2891	1208	1683	1.15
IDEAS	Johansson, 1978	Mechanical/Fixing	1	1	1977	1	598	292	306	1.15
IDEAS	Johansson, 1978	Mechanical/Fixing	1	1	1977	2	3436	1755	1681	1.15
IDEAS	Johansson, 1978	Electronics	1	1	1977	1	598	292	306	1.15
IDEAS	Johansson, 1978	Electronics	1	1	1977	2	3436	1755	1681	1.20
KGIS-E	Kuder, 1964	Mechanical	1	1	1963	1	4109	2080	2029	2.12
KGIS-E	Kuder, 1964	Mechanical	1	1	1963	2	5704	2766	2938	1.83
KGIS-E	Kuder and Zytowski, 1988	Mechanical	1	1	1987	1	5894	2714	3180	1.86
KGIS-E	Kuder and Zytowski, 1988	Mechanical	1	1	1987	2	7113	3402	3711	1.91
KOIS	Kuder and Zytowski, 1991	Mechanical	1	1	1985		3214	1583	1631	1.14
KCS	Zytowski, 2007	Mechanical	1	1	2007		3619	1663	1956	0.97
CDI	Jackson, 1986	Industrial Arts	1	1	1986	3	1000	500	500	1.31
CDI	Jackson, 2003	Industrial Arts	1	1	2003	1	212	114	98	2.21
CDI	Jackson, 2003	Industrial Arts	1	1	2003	2	737	385	352	1.77
CDI	Jackson, 2003	Industrial Arts	1	1	2003	3	386	206	180	1.40
CDI	Jackson, 2003	Industrial Arts	1	1	2003	4	392	171	221	1.06
CDI	Jackson, 2003	Industrial Arts	1	1	2003	5	317	148	169	0.98
CDI	Jackson, 2003	Industrial Arts	1	1	2003	5	276	145	131	1.26
OASIS:IS	Parker, 2002	Mechanical	1	2	1991	2	1091	551	540	1.25
CCQ-S	Chronicle Guidance Publications, 1992	Mechanical	1	2	1990	1	1536	797	739	1.08
CCQ-L	Chronicle Guidance Publications, 1992	Mechanical	1	2	1990	2	1311	661	650	1.08
GOCL II	Gordon, 1981	Technology–Mechanical	1	1	1981	2	359	168	191	0.62
VRII	Vocational Research Institute, 1988	Mechanical	1	1	1985	2	856	429	427	1.51
VRII	Vocational Research Institute, 1988	Mechanical	1	1	1985	4	525	198	327	0.71
WOWI	Ripley et al., 2001	Mechanical and Electrical Work	1	1	1997		169,436	78,564	90,872	1.07
COMPUTER SCIENCE										
SII	Donnay et al., 2005	Programming and Information System	2	1	2002	5	2250	1125	1125	0.38

(Continued)

Table 1 | Continued

Inventory	References	Scale	Moderator				Sample size			Effect size
			Complexity	Gender-balanced	Sample year	Age group	N	MN	FN	
MATHEMATICS										
SII	Donnay et al., 2005	Mathematics	2	1	2002	5	2250	1125	1125	0.46
SII	Harmon et al., 1994	Mathematics	2	1	1994	5	18,951	9484	9467	0.24
SII	Hansen and Campbell, 1985	Mathematics	2	1	1985	5	600	300	300	0.26
SII	Campbell, 1974	Mathematics	2	1	1974	5	600	300	300	0.35
CAI-E	Johansson, 2003	Mathematics	2	2	1986	5	900	450	450	0.22
CISS	Campbell et al., 1992	Mathematics	2	1	1992	5	5241	3442	1799	0.35
JVIS	Jackson, 2000	Mathematics	2	1	1999	2	2380	1190	1190	0.55
JVIS	Jackson, 2000	Mathematics	2	1	1999	5	1120	560	560	0.45
JVIS	Jackson, 1977	Mathematics	2	1	1977	3	1000	500	500	0.55
APPLIED MATHEMATICS										
SII	Harmon et al., 1994	Data Management	1	1	1994	5	18,951	9484	9467	0.16
CAI-V	Johansson, 1984	Numbers	1	2	1976	5	1500	750	750	0.05
OVIS-II	Winefordner, 1983	Numerical	1	2	1980	1	8917	4425	4492	0.29
OVIS-II	Winefordner, 1983	Numerical	1	2	1980	2	6315	3109	3206	0.27
OVIS-II	Winefordner, 1983	Numerical	1	2	1980	3	2780	1052	1728	0.62
OVIS	D'Costa et al., 1970	Numerical	1	1	1969	2	46,015	23,164	22,851	0.52
IDEAS	Johansson, 1996	Mathematics	1	2	1989	1	1770	820	950	0.05
IDEAS	Johansson, 1996	Mathematics	1	2	1989	2	2891	1208	1683	0.00
IDEAS	Johansson, 1978	Numbers	1	1	1977	1	598	292	306	0.00
IDEAS	Johansson, 1978	Numbers	1	1	1977	2	3436	1755	1681	0.05
KGIS-E	Kuder, 1964	Computational	1	1	1963	1	4109	2080	2029	0.30
KGIS-E	Kuder, 1964	Computational	1	1	1963	2	5704	2766	2938	0.44
KGIS-E	Kuder and Zytowski, 1988	Computational	1	1	1987	1	5894	2714	3180	0.18
KGIS-E	Kuder and Zytowski, 1988	Computational	1	1	1987	2	7113	3402	3711	0.27
KOIS	Kuder and Zytowski, 1991	Computational	1	1	1985		3214	1583	1631	0.27
KCS	Zytowski, 2007	Computational	1	1	2007		3619	1663	1956	0.20

d, inverse variance weighted effect size; CAI-E, Career Assessment Inventory–Enhanced Version; CAI-V, Career Assessment Inventory–Vocational Version; CCQ-S, Chronicle Career Quest (Form S); CCQ-L, Chronicle Career Quest (Form L); CDI, Career Decision Inventory; CII-1, Career Interest Inventory (Level 1); CII-2, Career Interest Inventory (Level 2); CISS, Campbell Interest and Skill Survey; COPS, Career Occupational Preference System Interest Inventory; COPS-R, Career Occupational Preference System Interest Inventory–Revised; GOCL II, Gordon Occupational Check List II; GZII, Guilford–Zimmerman Interest Inventory; IDEAS, Interest Determination, Exploration and Assessment System; JVIS, Jackson Vocational Interest Survey; KGIS-E, Kuder General Interest Survey (Form E); KOIS, Kuder Occupational Interest Survey; KCS, Kuder Career Search with Person Match; OASIS:IS, Occupational Aptitude Survey and Interest Schedule: Interest Schedule; OVIS, Ohio Vocational Interest Survey; SII, Strong Interest Inventory; VII, Vocational Interest Inventory; VII-R, Vocational Interest Inventory–Revised; VRII, Vocational Research Interest Inventory; WOWI, World of Work Inventory. In the coding for job complexity, 1, technical level, 2, professional level. For item development strategy (Gender_balanced), 1 represents an overlap of male and female scores of less than 75% or cases in which more than 33% of the items have response differences larger than 15%; 2 represents an overlap of male and female scores from 75 to 85% or 10 to 33% of the items have response differences larger than 15%; 3 represents an overlap of male and female scores larger than 85% or in which no more than 10% of the items have response differences larger than 15%. Age group was coded as the following: 1, middle school students or 12–14 years old; 2, high school students or 15–18 years old; 3, college students or 19–22 years old; 4, emerging working adults or 23–30 years old; 5, experienced working adults or 31 years and older.

job requirements for over 900 U.S. occupations, including occupational interest profiles (OIPs; Rounds et al., 1999) and levels of required abilities (McCoy et al., 1999; Donsbach et al., 2003). The OIPs are organized using Holland's (1997) RIASEC typology for describing work environments. The scores on each OIP indicate how well the occupation represents the six types of work environments. For example, the *Realistic* score for an occupation indicates how characteristic the occupation is of a things-oriented work environment; the *Social* score for an occupation indicates

how descriptive the occupation is of a people-oriented work environment. Therefore, the *Realistic* and *Social* scores on the OIPs were used to represent the things- and people-orientations for an occupation, respectively. Both are on a scale from 1 to 7, with higher scores indicating stronger things- or people-orientation. The O*NET system includes scores on two types of quantitative ability required by each occupation: Mathematical Reasoning (i.e., the ability to choose the right mathematical methods or formulas to solve a problem) and Number Facility (i.e., the ability to

add, subtract, multiply, or divide quickly and correctly). Because scores for the two types of quantitative ability are highly correlated ($r > 0.90$), in the current study the average score was taken to represent the level of quantitative ability required by each occupation. The score for required level of quantitative ability ranged from 0 to 6, with higher scores indicating greater ability required.

The occupational-level characteristics and statistics were then aggregated to each of the 13 STEM fields following the SOC system. For example, the people-orientation, things-orientation, and required level of quantitative ability for *Physical Sciences* were calculated by averaging the information from all the occupations nested within it, including Astronomers and Physicists, Atmospheric and Space Scientists, Chemists and Materials Scientists, Environmental Scientists, and Geoscientists. The percentage of women in *Physical Sciences* was calculated by dividing the total number of females employed in the above occupations by the total number of males and females employed. For *Mathematics*, the occupational characteristics and statistics were calculated from the data for Mathematicians and Statisticians. Similar calculations were performed for the rest of the STEM fields.

CODING OF ADDITIONAL MODERATORS

As discussed previously, we coded the complexity of the activities measured by a basic interest scale (professional-level = 2, technical-level = 1). The age group of a sample was coded based on the sample description and mean age of the sample reported in an interest inventory technical manual (middle school students or 12–14 years old = 1, high school students or 15–18 years old = 2, college students or 19–22 years old = 3, emerging working adults or 23–30 years old = 4, and experienced working adults or 31 years and older = 5). The years of data collection were also identified from the interest inventory technical manuals, ranging from 1963 to 2007. Information on item development strategy, or the degree to which an interest inventory was developed to be gender-balanced, was obtained from Su et al. (2009) and was coded as the following: overlap of male and female interest scores was less than 75% or more than 33% of the items had response differences larger than 15% = 1; overlap of male and female scores was between 75 and 85% or 10 to 33% of the items had response differences larger than 15% = 2; overlap of male and female scores was larger than 85% or no more than 10% of the items had response differences larger than 15% = 3. Coding for these additional moderators, along with the sample sizes by gender and total sample size for each sample, are listed in Table 1.

ANALYTICAL PROCEDURES

To examine gender differences in basic interests across STEM fields, we first calculated the standardized mean difference between males and females (Cohen's d) for each basic interest scale. This step yielded a total of 173 effect sizes, presented in Table 1. In the case where an interest inventory had more than one basic interest scales assessing a STEM field (e.g., both a *Mechanics* scale and an *Electronics* scale for the field of Mechanics and Electronics), we averaged the effect sizes within sample to avoid statistical dependence, creating 168 independent effect sizes. We then followed the procedures outlined in Hedges and Olkin

(1985) and Lipsey and Wilson (2001) to calculate the standard error and inverse-variance weight for each effect size, correct the effect sizes for small-sample-size bias, and synthesize the effect sizes. As discussed previously, we expected to find heterogeneity among the effect sizes. Instead of focusing on the grand mean gender difference in interests across all STEM fields, the main goal of our study was to understand how the average gender difference in interests varies by STEM field. Therefore, we conducted a meta-analytic analog of (inverse-variance weighted) analysis of variance (ANOVA) to compare the average gender differences in interests in different STEM fields, using a mixed-effects model (cf. Viechtbauer, 2008, for the rationale to start with the mixed-effects model for meta-analyses that are focusing on moderators).

Next, to understand the occupational characteristics associated with gender differences in interests and other variables that potentially moderate the effect sizes, we conducted an inverse-variance weighted meta-regression to evaluate the effects of the people-orientation, things-orientation, required level of quantitative ability, and job complexity (professional- vs. technical-level) of a STEM field as well as the age group of a sample and the year of data collection, again using a mixed-effects model. The weighted ANOVA and weighted meta-regression analysis were both performed using the statistical macros provided by Wilson (2005).

To examine the relationship between the gender differences in interests and gender composition within STEM fields, we conducted correlation and regression analyses at the occupational level, using occupational characteristics and statistics aggregated to the 13 STEM fields.

RESULTS

As expected, we found gender differences in interests to be heterogeneous and to vary largely across the 13 STEM fields. We summarized the effect sizes of gender differences in interests by STEM field in Table 2. In addition to the weighted mean effect size, d , we reported k , the number of effect sizes used to compute each mean effect size, N , the number of total respondents within a STEM field, as well as the 95% confidence interval and 90% credibility values for each mean effect size². A positive d value indicates that men had stronger interests in the STEM field than women and a negative d -value indicates stronger interests for women.

The most notable finding was that gender differences in interests varied greatly by STEM field: the largest gender differences in interests were observed in Engineering disciplines ($d = 0.83, 0.89$, and 1.21 for Engineering—professional level, Engineering Technicians, and Mechanics and Electronics, respectively), favoring men. In contrast, no significant gender differences in interests were found in Biological and Medical sciences, neither in the technical aspects of scientific activities. In Social Sciences and Medical Services, arguably the most

²Note that there is only one basic interest scale that specifically assessed interest in Computer Science. Therefore, inferential statistics could not be calculated. However, we reported the single effect size as it was from one of the most highly regarded, well-sampled interest inventory—the Strong Interest Inventory (Donnay et al., 2005)—and provided useful information as a reference in the study.

Table 2 | Weighted mean effect sizes and distribution of heterogeneity by STEM field.

Basic interest scale	<i>k</i>	<i>N</i>	<i>d</i>	<i>SE</i>	Lower 95%CI	Upper 95%CI	Lower 90%CV	Upper 90%CV	<i>Q_W</i>	<i>p</i>
Physical Sciences	15	57,669	0.56	0.06	0.43	0.69	0.09	1.03	16.84	0.26
Natural Sciences	6	15,436	0.41	0.11	0.21	0.62	0.17	0.66	1.66	0.89
Biological Science	3	4,500	0.19	0.15	−0.11	0.48	0.10	0.27	0.08	0.96
Medical Science	16	98,919	−0.04	0.06	−0.17	0.08	−0.28	0.20	4.74	0.99
Social Sciences	10	98,022	− 0.33	0.08	− 0.49	− 0.17	− 0.63	− 0.03	4.62	0.87
Science Technicians	5	23,846	0.14	0.12	−0.09	0.37	0.01	0.27	0.37	0.98
Medical Services	26	193,130	− 0.40	0.05	− 0.50	− 0.30	− 0.77	− 0.03	18.84	0.80
Engineering	22	355,531	0.83	0.06	0.72	0.94	0.29	1.36	32.88	0.05
Engineering Technicians	8	49,322	0.89	0.10	0.71	1.08	0.64	1.15	2.59	0.92
Mechanics and Electronics	31	255,508	1.21	0.05	1.12	1.30	0.50	1.92	78.15	0.00
Computer Science	1	2,250	0.38	—	—	—	—	—	—	—
Mathematics	9	33,042	0.38	0.09	0.21	0.55	0.17	0.59	1.95	0.98
Applied Mathematics	26	122,826	0.23	0.06	0.11	0.35	−0.07	0.53	7.70	0.94
Within-field heterogeneity (<i>Q_W</i>):								170.42 (<i>df</i> = 155)		0.19
Between-field heterogeneity (<i>Q_B</i>):								776.45 (<i>df</i> = 12)		0.00
Total heterogeneity (<i>Q_T</i>):								946.88 (<i>df</i> = 167)		0.00

k, Number of effect sizes; *N*, number of respondents; *d*, inverse variance weighted effect sizes, a positive *d*-value indicates gender difference favoring men and a negative *d*-value indicates gender difference favoring women; *SE*, standard error for *d*; *CI*, confidence interval; *CV*, credibility value; *Q*, heterogeneity statistic; *p*, probability of significance value associated with the *Q* statistic; bolded confidence intervals and credibility values indicate that 0 is not included within the interval; bolded *Q* statistic and corresponding *p*-value indicate that there was significant total heterogeneity between studies and significant heterogeneity among the effect sizes across STEM fields.

people-oriented fields, women exhibited stronger interests than men ($d = -0.33$ and -0.40 , respectively). Importantly, results from inverse-variance weighted ANOVA showed that the majority of heterogeneity among the effect sizes was introduced by dissimilarities between STEM fields ($Q_B = 776.45$, $df = 12$, $p < 0.001$), rather than from within STEM fields ($Q_W = 170.42$, $df = 155$, $p = 0.19$). The observed gender differences in interests within each STEM field were homogeneous for 11 of the 13 STEM fields. Only two exceptions—Engineering, and Mechanics and Electronics—had significant within-field variations, with effect sizes ranging 0.19 to 1.55 for Engineering and from 0.54 to 2.21 for Mechanics and Electronics.

Table 3 presents findings from the meta-regression on the effects of covariates of gender differences in interests, including the people-orientation, things-orientation, level of quantitative ability required, and job complexity (professional- vs. technical-level) of a STEM field as well as the age group of a sample and year of data collection. Consistent with our hypotheses 1a and 1b, gender differences of interests in various STEM fields can be explained by the people-orientation and things-orientation of the disciplines. The size of gender differences in interests (favoring men) increased with higher things-orientation of a STEM field ($B = 0.18$, $\beta = 0.48$, $p < 0.001$) and decreased with higher people-orientation ($B = -0.19$, $\beta = -0.60$, $p < 0.001$). In contrast, the level of quantitative ability required did not predict differential interests of men and women in a STEM field ($B = 0.02$, $\beta = 0.03$, $p = 0.68$). Hypothesis 2 was also supported. Job complexity and gender-balanced item development strategy each

had a small effect (smaller gender differences in interests at the professional level compared to the technical level, $B = -0.10$, $\beta = -0.08$, $p = 0.11$, and small gender differences in interests with more aggressive gender-balanced item development strategy, $B = -0.08$, $\beta = -0.07$, $p = 0.08$), yet neither was significant at the $p < 0.05$ level. The age group of a sample and the year of data collection did not influence the size of gender differences in interests. The meta-regression model (primarily people- and things-orientations) explained 76.98% of the total between-study heterogeneity ($Q_M = 532.87$, $df = 7$, $p < 0.001$) and the residual heterogeneity was not significant ($Q_E = 159.37$, $df = 150$, $p = 0.28$), indicating that people-orientation and things-orientation of the STEM fields were the main contributors to the variation in effect sizes across STEM fields.

Finally, we looked at the gender composition in STEM occupations and examined its association with gender differences in interests and various occupational characteristics. In **Table 4**, we report the percentage of women by STEM field, along with the level of quantitative ability required, things-, and people-orientations for each field. We again present the effect size of gender difference in interests (d) for each STEM field and report two additional statistics³ associated with d : (1) We calculated the percentage of overlap (Bhattacharyya coefficient) between male and female interest distributions given the effect size of gender difference in interests for each STEM field. This statistic provides

³Syntaxes for calculating these statistics in R are available from the first author upon request.

Table 3 | Meta regression coefficients for covariates of gender differences in STEM interests.

Model	<i>B</i>	<i>SE</i>	Lower 95%CI	Upper 95%CI	<i>Z</i>	β	<i>p</i>
Constant	3.46	4.815	−5.98	12.90	0.72	0.00	0.47
Things-orientation	0.18	0.024	0.14	0.23	7.65	0.48	0.00
People-orientation	−0.19	0.018	−0.22	−0.15	−10.67	−0.60	0.00
Required level of quantitative ability	0.02	0.053	−0.08	0.13	0.41	0.03	0.68
Job complexity	−0.10	0.062	−0.22	0.02	−1.60	−0.08	0.11
Gender-balanced item development strategy	−0.08	0.048	−0.18	0.01	−1.76	−0.07	0.08
Year of data collection	−0.00	0.002	−0.01	0.00	−0.66	−0.03	0.51
Age group of sample	−0.02	0.017	−0.05	0.02	−1.02	−0.04	0.31
Model explained heterogeneity (Q_M):					532.87 (<i>df</i> = 7)		0.00
Residual heterogeneity (Q_E):					159.37 (<i>df</i> = 150)		0.28
Total heterogeneity (Q_T):					692.24 (<i>df</i> = 157)		0.00
R2 analog (amount of heterogeneity accounted for):							76.98%

B, unstandardized regression coefficient; *SE*, standard error for *B*; *CI*, confidence interval; *Z*, standard score for *B*; β , standardized regression coefficient; *p*, probability of significance value for regression coefficients; bolded confidence intervals indicate that 0 is not included within the interval; bolded *Q*-statistic and corresponding *p*-value indicate that total heterogeneity between studies was significant and the model explained a significant amount of heterogeneity.

Table 4 | Occupational characteristics, gender differences in interests, and percentage of females by STEM field.

STEM Field	Quantitative ability	Things- orientation	People- orientation	M–F <i>d</i> in interests	M–F interest overlap (%)	<i>p</i> (<i>F</i>) in top 10% interests (%)	<i>p</i> (<i>F</i>) in STEM field (%)
Physical Sciences	4.14	4.89	1.67	0.56	77.84	35.79	32.80
Natural Sciences	4.00	4.82	1.64	0.41	83.64	39.47	35.41
Biological Science	3.68	4.67	1.58	0.19	92.54	45.18	45.74
Medical Science	3.07	4.61	5.13	−0.04	98.30	51.10	43.13
Social Sciences	3.09	2.22	3.39	−0.33	86.76	58.53	58.46
Science Technicians	3.16	5.27	1.76	0.14	94.47	46.45	44.66
Medical Services	2.59	4.02	6.41	−0.40	84.25	60.14	89.41
Engineering	3.97	5.75	1.46	0.83	67.91	29.61	10.98
Engineering Technicians	3.27	6.26	1.11	0.89	65.46	28.12	12.18
Mechanics and Electronics	2.09	6.95	1.23	1.21	54.47	21.61	2.91
Computer Science	3.33	4.51	1.76	0.38*	84.93	40.30	23.75
Mathematics	5.44	2.00	1.00	0.38	84.86	40.25	39.58
Applied Mathematics	4.21	1.97	1.86	0.23	90.80	44.06	58.57

*Estimated based on one interest inventory. M–F, Male–Female; *d*, inverse variance weighted effect sizes, a positive *d*-value indicates gender difference favoring men and a negative *d*-value indicates gender difference favoring women; *p*(*F*), percentage of females.

an additional, intuitive metric to represent the similarity and dissimilarity of men's and women's interests. A higher percentage of overlap indicates more similar interests between men and women, and a lower percentage of overlap indicates more dissimilar interests. (2) We calculated the percentage of women within the top 10% of the total population in the interest distribution. This statistic provides an index on how well women are represented among those who are most strongly interested in a STEM field. Assuming that individuals at the right tail (highest 10%) of a population interest distribution are likely to choose a career in

that basic interest area (e.g., *Mathematics*), this statistic also represents the hypothetical/projected percentage of women who would work in a STEM field given the gender difference in interests. These statistics can provide further insight into men and women's differential interests in various STEM fields and a more straightforward comparison with the actual gender distribution in each field.

Table 5 presents the correlations among occupational characteristics, gender differences in interests, and the percentages of women across STEM fields. As expected, people-orientation and

Table 5 | Correlations among occupational characteristics, gender differences in interests, and percentages of females in STEM fields.

	Quantitative ability	People-orientation	Things-orientation	M-F <i>d</i> in interests
People-orientation	−0.44			
Things-orientation	−0.58	−0.21		
M-F <i>d</i> in interests	0.00	−0.73**	0.65*	
Percentage of females	−0.00	0.72**	−0.66*	−0.89**

**Correlation is significant at the 0.01 level; *Correlation is significant at the 0.05 level.

things-orientation were associated with the percentage of women in a STEM field ($r = 0.72$, $p < 0.01$, and $r = -0.66$, $p < 0.05$, respectively). The percentages of women were higher in STEM fields that are more people-oriented and less things-oriented. The percentages of women in STEM fields were also very strongly correlated with gender differences in interests ($r = -0.89$, $p < 0.01$). The percentages of women were higher in STEM fields in which men and women were more equally interested or those for which women had stronger interests.

Further, hierarchical regression analysis showed that, after controlling for the effect of gender differences in interests, the effect of people-orientation decreased substantially and was no longer significant ($\beta = 0.14$, $p = 0.50$ for people-orientation; $\beta = -0.79$, $p < 0.01$ for gender differences in interests). Similarly, after controlling for the effect of gender differences in interests, the effect of things-orientation decreased substantially and was no longer significant ($\beta = -0.14$, $p = 0.47$ for things-orientation; $\beta = -0.80$, $p < 0.01$ for gender differences in interests). These results indicated that the effects of people- and things-orientations on the gender composition (percentage of women) in STEM fields were mediated through the differential interests of men and women. Hypotheses 3, 4a, and 4b were supported. Consistent with Hypothesis 5, the percentage of women in a STEM field was not associated with the level of quantitative ability required by the field.

To visualize the relationship between gender differences in interests and the gender composition across STEM fields, we plotted the projected percentages of women given the gender differences in interests in comparison with the actual percentages of women in various STEM fields in **Figure 1**. As shown in **Figure 1**, the actual percentages of women closely mirror the projected percentages of women given the gender differences in interests in *Mathematics* and the sciences (*Physical Sciences*, *Natural Sciences*, *Biological Science*, *Medical Science*, *Social Sciences*, and *Science-Technicians*). However, the actual percentages of women fell short of the predicted percentages based on interests in the Engineering disciplines (*Engineering*, *Engineering-Technicians*, *Mechanics and Electronics*, and *Computer Science*). The percentages of women exceeded the predicted percentages based on interests in *Applied Mathematics* and *Medical Services*. These results suggest that men and women's participation in these fields were potentially influenced by factors other than interests.

DISCUSSION

Increasing the representation of women in the STEM workforce poses one of the most critical challenges for our society. To date, research to understand gender disparities in STEM careers typically treated all the STEM fields as a whole and emphasized the similarities among STEM fields rather than their dissimilarities. We argue that STEM fields are heterogeneous. Understanding men's and women's career choices across different STEM fields is as meaningful as understanding the career choices between STEM and non-STEM fields. Therefore, we examined gender differences in basic interests across different STEM fields.

We found drastically different levels of gender differences in basic interests within STEM fields. Large to very large gender differences in interests favoring men were observed in engineering-related fields ($d = 0.83$ for *Engineering*—professional level, $d = 0.89$ for *Engineering Technicians*, and $d = 1.21$ for *Mechanics and Electronics*). Small to moderate gender differences in interests favoring men were observed for mathematical careers ($d = 0.38$ for *Mathematics*, and $d = 0.23$ for *Applied Mathematics*). Gender differences in interests vary largely in the sciences, ranging from moderate, favoring men, in *Physical Sciences* ($d = 0.56$), to non-significant ($d = 0.19$ for *Biological Science*, $d = 0.14$ for *Science Technicians*, and $d = -0.04$ for *Medical Science*), and to small to moderate, favoring women ($d = -0.33$ for *Social Sciences*, and $d = -0.40$ for *Medical Services*). These findings provide refined information about men and women's interests in sub-disciplines of STEM. Measuring interests at the basic interest level can produce tailored results about career preferences and can facilitate career guidance for individuals in choosing a STEM career that best matches their interests. Researchers may also gain a clearer understanding of the relationship between interests and career choices by using basic interest measures.

Through investigating gender differences in basic interests across various STEM fields and the occupational characteristics associated with these gender differences in interests, we offer a preference-based explanation for why women are underrepresented in some STEM fields, but not others. Specifically, we argue that individuals' interests are powerful predictors of their occupational membership. Individuals are oriented toward work environments that are congruent with their interests. Men's and women's differences in basic interests lead to unbalanced gender composition in different sectors of the world of work. Two interest dimensions—*Realistic* interests (interest in working with things and gadgets) and *Social* interests (interest in working with people and helping people) may be the most salient in characterizing men and women's differential career preferences, with men having substantially stronger interests in working with things and women preferring working with people. As such, there tend to be larger gender differences in interests (favoring men) for more things-oriented and less people-oriented occupational fields. Overall, STEM fields tend to be high in things-orientation and low in people-orientation. As a result, women on average are less likely to be interested in STEM fields than men, which translate to the lower percentages of women in the STEM workforce. Nonetheless, because STEM disciplines also vary in their things- and people-orientation, women tend to gravitate toward more people-oriented fields within STEM, such as

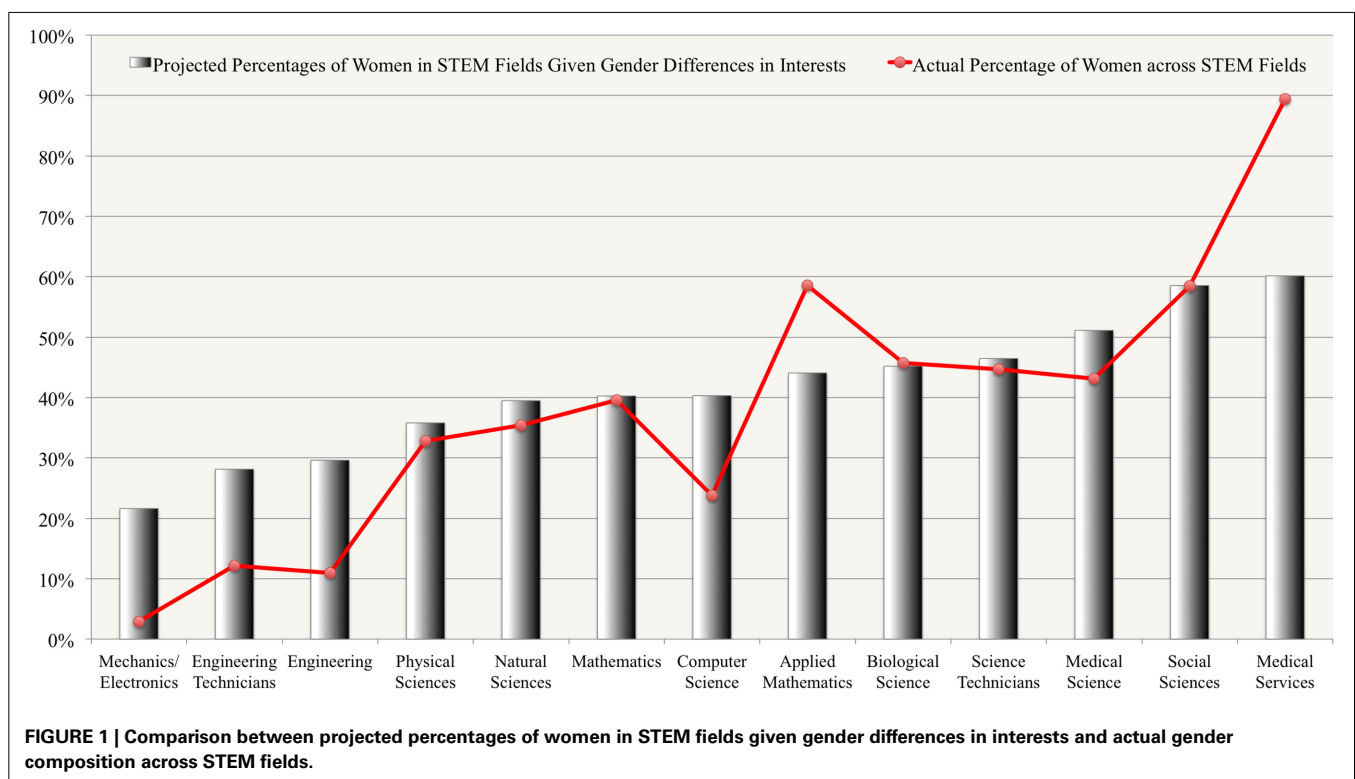
Medical Science and *Social Sciences*, as a function of higher *Social* interests.

The current study found the percentages of women within most STEM fields to mirror the gender differences in basic interests in those fields, lending support to the preference-based explanation for gender disparities in STEM careers. Although the projected percentages of females in STEM fields based on interests are only approximations, they provide useful yardsticks for comparing different STEM fields. Information from **Figure 1** allows us to identify sub-disciplines of STEM where the shortages of women reflect gender differences in interests and other sub-disciplines where the underrepresentation of women exhibits unexpected patterns. For example, in mathematics and sciences, the actual gender composition is closely aligned with gender differences in interests; however, there are discrepancies between the projected percentages of women based on interests and the actual gender composition in the engineering-related fields and *Medical Services*. The actual percentages of women in engineering-related fields (10.98% for *Engineering*—professional level, 12.18% for *Engineering Technicians*, merely 2.91% for *Mechanics and Electronics*) are even lower than what would be expected based on women's lower interests than men (29.61, 28.12, and 21.61%, respectively). In contrast, the actual percentage of women in *Medical Services* (89.41%) largely exceeded what would be expected based on women's higher interests in this field (60.14%). These results indicate the existence of other factors that escalated the gender disparities in these STEM careers. A few potential factors suggested by the literature include preference for work-life balance (e.g., Ferriman et al., 2009), gender stereotyping and gender role schema in individuals' career

decision-making (e.g., Konrad et al., 2000), and implicit bias in employers' selection process (e.g., Moss-Racusin et al., 2012). It is beyond the scope of this article to provide a detailed review of these alternative factors contributing to the gender disparities in the STEM fields (for a comprehensive review, see Ceci et al., 2014). However, the current study points out specific STEM fields where attention to these alternative influences may be most fruitful.

Despite the importance of quantitative ability for STEM careers, we showed that the level of quantitative ability required by a STEM discipline was not associated with men and women's differential interests and representation in that field. To clarify, this result does not mean that quantitative ability is not a consideration in STEM career choices. Instead, it means that the consideration of quantitative ability *at the occupational level* is *equally* important for men and women when choosing a STEM career. *At the individual level*, existing literature (e.g., Lubinski et al., 2001; Wai et al., 2010) has shown that individuals with higher quantitative ability, regardless of their gender, are interested in activities and work environments that require higher levels of quantitative ability and are more likely to choose an occupational field with higher quantitative ability requirement, such as the STEM fields. Individuals with lower quantitative ability, regardless of their gender, are not prepared for entering STEM careers.

Earlier in this article, we discussed the dynamic and reciprocal relationship between interests, knowledge acquisition, and ability development. As previously noted, interests serve as a source of intrinsic motivation for individuals to engage in the activities that they like and accumulate knowledge and skills associated



with these activities. Therefore, individuals' interests at an early age may have a profound influence on their ability development through directing the learning process. For example, a girl who is interested in people-oriented activities may choose to focus on classes and extracurricular activities that fulfill her people interests, such as social studies and volunteering at an animal shelter, and may avoid mathematics classes and activities that cultivate the development of quantitative skills and ability because they are low in people-orientation. The lack of development in quantitative skills and ability may further discourage her interest in math-related activities, which in turn impede future learning in these areas. In the long run, the girl may not be equipped with the quantitative skills and ability needed for her to be eligible for or successful in a people-oriented STEM field that she wants to pursue, such as medical science. Therefore, although the level of quantitative ability required in a STEM field was not found to differentially influence men's and women's interests and career choices, interests play a critical role in the early development of quantitative ability. Boys or girls who are disinterested and "turn off" learning in quantitative-related activities are equally unlikely to be successful in pursuing a STEM career later on. As such, (dis)interest *constrains* one's options in educational and occupational pursuits indirectly through affecting ability development.

On the other hand, some researchers have advanced a breadth-based model to explain women's underrepresentation in STEM fields (Valla and Ceci, 2014; also see, Lubinski and Benbow, 2006). The breadth-based model states that females, more likely than males, have interests that promote the development of more symmetrical, competing levels of quantitative and verbal abilities, which in turn afford them with broader career choices. As a result, more females may opt for careers that allow them to express their verbal and people-related skills and abilities, such as law or social sciences, even when they also have the interests and adequate quantitative ability to pursue other STEM fields. This perspective is consistent with empirical findings such as those reported in Woodcock et al. (2013) that people-orientation moderated the relationship between things-orientation and the choice of a STEM major such that students high in things-orientation are less likely to choose a STEM major when their people-orientation is also high. Similarly, Wang et al. (2013) analyzed data from a longitudinal study and reported that mathematically capable twelfth graders who also had high verbal skills were less likely to pursue STEM careers when they were 33 years old than were individuals who had high math skills but moderate verbal skills. Because women were overrepresented in the high math and high verbal skills group, fewer mathematically talented women entered STEM careers compared to their male peers. Therefore, according to the breadth-based model, interests do not constrain but rather *broaden* women's career choices through influencing more balanced ability development.

We acknowledge that both processes—*constraining* and *broadening*—may happen in a parallel manner. As discussed earlier in this article and in another paper (Su et al., 2009), individuals engage in both inter-personal and intra-personal comparisons while making educational and career choices. The *constraining* process happens, from the inter-personal perspective, when

individuals are selected out or self-select out of STEM fields for not having high quantitative ability compared to other individuals; the *broadening* process happens, from the intra-personal perspective, when individuals evaluate multiple interests and talents within themselves and weigh other options besides STEM careers. Therefore, we urge researchers to examine the indirect effect of interests on the educational and career attainment in STEM fields through learning and ability development in addition to the direct influence of interests on STEM career choices.

POTENTIAL INTERVENTIONS TARGETING INTERESTS AND WORK ENVIRONMENTS

The current findings have implications for potential interventions to increase women's representation in the STEM workforce. At the individual level, the current findings suggest that interests are critical predictors of occupational membership in STEM fields. Highlighting the societal relevance of STEM knowledge, skills, and careers and their value in improving people's lives may prove to be an effective way for appealing to females' *Social* interests and getting more females to engage in STEM activities (Eccles, 2009; Valla and Ceci, 2014). For example, it has been demonstrated that using a science-technology-society (STS) approach to teaching science in high school improved attitudes toward science, particularly for girls (Bennett et al., 2007). In another experimental study, Harackiewicz et al. (2012) showed that mailing parents brochures about how to help their adolescents in tenth and eleventh grade see the values of mathematics and science to their personal lives increased the adolescents' mathematics and science course-taking by almost one semester. These interventions provide promising ways for educators and parents to increase students' interests and engagement in STEM activities. We note, however, that evidence on the effectiveness of these interventions is still preliminary. More research is needed to quantify the effect sizes of the improvement in students' attitudes, interests, and behaviors, and in particular, the long-term outcomes of the interventions, such as participation in the workforce. We call for more interventions that integrate students' people interests into STEM education and that increase students' perception of task values of STEM activities and careers, as well as more research that uses a longitudinal design to evaluate such interventions. On the other hand, while the literature has consistently shown the influence of social contexts (e.g., parents, schools) on students' interest development, particularly the development of differential interests for boys and girls (e.g., Hartung et al., 2005; Jacobs et al., 2005), little is known about the link between biological factors (e.g., brain structure, hormones) and interest development. To the extent that gender differences in interests are explained by biological factors, the effectiveness of social and educational interventions for increasing girls' interests in STEM fields may be constrained. More research is needed to provide a comprehensive picture of why such large gender differences in interests exist and how they are developed.

Moreover, we note that the timing is important for an intervention that targets individuals' interests, particularly given the relationship between interests and ability development. Few interventions to date have reached individuals before high school, yet the mutual influence between interests and ability development

start from a much earlier age. Although little research has examined the career exploration and interest development of preadolescent children, the research that does exist suggests that children do use their interests to guide learning and formulation of career goals before reaching teen-age years (Hartung et al., 2005). For example, a study that surveyed the finalists in the Westinghouse Science Competition and members of the National Academy of Sciences (NAS) showed that the respondents were already certain about their interest in science at an average age of 11 years old, and as early as 4 years old (Feist, 2006). The awareness of interests early on promoted early engagement in research experiences and in turn contributed to the development of scientific talent and lifetime research productivity. Further, children's perceptions of occupations including the traditional sex-type of occupations also start to form during grade school years, which contribute the development of their differential career preferences (Hartung et al., 2005). It was reported that children as young as 4 years of age express occupational preferences along sex-based distinctions (Trice and Rush, 1995). Given this research, we assert that interventions aiming at increasing individuals' interests in STEM fields and reforming individuals' perceptions of STEM careers need to occur at early ages.

Given the importance of interests for individuals' cognitive development and career exploration starting from an early age, it is necessary to assess interests periodically when they begin to form. Just as a standardized achievement test or other types of cognitive assessments that give students, parents, and educators feedback regarding the students' knowledge acquisition and skill development, measuring interests at a regular basis would provide students, parents, and educators with information regarding students' interest development that can be used to guide students' involvement in curricular and non-curricular activities and to facilitate students' career exploration. We propose a national barometer of basic interests to be developed and administered in K-12 education annually. Such an index would be particularly useful for monitoring the development of gender differences in interests and for guiding girls with STEM interests to engage in STEM activities and explore STEM careers.

At the institutional level, work environments in STEM fields can be reconstructed to increase their people-orientation and to better fulfill women's people interests. Although the analyses in the current study used STEM fields as units and focused on the heterogeneity in people-orientation across STEM fields, we note that the work environments within a STEM discipline can vary as well. For example, different universities or different organizations may have different culture, climate, and practices that provide individuals with different experiences. These salient, more proximal environments are likely to have the largest impact on individual behaviors when assessing their fit with individual interests (Holland, 1997). To the extent an academic program or an organization can implement interventions that enhance its people-orientation, such as incorporating mentoring and team-working activities and emphasizing communication (e.g., Seat et al., 2001), women would be more likely to find such work environment congruent with their interests and are more likely to choose and stay in such work environment. More research is needed to examine the effectiveness of such workplace interventions on

women's career choice, job satisfaction, and retention in STEM fields.

LIMITATIONS AND FUTURE RESEARCH

As we have mentioned earlier, findings from the current study are based on occupational level of analysis and should only be interpreted at the occupational level. An individual level analysis may reveal greater role of quantitative ability for STEM careers, as previous literature suggested. Nonetheless, given findings from the current study, we expect people and things interests at the individual level to strongly influence individuals' career choice and attainment and expect these relationships to account for the effect of gender.

The current study categorized basic interest scales into 13 STEM fields. While we have demonstrated how these 13 STEM fields differ from each other, the basic interest data did not allow us to perform comparisons of STEM occupations at a more refined level. Even within a sub-discipline of STEM, we may still identify occupations that are heterogeneous in terms of their occupational characteristics. For examples, economics and psychology are both nested within social sciences, yet economics is higher in its level of required quantitative ability (4.25 compared to 3.08 for Psychology) and things-orientation (2.33 compared to 1.70) and is substantially lower in its people-orientation (1.67 compared to 5.04). We expect these differences in occupational characteristics to influence the gender differences in interests and the actual gender composition in economics and psychology. Indeed, women constitute a much smaller percentage among the economists compared to psychologists (21.74% compared to 77.42%). We expect the current findings to replicate in future research examining STEM occupations at a finer level.

Lastly, the current findings are correlational and no causal inferences should be made. By conducting a meta-analysis and pooling together many different "slices" at different stages of the developmental process, we partially alleviated the limitations of using cross-sectional data and showed that age did not moderate the size of gender differences in basic interests. However, to truly understand how interests and cognitive ability unfold and interact to influence individuals' career development, more longitudinal studies like Lubinski and Benbow (2006) are needed in the future. To replicate and complement the findings from the current study, experimental studies are needed to establish causal relationships between the things- and people-orientations of a work environment and individuals' interests and career choices.

CONCLUSION

To understand the reasons for women's underrepresentation in STEM fields, more attention needs to be paid to interests. In the current study, we showed that women's interests in more people-oriented, and less things-oriented work environments was a key factor that influenced their career choice in STEM fields. Importantly, not only the choices between STEM and non-STEM careers but also the choices within STEM careers reflect individuals' interest patterns. Interventions at the individual level targeting the development of interests and those at the institutional level aiming at creating educational and work environments that better accommodate women's people interests may prove to be

fruitful. In addition, findings from the current study highlight the discrepancies in some STEM fields where the number of women did not meet their level of interests, indicating other factors at work. Realizing that the issue of women's underrepresentation is not identical across all STEM fields and the mechanisms contributing to the gender disparities are overlapping yet different is important for designing future investigations and interventions to understand and increase women's representation in STEM using a multivariate approach.

REFERENCES

- Ackerman, P. L. (1996). A theory of adult intellectual development: process, personality, interests, and knowledge. *Intelligence* 22, 227–257. doi: 10.1016/S0160-2896(96)90016-1
- Bennett, J., Lubben, F., and Hogarth, S. (2007). Bringing science to life: a synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Sci. Educ.* 91, 347–370. doi: 10.1002/sci.20186
- Campbell, D. P. (1971). *Handbook for the Strong Vocational Interest Blank*. Stanford, CA: Stanford University Press.
- *Campbell, D. P. (1974). *Manual for the Strong–Campbell Interest Inventory: T325 (Merged Form)*. Stanford, CA: Stanford University Press.
- *Campbell, D. P., Hyne, S. A., and Nilsen, D. L. (1992). *Manual for the Campbell Interest and Skill Survey: CISS*. Minneapolis, MN: National Computer Systems.
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: explaining the gap. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., William, W. M., Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412
- *Chronicle Guidance Publications. (1992). *Chronicle Career Quest: Technical Manual*. Moravia, NY: Chronicle Guidance Publications.
- *D'Costa, A. G., Winefordner, D. W., Odgers, J. G., and Koons, P. B. Jr. (1970). *Ohio Vocational Interest Survey Manual for Interpreting*. New York, NY: Harcourt Brace Jovanovich.
- Dawis, R. V., and Lofquist, L. H. (1984). *A Psychological Theory of Work Adjustment*. Minneapolis, MN: University of Minnesota Press.
- Denissen, J. J., Zarrett, N. R., and Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Dev.* 78, 430–447. doi: 10.1111/j.1467-8624.2007.01007.x
- Diekmann, A. B., Brown, E. R., Johnston, A. M., and Clark, E. K. (2010). Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342
- Diekmann, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., and Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to STEM careers: evidence for a goal congruity perspective. *J. Pers. Soc. Psychol.* 101, 902–918. doi: 10.1037/a0025199
- *Donnay, D. A. C., Morris, M. L., Schaubhut, N. A., and Thompson, R. C. (2005). *Strong Interest Inventory Manual: Research, Development, and Strategies for Interpretation*. Mountain View, CA: CPP.
- Donsbach, J., Tsacoumis, S., Sager, C., and Updegraff, J. (2003). *O*NET Analyst Occupational Abilities Ratings: Procedures*. Raleigh, NC: National Center for O*NET Development. Available online at: http://www.onetcenter.org/dl_files/AnalystProc.pdf
- Eccles, J. S. (1994). Understanding women's educational and occupational choices: applying the Eccles et al. model of achievement-related choices. *Psychol. Women Q.* 18, 585–609. doi: 10.1111/j.1471-6402.1994.tb01049.x
- Eccles, J. S. (2007). “Where are all the women? Gender differences in participation in physical science and engineering,” in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, eds S. J. Ceci, and W. M. Williams (Washington, DC: American Psychological Association), 199–210. doi: 10.1037/11546-016
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368
- Feist, G. J. (2006). How development and personality influence scientific thought, interest, and achievement. *Rev. Gen. Psychol.* 10, 163–182. doi: 10.1037/1089-2680.10.2.163
- Ferriman, K., Lubinski, D., and Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: developmental changes and sex differences during emerging adulthood and parenthood. *J. Pers. Soc. Psychol.* 97, 517–532. doi: 10.1037/a0016030
- *Gordon, L. V. (1981). *Manual for the Gordon Occupational Check List II*. Cleveland, OH: Psychological Corporation.
- Graziano, W. G., Habashi, M. M., and Woodcock, A. (2011). Exploring and measuring differences in person-thing orientation. *Pers. Individ. Dif.* 51, 28–33. doi: 10.1016/j.paid.2011.03.004
- *Guilford, J. S., and Zimmerman, W. S. (1989). *Manual for the Guilford–Zimmerman Interest Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- *Hansen, J. C., and Campbell, D. P. (1985). *Manual for the Strong Interest Inventory: Form T325 of the Strong Vocational Interests Blanks, 4th Edn.* Stanford, CA: Stanford University Press.
- Harackiewicz, J. M., Rozek, C. R., Hulleman, C. S., and Hyde, J. S. (2012). Helping parents motivate their teens in mathematics and science: an experimental test. *Psychol. Sci.* 23, 899–906. doi: 10.1177/0956797611435530
- *Harmon, L. W., Hansen, J. C., Borgen, F. H., and Hammer, A. L. (1994). *Strong Interest Inventory Applications and Technical Guide*. Stanford, CA: Stanford University Press.
- Hartung, P. J., Porfeli, E. J., and Vondracek, F. W. (2005). Child vocational development: a review and reconsideration. *J. Vocat. Behav.* 66, 385–419. doi: 10.1016/j.jvb.2004.05.006
- Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. doi: 10.1126/science.7604277
- Hedges, L. V., and Olkin, I. (1985). *Statistical Method for Meta-Analysis*. Orlando, FL: Academic Press.
- Hidi, S. (2001). Interest, reading and learning: theoretical and practical considerations. *Educ. Psychol. Rev.* 13, 191–210. doi: 10.1023/A:1016667621114
- Hogan, R., and Blake, R. (1999). John Holland's vocational typology and personality theory. *J. Vocat. Behav.* 55, 41–56. doi: 10.1006/jvbe.1999.1696
- Holland, J. L. (1959). A theory of vocational choice. *J. Couns. Psychol.* 6, 35–45. doi: 10.1037/h0040767
- Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments, 3rd Edn.* Odessa, FL: Psychological Assessment Resources.
- Hyde, J. S., and Linn, M. C. (2006). Gender similarities in mathematics and science. *Science* 314, 599–600. doi: 10.1126/science.1132154
- Hyde, J. S., Fennema, E., and Lamon, S. (1990). Gender differences in mathematics performance. A meta-analysis. *Psychol. Bull.* 107, 139–155. doi: 10.1037/0033-2909.107.2.139
- *Jackson, D. N. (1977). *Jackson Vocational Interest Survey Manual*. Port Huron, MI: Research Psychologists Press.
- *Jackson, D. N. (1986). *Career Decisions Inventory Manual*. Port Huron, MI: Research Psychologists Press.
- *Jackson, D. N. (2000). *Jackson Vocational Interest Survey Manual, 2nd Edn.* Port Huron, MI: SIGMA Assessment Systems.
- *Jackson, D. N. (2003). *Career Decisions Inventory Manual, 2nd Edn.* Port Huron, MI: SIGMA Assessment Systems.
- Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J. S., and Malanchuk, O. (2005). “I can, but I don't want to”: the impact of parents, interests, and activities on gender differences in math,” in *Gender Differences in Mathematics: An Integrative Psychological Approach*, eds A. M. Gallagher, and J. C. Kaufman (New York, NY: Cambridge University Press), 246–263.
- *Johansson, C. B. (1978). *Manual for IDEAS: Interest Determination, Exploration and Assessment System*. Minneapolis, MN: Interpretive Scoring Systems.
- *Johansson, C. B. (1984). *Manual for Career Assessment Inventory, 2nd Edn.* Minneapolis, MN: National Computer Systems.
- *Johansson, C. B. (1996). *Manual for IDEAS: Interest Determination, Exploration and Assessment System, Rev. Edn.* Minneapolis, MN: National Computer Systems.
- *Johansson, C. B., (2003). *Career Assessment Inventory: The Enhanced Version Manual, Rev. Edn.* Minneapolis, MN: National Computer Systems.
- *Knapp, R. R., and Knapp, L. (1979). *Technical Manual for the Career Occupational Preference System Interest Inventory (Form R)*. San Diego, CA: EDITS.

- Knapp, R. R., and Knapp, L. (1984). *COPSystem technical manual*. San Diego, CA: EdITS.
- *Knapp, R. R., Knapp, L., and Knapp-Lee, L. (1990). *Manual for the COPS Interest Inventory*. San Diego, CA: EdITS.
- Konrad, A. M., Ritchie, J. E., Lieb, P., and Corrigan, E. (2000). Sex differences and similarities in job attribute preferences: a meta-analysis. *Psychol. Bull.* 126, 593–641. doi: 10.1037/0033-2909.126.4.593
- Kuder, G. F. (1964). *Kuder General Interest Survey (Form E) manual*. Chicago: Science Research Associates.
- Kuder, F. G. (1977). *Activity Interests and Occupational Choice*. Chicago, IL: Science Research Associates.
- *Kuder, F., and Zytowski, D. G. (1988). *Kuder General Interest Survey (Form E): Preliminary Manual*. Adel, IA: National Career Assessment Services.
- *Kuder, F., and Zytowski, D. G. (1991). *Kuder Occupational Interest Survey (Form DD): General Manual, 3rd Edn*. Adel, IA: National Career Assessment Services.
- Liao, H.-Y., Armstrong, P. I., and Rounds, J. (2008). Development and initial validation of public domain basic interest markers [Monograph]. *J. Vocat. Behav.* 73, 159–183. doi: 10.1016/j.jvb.2007.12.002
- Lipsey, M. W., and Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: “Sinking shaft at a few critical points.” *Annu. Rev. Psychol.* 51, 405–444. doi: 10.1146/annurev.psych.51.1.405
- Lubinski, D., and Benbow, C. P. (1992). Gender differences in abilities and preferences among the gifted: implications for the math-science pipeline. *Curr. Dir. Psychol. Sci.* 1, 61–66. doi: 10.1111/1467-8721.ep11509746
- Lubinski, D., and Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: uncovering antecedents for the development of math-science expertise. *Perspect. Psychol. Sci.* 1, 316–345. doi: 10.1111/j.1745-6916.2006.00019.x
- Lubinski, D., Benbow, C. P., Shea, D. L., Eftekhari-Sanjani, H., and Halvorson, M. B. (2001). Men and women at promise for scientific excellence: similarity not dissimilarity. *Psychol. Sci.* 12, 309–317. doi: 10.1111/1467-9280.00357
- *Lunneborg, P. W. (1981). *Vocational Interest Inventory (VII) Manual*. Los Angeles, CA: Western Psychological Services.
- *Lunneborg, P. W. (1993). *Vocational Interest Inventory—Revised (VII-R) Manual*. Los Angeles, CA: Western Psychological Services.
- McCarty, M. K., Monteith, M. J., Kaiser, C. R. (2014). Communally constrained decisions in workplace contexts. *J. Exp. Soc. Psychol.* 55, 175–187. doi: 10.1016/j.jesp.2014.07.007
- McCoy, R., Campbell, J., Oswald, F., Rivkin, D., and Lewis, P. (1999). *Generation and Use of Occupational Ability Profiles for Exploring O*NET Occupational Units*. Raleigh, NC: National Center for O*NET Development. Available online at: <http://www.onetcenter.org/reports/GenUseAP.html>
- Meece, J., Eccles-Parsons, J., Kaczala, C. M., Goff, S., and Futterman, R. (1982). Sex differences in math achievement: toward a model of academic choice. *Psychol. Bull.* 91, 324–348. doi: 10.1037/0033-2909.91.2.324
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty’s subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16474–16479. doi: 10.1073/pnas.1211286109
- National Center for O*NET Development. (2014). *Production Database—O*NET® 18.1*. Available online at: <http://www.onetcenter.org/database.html>
- National Science Foundation. (2013). Women, minorities, and persons with disabilities in science and engineering. Available online at: http://www.nsf.gov/statistics/wmpd/2013/pdf/nsf13304_digest.pdf
- *Parker, R. M. (2002). *Occupational Aptitude Survey and Interest Schedule—3rd Edn: Interest Schedule Examiner’s Manual*. Austin, TX: Pro-Ed.
- Pervin, L. A. (1968). Performance and satisfaction as a function of individual-environment fit. *Psychol. Bull.* 69, 56–68. doi: 10.1037/h0025271
- Prediger, D. J. (1982). Dimensions underlying Holland’s hexagon: missing link between interests and occupations? *J. Vocat. Behav.* 21, 259–287. doi: 10.1016/0001-8791(82)90036-7
- *Psychological Corporation. (1991). *Career Interest Inventory: Counselor’s Manual*. San Antonio, TX: Psychological Corporation.
- *Ripley, R. E., Neidert, G. P. M., and Ortman, N. L. (2001). *World of Work Inventory Interpretation Manual*. Tempe, AZ: World of Work.
- Robertson, K. F., Smeets, S., Lubinski, D., and Benbow, C. P. (2010). Beyond the threshold hypothesis: even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Curr. Dir. Psychol. Sci.* 19, 346–351. doi: 10.1177/0963721410391442
- Rounds, J., and Su, R. (2014). The nature and power of interests. *Curr. Dir. Psychol. Sci.* 23, 98–103. doi: 10.1177/0963721414522812
- Rounds, J., Smith, T., Hubert, L., Lewis, P., and Rivkin, D. (1999). *Development of Occupational Interest Profiles (OIPs) for the O*NET part 1: Report*. Raleigh, NC: National Center for O*NET Development. Available online at: <http://www.onetcenter.org/reports/OIP.html>
- Savickas, M. L., and Spokane, A. R. (eds.). (1999). *Vocational Interests: Meaning, Measurement, and Counseling Use*. Palo Alto, CA: Davies-Black.
- Schneider, B. (1987). The people make the place. *Pers. Psychol.* 40, 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x
- Seat, E., Parsons, J. R., and Poppen, W. A. (2001). Enabling engineering performance skills: a program to teach communication, leadership, and teamwork. *J. Eng. Educ.* 90, 7–12. doi: 10.1002/j.2168-9830.2001.tb00561.x
- Silvia, P. J. (2006). *Exploring the Psychology of Interest*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195158557.001.0001
- Strong, E. K. Jr. (1943). *The Vocational Interests of Men and Women*. Stanford, CA: Stanford University Press.
- Su, R., Murdock, C. D., and Rounds, J. (2014). “Person-environment fit,” in *APA Handbook of Career Intervention*, eds P. Hartung, M. Savickas, and B. Walsh (Washington, DC: American Psychological Association), 81–98.
- Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364
- Tay, L., Su, R., and Rounds, J. (2011). People-things and data-ideas: Bipolar dimensions? *J. Couns. Psychol.* 58, 424–440. doi: 10.1037/a0023488
- Thorndike, E. L. (1911). *Individuality*. Boston, MA: Houghton, Mifflin.
- Trice, A. D., and Rush, K. (1995). Sex-stereotyping in four-year-olds’ occupational aspirations. *Percept. Mot. Skills* 81, 701–702. doi: 10.2466/pms.1995.81.2.701
- U.S. Bureau of Labor Statistics. (2010). *Standard Occupational Classification*. Available online at: <http://www.bls.gov/soc/#classification>
- U.S. Bureau of Labor Statistics. (2014). *Women in the Labor Force: A Databook*. Available online at: <http://www.bls.gov/cps/wlf-databook-2013.pdf>
- U.S. Department of Labor. (2007). *The STEM Workforce Challenge: The Role of the Public Workforce System in a National Solution for a Competitive Science, Technology, Engineering, and Mathematics (STEM) Workforce*. Available online at: http://www.doleta.gov/Youth_services/pdf/STEM_Report_4%2007.pdf
- U.S. Immigration and Customs Enforcement. (2012). *STEM-Designated Degree Programs List*. Available online at: <http://www.ice.gov/sites/default/files/documents/Document/2014/stem-list.pdf>
- Valla, J. M., and Ceci, S. J. (2014). Breadth-based models of women’s underrepresentation in STEM fields: an integrative commentary on Schmidt (2011) and Nye et al. (2012). *Perspect. Psychol. Sci.* 9, 219–224. doi: 10.1177/1745691614522067
- Viechtbauer, W. (2008). “Analysis of moderator effects in meta-analysis,” in *Best Practices in Quantitative Methods*, ed J. W. Osborne (Thousand Oaks, CA: Sage Publications), 471–487. doi: 10.4135/9781412995627.d37
- *Vocational Research Institute. (1988). *Manual for the Vocational Research Interest Inventory*. Philadelphia, PA: Vocational Research Institute.
- Wai, J., Lubinski, D., Benbow, C. P., and Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: a 25-year longitudinal study. *J. Educ. Psychol.* 102, 860–871. doi: 10.1037/a0019454
- Wang, M. T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in STEM career choice. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937
- Wilson, D. B. (2005). *Meta-Analysis Macros for SAS, SPSS, and Stata*. Available online at: <http://mason.gmu.edu/~dwilsonb/ma.html>

- *Winefordner, D. (1983). *OVIS II: Ohio Vocational Interest Survey Manual for Interpreting*, 2nd Edn. Cleveland, OH: Psychology Corporation.
- Woodcock, A., Graziano, W. G., Branch, S. E., Habashi, M. M., Ngambeki, I., and Evangelou, D. (2013). Person and thing orientations: psychological correlates and predictive utility. *Soc. Psychol. Pers. Sci.* 4, 116–123. doi: 10.1177/1948550612444320
- *Zytowski, D. G. (2007). *Kuder Career Search with Person Match: Technical Manual (Version 1.1)*. Available online at: <http://www.kuder.com/downloads/kcs-tech-manual.pdf>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 November 2014; paper pending published: 05 January 2015; accepted: 05 February 2015; published online: 25 February 2015.

Citation: Su R and Rounds J (2015) All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Front. Psychol.* 6:189. doi: 10.3389/fpsyg.2015.00189

This article was submitted to *Developmental Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Su and Rounds. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

*References marked with an asterisk indicate studies included in the meta-analysis.



Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in STEM careers

Ming-Te Wang*, Jessica Degol and Feifei Ye

University of Pittsburgh, Pittsburgh, PA, USA

Edited by:

Stephen J. Ceci, Cornell University, USA

Reviewed by:

Jonathan Wai, Duke University, USA
Joni M. Lakin, Auburn University, USA

*Correspondence:

Ming-Te Wang, University of Pittsburgh, 230 South Bouquet Street, Pittsburgh, PA 15260, USA
e-mail: mtwang@pitt.edu

Although young women now obtain higher course grades in math than boys and are just as likely to be enrolled in advanced math courses in high school, females continue to be underrepresented in some Science, Technology, Engineering, and Mathematics (STEM) occupations. This study drew on expectancy-value theory to assess (1) which intellectual and motivational factors in high school predict gender differences in career choices and (2) whether students' motivational beliefs mediated the pathway of gender on STEM career via math achievement by using a national longitudinal sample in the United States. We found that math achievement in 12th grade mediated the association between gender and attainment of a STEM career by the early to mid-thirties. However, math achievement was not the only factor distinguishing gender differences in STEM occupations. Even though math achievement explained career differences between men and women, math task value partially explained the gender differences in STEM career attainment that were attributed to math achievement. The identification of potential factors of women's underrepresentation in STEM will enhance our ability to design intervention programs that are optimally tailored to female needs to impact STEM achievement and occupational choices.

Keywords: gender gap, STEM, math achievement, career choice, motivation

INTRODUCTION

Although girls now obtain higher course grades in math than boys and are just as likely to be enrolled in advanced math courses in high school, females continue to be underrepresented in some Science, Technology, Engineering, and Mathematics (STEM) occupations (National Science Foundation, 2011). For example, in 2010 among employed individuals whose highest degree was a Bachelor's, females comprised around 42% of the workforce in mathematics, 11% of the workforce in engineering, 23% of the workforce in computer and information sciences, and 34% of the workforce in physical sciences (National Science Foundation, 2014).

Career aspirations based on individual competencies, values, and perceived compatibility of competencies and values, are formulated in adolescence and shape the academic pathways that lead to the STEM pipeline (Tai et al., 2006). It is very difficult to initiate a STEM trajectory after beginning college, due to the very constrained and prescribed curricula in STEM fields (Tyson, 2011). Therefore, in order to prevent many talented and capable young women from opting out of the STEM pipeline, it is important to identify the intellectual and psychological factors that surface in the elementary and secondary school years and predict later career choice (Maltese and Tai, 2011; Ceci et al., 2014). In turn, our ability to design intervention programs to impact STEM achievement and occupational choices through these factors will be more optimally tailored to females.

Despite many researchers dedicating themselves to studying the gender gap in STEM fields, the extant literature is limited in

several ways. Current reform efforts primarily focus on improving students' exposure to and performance in advanced-level math courses in high school as a way to address the gender gap in STEM. While encouraging math achievement and enrollment in advanced courses is an important step in setting the foundation for the successful attainment of STEM careers, it alone does not account for the complex motivational factors that influence STEM career choice (Eccles, 2009). In fact, neither mathematical aptitude, nor advanced math course enrollment are strongly predictive of student enjoyment in math-related activities or career choice (Wang and Degol, 2013). Instead, students' motivational beliefs (e.g., competence beliefs, attitudes, values, interest) about math learning are more critical determinants of future educational and career choices (Maltese and Tai, 2010, 2011). While the importance of motivational beliefs has been widely recognized, most studies are limited to STEM performance or college major as the outcome and very few longitudinal studies have addressed the underlying factors in the high school years that motivate girls to pursue actual STEM careers in adulthood (Lubinski and Benbow, 2006).

Although ability self-concept (feeling competent to succeed) has been shown to be an important predictor of academic performance (Guiso et al., 2008; Hill et al., 2010), personal interest and perceived task value play highly important roles in shaping individual achievement and career choices, and can be more influential than academic self-concept (Eccles, 2009). For example, studies show that in early adolescence, girls and boys tend to endorse different work preferences and lifestyle values

(Ferriman et al., 2009). These personal interests and task values can rest outside of students' perceptions of their own intellectual abilities, and may contribute to the gender gaps in STEM performance and career choices. However, it is unclear whether students' motivational beliefs (subjective task values in particular) mediate the relation between gender and STEM career through math achievement.

In this study, we draw on Eccles' (2009) expectancy-value theory to assess which intellectual competencies and motivational beliefs move individuals toward or away from STEM careers. Expectancy-value theory posits that achievement-related choices, such as occupation selection, are most directly influenced by intellectual competencies, ability self-concepts, and the subjective task value attached to the various options. Subjective task value is comprised of interest value (liking or enjoyment), utility value (the instrumental value of the task for helping to fulfill personal goals), attainment value (the link between the task and one's sense of self, identity, and core personal values), and cost (what may be given up by making a specific choice). Career choices are ultimately made after a number of options, and their various components (e.g., money, authority, social connection) are evaluated and identified as either fitting personal goals or not. Gender differences in career choices reflect gendered differences in relative intellectual competencies, ability self-concepts, and the relative subjective task value of each option under consideration.

INTELLECTUAL COMPETENCIES

There are small average gender differences between boys and girls on some indicators of intellectual competencies: girls outperform boys in some tests on verbal skills (Park et al., 2008); and girls earn slightly higher grades in all school subjects, including high school math and science (Hyde et al., 2008). Furthermore, differences in the proportion of males and females scoring in the extreme right tail of high stakes math and reading standardized tests have been consistently detected. Males outnumber females in the top 0.01% of the distribution in the SAT and ACT math subtests by 4:1 and 3:1, respectively, while females have a slight advantage on the verbal subtests (Wai et al., 2012). These findings lead to the conclusion that intellectual aptitude, at least by itself, is not the dominant factor in the underrepresentation of women in STEM fields (Ceci and Williams, 2010).

ABILITY SELF-CONCEPTS

Expectations for success, confidence in one's abilities to succeed, and personal efficacy have emerged as important predictors of academic achievement and activity involvement (Wigfield et al., 2006). Both boys and girls who rate their math competence highly are more likely to enroll in advanced math courses and receive higher grades in math (Pajares, 2005). Additionally, high school girls tend to rate their math competence lower than boys with similar math grades (Correll, 2001); a finding of particular interest given that poor math self-concept or perceived competence may play a role in female underperformance in mathematics (Durik et al., 2006). Yet intellectual competencies or competence beliefs are a necessary—but not sufficient—predictor of career choices (Joyce and Farenga, 2000). As suggested by expectancy-value theory, career choices depend not only on confidence in one's abilities

to succeed, but also on subjective task values—the value one attaches to relevant subject domains and the goals associated with these domains.

SUBJECTIVE TASK VALUES

Research on subjective task values shows a number of potentially interrelating effects and gender variations. For instance, despite similarities in math performance, girls' 'liking' of math decreases on average as they move through adolescence to a greater extent than boys' (Koller et al., 2001). Girls also are more likely to express greater interest in English than math when compared to boys (Jacobs et al., 2002). These findings, in combination with research showing that even females with high math-aptitude tend to express less interest in math-intensive careers (Lubinski and Benbow, 2006), suggest that differential interest and task value in math may contribute to the underrepresentation of women in STEM fields.

Gender differences in occupational and lifestyle values (forms of utility and attainment values) are also potentially important contributing factors to women's underrepresentation in STEM fields (Lubinski et al., 2001; Ferriman et al., 2009). For example, females are typically more interested in socially oriented careers, while males are more interested in working with objects (Su et al., 2009; Diekmann et al., 2011). Meanwhile, women are more likely to value the development of altruistic, reciprocal relationships more than men (Schwartz and Rubel, 2005). This phenomena is illustrated by the fact that women tend to put more value on jobs that allow them to help others and make meaningful contributions to society (communion/affiliative orientation; Abele and Spurk, 2011) and math-intensive careers are usually viewed as being object-oriented (Webb et al., 2002) and less social (Hill et al., 2010).

Finally, research on how priorities beyond career fulfillment help shape females' decisions to refrain from entering STEM fields, indicate that life values and 'sense of fit' are important factors. Per Hakim (2006), women tend to prefer more home-centered lifestyles, whereas men tend to prefer more work-committed lifestyles, and math-related careers are not perceived by females as accommodating to their desired work-family balance. Because work-family balance is highly relevant to career-aged women, most studies have been conducted with adult females; however, this gap in the literature makes it unclear whether family work balance is an important predictor of career choices for high school students.

The current study investigates (1) which intellectual and motivational factors in high school predict gender differences in career choices and (2) whether students' motivational beliefs mediated the pathway of gender on STEM career through math achievement. Two sets of analyses were conducted to this end. In the first set of analyses, we used hierarchical logistic regression to test whether math ability self-concept and subjective task values (i.e., math interest, social and family values, and desired job characteristics) at 12th grade predicted gender differences in the selection of STEM vs. non-STEM careers, while holding math and reading ability, and family socioeconomic status constant. In the second set of analyses, we tested the role of subjective task values and

math achievement as potential mediators for predicting gender differences in selecting STEM and non-STEM careers.

MATERIALS AND METHODS

PARTICIPANTS

We used data from the Longitudinal Study of American Youth, a large-scale national study initiated in 1987 that followed two cohorts of students through middle school, high school, and at various stages beyond high school, focusing predominantly on student, family, and school characteristics that influence student achievement, interest, and occupational proclivities toward math and science. The base-year sample consisted of 3,116 students in the 7th grade (mean age = 12 years, cohort II) and 2,829 10th graders (mean age = 15 years, cohort I) from 50 public school systems across the country. Schools were classified as urban (25%), suburban (42%), and rural (33%). Selected schools are considered representative of secondary schools across the country. Each year participants were given standardized tests of math achievement in addition to completing questionnaires about their experiences and attitudes on STEM-related learning. Reading achievement was also assessed for both cohorts in 12th grade. In 2007, when the original study participants were between 33 and 37 years of age, a sample of 3,689 original participants (76% response rate) completed the telephone interview surveys, updating their educational and occupational history from post high school into their mid-thirties. Data used in this study were mainly from two waves: 12th grade and the 2007 follow-up, 14 or 17 years postsecondary school, depending on the cohort. At 12th grade, 75% were White, and 49% were female adolescents.

To determine whether the students who participated in 12th grade differed from those who dropped out between the ages of 33–37, a series of independent samples contingency table analyses and *t*-tests were conducted with all independent, outcome, and demographic variables at 12th grade. Results revealed that those who dropped out of the study did not differ from those who participated in the study at 12th grade. We used full information maximum likelihood (FIML) estimation in Mplus 7.3 to account for missing data in all analysis, as FIML was recommended as the most appropriate approach to handle missing data when data are missing at least at random (Allison, 2012).

MEASURES

STEM occupation

Participants' occupations at ages 33–37 were self-reported in a telephone interview conducted in 2007. We operationalized occupations into two categories: (1) non-STEM, consisting of careers in the fine arts, literature, business, education, and social sciences, and (2) STEM jobs, consisting of occupations in mathematics, engineering, computer science, life science, medical science, and physical science.

Math and reading achievement

Standardized math scores were used from tests taken by students in the spring of 12th grade. The test was developed by the (National Assessment of Educational Progress [NAEP], 1986) to measure students' knowledge of math, the application and utilization of math knowledge, and integration of math knowledge. A

standardized test of reading developed by the Educational Testing Service was used to measure students' reading comprehension in the spring of 12th grade. Multiple-group item-response theory (IRT) methods were used to scale scores on a metric with a mean of 50 and a SD of 10 (Miller and Kimmel, 2010).

Math ability self-concept

Students completed a survey in the fall of 12th grade indicating their math ability self-concept. The math ability self-concept scale (Bleeker and Jacobs, 2004) included three items that measured students' perceived abilities and expectancy for success in math (e.g., "I am good at math," "I usually understand math"). The academic self-concept scale was rated from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores reflecting higher math ability self-concept ($\alpha = 0.80$).

Subjective task values

In the fall of 12th grade, we measured students' interest values, utility values, and attainment values (Eccles et al., 1997):

Math task value. The math task value scale included five items that measured students' interest, enjoyment, and the value they attach to math (e.g., "I enjoy math," "Math is useful in everyday problems"). The math task value scale was rated from 1 (*strongly disagree*) to 5 (*strongly agree*) with higher scores reflecting higher task values in math ($\alpha = 0.75$).

Altruism, family values, and monetary values. Students rated the relative importance (1 = *not important*; 2 = *somewhat important*; 3 = *very important*) of a variety of future economic, social, and familial goals. Three separate constructs were generated indicating the extent to which youth exemplified altruistic values, family values, and monetary values. A total of four items were used to indicate the importance students attributed to having an active role in helping others in their communities, including changing social/economic wrongs, staying current on social issues, and helping others within the community ($\alpha = 0.77$). Family values were constructed using two items that reflected the importance students attributed to having children and prioritizing their family life in the future ($\alpha = 0.69$). Finally, monetary values measured the extent to which students valued making lots of money in the future. Higher scores indicate placing greater importance on altruism, family values, or monetary values.

Work preferences. Students completed a survey indicating qualities of a future career they would find preferable. Students checked a box to indicate whether they preferred a job with the characteristics listed (1 = *yes*; 0 = *no*). In order to examine the extent to which youth preferred to work with people or objects, two items were examined for the absence or presence of a checkmark: prefers a job that allows work with other people in teams, and prefers a job that allows work with numbers and formulas. Working with teams, therefore, represents a people-oriented job focus and working with numbers and formulas represents an object-oriented job focus.

Covariates

We controlled for several potential confounds related to individual career choices in STEM fields, including child gender (0 = *female*;

1 = male), child race/ethnicity (0 = White; 1 = others), parent education (0 = some college/HS or less; 1 = BA/BS or higher), and parent STEM employment (0 = parents do not work in STEM or technical field; 1 = at least one parent is employed in a STEM or technical profession). Parental education and employment were collected from parent reports.

RESULTS

We compared males and females on career choice, covariates, and independent variables. Chi-square tests were used for dichotomous variables, and independent sample *t*-tests for continuous variables (see **Table 1**). More males chose STEM careers than females, and preferred to work with numbers. Moreover, males had higher math achievement, math ability self-concept, math task value, and a greater preference for high-paying careers. In contrast, females had higher reading achievement, altruism, and family values.

We conducted hierarchical logistic regression to examine which intellectual and motivational factors were predictive of STEM careers and contributed to gender disparities in selection of STEM occupations, controlling for child gender and race, parent education, and parent STEM employment (see **Table 2**). All continuous predictors, including math achievement, reading achievement, altruism, family values, math ability self-concept, and math task value, were standardized to have mean of zero and SD of one. All dichotomous predictors were indicator coded. Variance Inflation Factor (VIF) was calculated for all predictors and there was no concern with multicollinearity (VIFs < 2.3).

In the first set of hierarchical logistic regression models, we included gender as the only predictor to show the gender disparity in STEM occupation. Males were 1.38 times ($p = 0.005$) as likely as females to choose STEM careers. Second, we added student race/ethnicity, parent education,

and parent STEM occupation, which all significantly predicted STEM occupation ($ps < 0.02$). The gender effect was still significant, but the odds ratio decreased to 1.36 ($p = 0.009$). Third, we added math and reading achievement, and found that only math achievement was significantly related to STEM occupation. Importantly, the gender effect was reduced to non-significance. Fourth, we added math ability self-concept and math task value, in which only math task value was positively associated with STEM occupation. Fifth, we added altruism, family values, and monetary importance, with only altruism negatively predicting STEM occupation. Finally, we added student's work preferences (e.g., either working with people or objects), both of which failed to significantly differentiate career choices.

In order to test the mediation effect of math task value and math achievement on gender differences in career choice, we adopted the outlined procedure in Baron and Kenny (1986). We first assessed the total direct effect of gender on STEM occupation with a logistic regression model while partialling out the effects of such covariates as race, parent education, parent occupation, and reading achievement from STEM occupation. Then we conducted two path models to tease out the mediation effects of math achievement, math task values, and altruism on gender difference in STEM occupation, while partialling out the effects of the covariates from all mediators and STEM occupation. In the first mediation model (**Figure 1A**), we tested only the mediation effect of math achievement. In the second mediation model (**Figure 1B**), we added math task values and altruism as additional mediators given that the hierarchical logistic regression results suggested that altruism and math task values were the only significant motivational predictors of STEM occupation. A direct relationship was modeled from math task value and altruism to math achievement. Then the indirect effects were tested

Table 1 | Descriptive statistics of the sample and tests of the difference between female and male ($N = 5,945$).

	Female	Male	<i>p</i> -value
Dependent variable			
Science, Technology, Engineering, and Mathematics (STEM) career (0 = no; 1 = yes; %)	8	10	0.005
Covariates			
Child race (0 = white/other; 1 = black/hispanic; %)	21	21	n.s.
Parent education (0 = some college/HS or less; 1 = BA/BS or higher; %)	29	31	n.s.
Parent STEM occupation (0 = non-STEM; 1 = at least one parent in STEM occupation; %)	17	16	n.s.
Math achievement score	67.24 (12.60)	69.48 (14.60)	<0.001
Reading achievement score	55.34 (27.30)	51.44 (29.60)	<0.001
Math ability self-concept	9.88 (2.85)	10.47 (2.34)	<0.001
Math task value	17.61 (3.80)	18.01 (3.56)	0.002
Altruism	8.13 (1.98)	7.96 (2.03)	0.013
Family values	5.25 (0.97)	4.98 (1.07)	<0.001
Monetary importance	2.23 (0.59)	2.41 (0.61)	<0.001
Prefer job: work with others (0 = no; 1 = yes; %)	53	51	n.s.
Prefer job: work with numbers (0 = no; 1 = yes; %)	14	20	<0.001

Independent sample *t*-test was used for continuous variables and Chi-square tests were used for binary variables. SD are in parentheses.

Table 2 | Hierarchical logistic regression to predict the choice of a STEM occupation.

Predictor	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Odds ratio
Gender (1 = male)	0.32 (0.12)**	0.31 (0.12)**	0.16 (0.12)	0.12 (0.13)	0.14 (0.13)	0.12 (0.13)	1.12
Child race (1 = black/hispanic)		-0.69 (0.21)**	-0.22 (0.22)	-0.35 (0.22)	-0.29 (0.23)	-0.31 (0.23)	0.74
Parent education		0.78 (0.12)**	0.44 (0.13)**	0.53 (0.13)**	0.55 (0.13)**	0.56 (0.13)**	1.75
Parent STEM occupation		0.33 (0.14)*	0.18 (0.14)	0.17 (0.15)	0.18 (0.15)	0.15 (0.15)	1.17
Math achievement score			0.79 (0.12)**	0.64 (0.12)**	0.66 (0.12)**	0.65 (0.13)**	1.91
Reading achievement score			0.02 (0.09)	0.01 (0.09)	0.02 (0.09)	0.02 (0.10)	1.02
Motivational beliefs and values							
Math ability self-concept				0.11 (0.10)	0.08 (0.10)	0.07 (0.11)	1.07
Math task value				0.40 (0.09)**	0.44 (0.09)**	0.42 (0.09)**	1.53
Altruism					-0.20 (0.08)*	-0.20 (0.08)*	0.82
Family values					0.09 (0.08)	0.08 (0.08)	1.08
Monetary importance					-0.08 (0.07)	-0.08 (0.07)	0.92
Prefer job: work with others						0.28 (0.15)	1.32
Prefer job: work with numbers						0.26 (0.17)	1.29

All continuous predictors were standardized, including math achievement score, reading achievement score, math ability self-concept, math task value, altruism, family values, and monetary importance. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

using bias-corrected and accelerated (BCa) bootstrap confidence interval (BCI).

The total direct effect of gender on STEM occupation was significant, $B = 0.36$, $p = 0.002$, odds ratio = 1.43. With math achievement as the only mediator in the model (Figure 1A), the direct effect from gender to STEM occupation became not significant, $B = 0.15$, $p = 0.21$, odds ratio = 1.17. When math task value and altruism were added as mediators in addition to math achievement (Figure 1B), the direct effect was further reduced to $B = 0.10$, $p = 0.41$, odds ratio = 1.11. The relative indirect effect (Huang et al., 2004), loosely interpreted as the proportion of the total effect that is mediated, was calculated to be $1 - 0.15/0.36 = 0.58$ with math achievement as the mediator, and $1 - 0.10/0.36 = 0.72$ with math achievement, math task value, and altruism as the mediators.

Table 3 presents indirect effects (unstandardized path coefficients) and their BCa BCI. In the first path model, math achievement significantly mediated the gender difference in STEM occupation, with the indirect effect estimated to be 0.17 (95% BCI: 0.11–0.24), indicating that, indirectly via math achievement, the odds of males choosing STEM occupations increased by 1.19 times that of females. In the second path model, males had higher math achievement and math task value, but lower altruism. Math task value was significantly associated with math achievement, while altruism was not. Both math task value and altruism were directly and significantly associated with STEM occupation, positively for math task value, and negatively for altruism. We found four significant indirect paths, including: (1) Gender → Math Task Value → STEM occupation, (2) Gender → Altruism → STEM occupation, (3) Gender → Math Task Value → Math Achievement → STEM occupation, and (4) Gender → Math Achievement → STEM occupation. The relative indirect effect indexes were 0.17, 0.06, 0.06, and 0.36 for these four significant indirect effects. It is noteworthy that math task value not only mediated the gender difference in STEM occupation, but its positive relationship with math achievement also accounted partially for the mediating effect of math achievement on the gender effect. By including math task value and altruism, the indirect effect of Gender → Math Achievement → STEM Occupation was reduced from 0.17 to 0.13, with the corresponding relative indirect effect index reduced from 0.58 to 0.36. However, the magnitude of the Gender → Math Task Value → Math Achievement → STEM occupation effect has a small value of 0.02. This is not surprising given that the effects of race, parent education, parent occupation, and reading achievement were controlled for with all mediators. Results showed that math achievement was significantly related with gender ($B = 0.19$, $p < 0.001$), race ($B = -0.43$, $p < 0.001$), parent education ($B = 0.23$, $p < 0.001$), parent occupation ($B = 0.13$, $p = 0.001$), and reading achievement ($B = 0.54$, $p < 0.001$). Math task value was significantly related with gender ($B = 0.13$, $p < 0.001$), race ($B = 0.25$, $p < 0.001$), and reading achievement ($B = 0.15$, $p < 0.001$).

DISCUSSION

Increasing opportunities for female participation in STEM fields is a pivotal social, economic, and political issue in the advancement

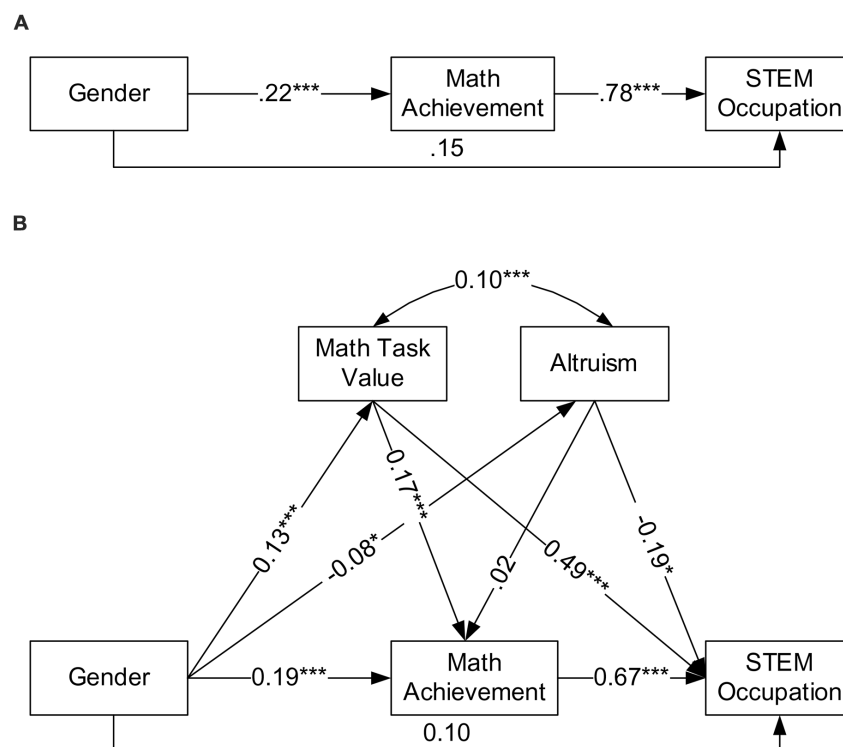


FIGURE 1 | Mediation effects of math task value, altruism, and math achievement on gender difference in STEM occupation.

(A) Presents results with math achievement as the only mediator and **(B)** presents results with math task value and altruism added

as mediators in addition to math achievement. All coefficients are unstandardized, adjusted with covariates of race, parent education, parent occupation, and reading achievement. * $p < 0.05$; *** $p < 0.001$.

Table 3 | Mediation effect of math achievement, task value, and altruism on gender difference in STEM occupation.

	Mediation model 1		Mediation model 2	
	Estimate	Bootstrap 95% CI	Estimate	Bootstrap 95% CI
Indirect path				
Gender → math task value → STEM occupation			0.06	(0.03, 0.11)
Gender → altruism → STEM occupation			0.02	(0.002, 0.04)
Gender → math task value → math achievement → STEM occupation			0.02	(0.006, 0.03)
Gender → altruism → math achievement → STEM occupation			0.001	(−0.005, 0.000)
Gender → math achievement → STEM occupation	0.17	(0.11, 0.24)	0.13	(0.08, 0.19)

of female interests. In order to elucidate the factors associated with females' underrepresentation in STEM, the current study examined which factors predicted gender differences in the selection of STEM occupations, and whether math task values and altruism mediated the pathway of the gender effect on STEM career choice through math achievement. Identifying potential barriers that keep women from fulfilling their potential in STEM fields will help inform intervention efforts targeting the removal of these barriers.

We found that math achievement in 12th grade mediated the association between gender and attainment of a STEM career by the early to mid-thirties. Women were less likely than men

to pursue a career in STEM, but this relation was explained by gender differences in math achievement in high school. Our results show that women, on average, had lower math standardized scores than men, and unsurprisingly, individuals with higher math achievement were more likely to attain a career in STEM. However, math achievement was not the only factor distinguishing gender differences in STEM occupations. Math task value partially mediated the pathways among gender, math achievement, and STEM careers. Women had lower math task values than men, and lower math task value was associated with lower math achievement and lower likelihood of pursuing STEM careers. Essentially, despite math achievement explaining career

differences between men and women, math task value also contributed to the gender differences in STEM career attainment that were attributed to math achievement. These findings shed some light onto the complex ways that ability self-concept and subjective task values operate in promoting STEM career selection. Expectancy-value theory posits that individuals consider multiple factors when selecting potential careers, including prior achievement, perceived competencies, and task values. In line with this theory, math task value and altruism (a form of utility value) predicted STEM career, but math ability self-concept did not. Previous research has found that ability self-concept is predictive of academic achievement, but, unlike subjective task values, it is not consistently linked to educational or career choices (Durik et al., 2006). In other words, believing that you are good at a task may further enhance your performance in the task, but it does not mean that you enjoy the task and will continue to pursue it.

How do these findings relate to factors associated with females' underrepresentation in STEM fields? Increasing math achievement is important for increasing women's representation in STEM, but achievement alone may not be sufficient. We know that achievement matters for STEM enrollment; many of the most mathematically talented individuals eventually achieve prestigious careers in STEM fields (Wai et al., 2010). Historically, women's underperformance in quantitative reasoning skills relative to men's has been considered one of the main factors in women's decisions to opt out of STEM fields (Halpern, 2007). In response, public focus and political initiatives have centered on increasing female math performance and advanced math course enrollment. However, converging evidence from the current study and other research has demonstrated that increasing quantitative skills alone will not effectively lead to greater female participation in STEM (Ceci and Williams, 2011; Maltese and Tai, 2011; Riegle-Crumb et al., 2012). While gender differences in attainment of STEM careers was explained by lower female performance on standardized math tests in our first model, the second model demonstrated that this pathway (gender to achievement to STEM career) was partially attributed to gender differences in math task values.

Girls consistently express less interest in math (Jacobs et al., 2002) and view math and STEM careers as less aligned with their personal career interests and goals (Su et al., 2009). Studies have shown that greater interest and greater perceived importance and utility value of math may lead to greater investment in and persistence in math activities, which ultimately lead to higher math achievement (Wigfield and Eccles, 2002; Wang, 2012). Therefore, aside from promoting greater math achievement, current policy initiatives also need to target the development of math task values: encouraging interest in math and its utility value. When women see STEM fields as useful, widely applicable, and viable career options they will be more likely to opt into them.

Given that math interest and task values are linked to academic performance, the benefits derived from enhancing math task value may be twofold. If math task values impact math achievement and selection of STEM occupations, then intervening to promote subjective task values in math should not only increase STEM persistence in the long run, but also enhance math achievement. Since math achievement positively predicted long term

decisions in STEM career, the developmental impacts of interventions that seek to increase math task value could be exponential compared to programs that target math skills alone. Furthermore, targeting math task values may not only lead to increases in math achievement, but improved math performance may actually further enhance math task value, given that the two are reciprocally linked over time. Exclusively focusing on math achievement as a path to STEM persistence is a unimodal answer, while increasing math task values in addition to achievement is a multimodal solution that could activate multiple pathways to a STEM career.

Furthermore, since our study shows that math task values begin to predict students' STEM attainment as early as high school, early intervention is vital. Recent studies have shown that interest and career aspirations in STEM emerge prior to entry into high school, and that by 12th grade the decision to major in a STEM vs. non-STEM career is largely solidified for many students (Maltese and Tai, 2011). While there are an increasing number of programs that target student interest, enjoyment, and engagement in STEM (e.g., Detroit Area Pre-College Engineering Program, Great Explorations in Math and Science, Project Lead the Way), these crucial motivating factors should become a greater focus of all K-12 interventions. Particularly, given that increases in STEM course taking and achievement among females have not led to comparable increases in STEM workforce participation, programs need to strengthen teacher training and redesign curriculum to include targeted strategies for dispelling gender stereotypes and increasing female interest in STEM.

Our findings suggest that enhancing women's math task value may be instrumental in inspiring larger numbers of women to seriously consider STEM fields as viable career options. But how would this look in practice? We know that students are more engaged in classrooms that incorporate hands-on learning, creative thinking, and challenging real-world applications of problems and concepts (Marks, 2000). For girls and women in particular, it may be helpful to take a proactive approach that utilizes their unique strengths. For instance, a recent study showed that girls are more likely than boys to have both high verbal and math skills (Wang et al., 2013). Therefore, incorporating storytelling into math may not only capitalize on the strengths of girls' verbal skills but also increase female interest in math and science by making these subjects appear hands-on and practical. Additionally, specific teaching strategies such as focusing on women's historical contributions to these fields, and increasing girls' exposure and access to female scientists and engineers as career role models (Steinke et al., 2007), may help combat the pervasive math-gender stereotypes that affect girls' math identities as young as 6 years of age (Cvencek et al., 2011).

Altruism can also be emphasized. Women view helping people and contributing to the greater good as highly important career goals (Su et al., 2009; Abele and Spurk, 2011), which are not perceived to be in line with STEM careers. Indeed, our study suggests that altruism mediated the gender effect on STEM occupations. Since it is plainly not the goal to make women less altruistic, STEM educators should place greater emphasis on demonstrating how female scientists can develop technologies and make discoveries

that greatly benefit people's lives. This is in line with the National Academy of Engineering [NAE]'s (2013) recent efforts to alter public perceptions by communicating that engineering is a helping profession that works on solving problems of human health and safety throughout the world.

Interestingly, differences in family and monetary values, and preferences for working with people or numbers did not explain gender differences in STEM careers. The lack of findings for family values was not unexpected; previous research has demonstrated that gender differences in work/family balance preferences do not typically emerge until the mid-30s or adulthood, when women are more likely to be raising children and building their families (Ferriman et al., 2009). Since family values were assessed in 12th grade, we can expect that work/family lifestyle preferences will not yet factor prominently in determining male/female differences in STEM choice. Similarly, monetary values may not have predicted gender differences in career choices at this age, because much like family values, concern over earned income is a distal issue that will be experienced more fully in adulthood. More immediate concerns over money, such as tuition costs and student loan debt, may have greater bearing in adolescence. Preferences for working with people or numbers, which typically differ along gender lines, also failed to explain gender differences in STEM careers. Most occupations allow for the opportunity to work with people, numbers, and objects to varying degrees. The key difference is in how prominently these aspects are featured in a career (e.g., interacting directly with people, such as teaching vs. interacting directly with numbers, such as engineering). However, enjoyment of working with numbers does not necessarily indicate a lack of enjoyment in working with people and vice versa, and both may be large components of the same career (e.g., teaching engineering students). For this reason, it is likely that these preferences may not explain gender differences in STEM career selection to the same degree as math task value and altruism.

CONCLUSION

Despite women's advances in the U.S. workforce, their entrance into lucrative STEM careers has been less successful, and these professions continue to be heavily male-dominated. The prestige and innovation surrounding math and science, along with their accompanying economic benefits, are not extended to women when they are non-participatory in these fields. Our study builds on well-established literature by identifying the intellectual and motivational factors contributing to women's underrepresentation in STEM. However, it is important to reiterate that generating greater female interest in STEM should not be equated with forcing unwanted career choices on them. We do not want to coerce women into STEM fields if they have no interest in them, and we do not want to undermine the importance and value of non-STEM careers. Instead, we seek to alter instructional approaches to math and science education to demonstrate how STEM careers can benefit society and provide opportunities for helping and interacting with others, thereby, merging women's personal task values and career aspirations. Furthermore, many adolescents may not truly understand what it means to obtain a degree in STEM (Fralick et al., 2009). Introducing youth to the

different majors they can pursue in STEM and the careers that these degrees will prepare them for can provide adolescents with a better understanding of the nature of these occupations. Ensuring that females are well informed of the full diversity of options available in STEM will enable math-competent females to better evaluate the utility and cost of different STEM careers. Our main goals are to present all of the STEM career opportunities available to women, remove misconceptions that operate as barriers to STEM enrollment, and empower women to make career decisions that best meet their needs for personal and occupational fulfillment.

REFERENCES

- Abele, A. E., and Spurk, D. (2011). The dual impact of gender and the influence of timing of parenthood on men's and women's careers development: longitudinal findings. *Int. J. Behav. Dev.* 35, 225–232. doi: 10.1177/0165025411398181
- Allison, P. (2012). Handling missing data by maximum likelihood. *Paper 312–2012 presented at SAS Global Forum 2012*, Orlando, FL.
- Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173
- Bleeker, M., and Jacobs, J. E. (2004). Achievement in math and science: do mothers' beliefs matter 12 years later? *J. Educ. Psychol.* 96, 97–109. doi: 10.1037/0022-0663.96.1.97
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., and Williams, W. M. (2010). Sex differences in math-intensive fields. *Curr. Dir. Psychol. Sci.* 19, 275–279. doi: 10.1177/0963721410383241
- Ceci, S. J., and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3157–3162. doi: 10.1073/pnas.1014871108
- Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. Soc.* 106, 1691–1730. doi: 10.1086/321299
- Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x
- Diekmann, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., and Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to STEM careers: evidence for a goal congruity perspective. *J. Pers. Soc. Psychol.* 101, 902–918. doi: 10.1037/a0025199
- Durik, A. M., Vida, M., and Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: a developmental analysis. *J. Educ. Psychol.* 98, 382–393. doi: 10.1037/0022-0663.98.2.382
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368
- Eccles, J. S., Lord, S. E., Roeser, R. W., Barber, B. L., and Jozefowicz, D. M. (1997). "The association of school transitions in early adolescence with developmental trajectories through high school," in *Health Risks and Developmental Transitions During Adolescence*, eds J. Schulenberg, J. I. Maggs, and K. Hurrelmann (New York: Cambridge University Press), 283–321.
- Ferriman, K., Lubinski, D., and Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: developmental changes and gender differences during emerging adulthood and parenthood. *J. Pers. Soc. Psychol.* 97, 517–532. doi: 10.1037/a0016030
- Fralick, B., Kearn, J., Thompson, S., and Lyons, J. (2009). How middle schoolers draw engineers and scientists. *J. Sci. Educ. Technol.* 18, 60–73. doi: 10.1007/s10956-008-9133-3
- Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science* 320, 1164–1165. doi: 10.1126/science.1154094
- Hakim, C. (2006). Women, careers, and work-life preferences. *Br. J. Guid. Couns.* 34, 279–294. doi: 10.1080/03069880600769118
- Halpern, D. F. (2007). "Science sex and good sense: why women are underrepresented in some areas of science and math," in *Why Aren't more Women*

- in Science? *Top Researchers Debate the Evidence*, eds S. J. Ceci and W. M. Williams (Washington, DC: American Psychological Association), 121–130. doi: 10.1037/11546-010
- Hill, C., Corbett, C., and St. Rose, A. (2010). *Why so few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: American Association of University Women.
- Huang, B., Sivaganesan, S., Succop, P., and Goodman, E. (2004). Statistical assessment of mediational effects for logistic mediational models. *Stat. Med.* 23, 2713–2728. doi: 10.1002/sim.1847
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., and Williams, C. C. (2008). Gender similarities characterize math performance. *Science* 321, 494–495. doi: 10.1126/science.1160364
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421
- Joyce, B. A., and Farenga, S. J. (2000). Young girls in science: academic ability, perceptions, and future participation in science. *Roeper Rev.* 22, 261–262. doi: 10.1080/02783190009554048
- Koller, O., Baumert, J., and Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *J. Res. Math. Educ.* 32, 448–470. doi: 10.2307/749801
- Lubinski, D., and Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: uncovering antecedents for the development of math-science expertise. *Perspect. Psychol. Sci.* 1, 316–345. doi: 10.1111/j.1745-6916.2006.00019.x
- Lubinski, D., Webb, R. M., Morelock, M. J., and Benbow, C. P. (2001). Top 1 in 10,000: a 10-year follow-up of the profoundly gifted. *J. Appl. Psychol.* 86, 718–729. doi: 10.1037/0021-9010.86.4.718
- Maltese, A. V., and Tai, R. H. (2010). Eyeballs in the fridge: sources of early interest in science. *Int. J. Sci. Educ.* 32, 669–685. doi: 10.1080/09500690902792385
- Maltese, A. V., and Tai, R. H. (2011). Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among US students. *Sci. Educ.* 95, 877–907. doi: 10.1002/sce.20441
- Marks, H. M. (2000). Student engagement in instructional activity: patterns in the elementary, middle, and high school years. *Am. Educ. Res. J.* 37, 153–184. doi: 10.3102/00028312037001153
- Miller, J. D., and Kimmel, L. (2010). *Longitudinal Study of American Youth User's Manual: Student, Parent, and School Data (1987-Continuing)*. Ann Arbor, MI: ICPSR.
- National Academy of Engineering [NAE]. (2013). *Messaging for Engineering: From Research to Action*. Washington, DC: The National Academies Press.
- National Assessment of Educational Progress [NAEP]. (1986). *Math Objectives 1985–1986 Assessment*. Princeton, NJ: Educational Testing Service.
- National Science Foundation. (2011). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2011*. Arlington, VA: National Science Foundation.
- National Science Foundation. (2014). *Survey of Doctorate Recipients, 2013. National Center for Science and Engineering Statistics*. Available at: <http://ncesdata.nsf.gov/doctoratework/2013/>
- Pajares, F. (2005). “Gender differences in mathematics self-efficacy beliefs,” in *Gender Differences in Mathematics: An Integrative Psychological Approach*, eds A. M. Gallagher and J. C. Kaufmann (Boston, MA: Cambridge University Press), 294–315.
- Park, G., Lubinski, D., and Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychol. Sci.* 19, 957–961. doi: 10.1111/j.1467-9280.2008.02182.x
- Riegle-Crumb, C., King, B., Grodsky, E., and Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *Am. Educ. Res. J.* 49, 1048–1073. doi: 10.3102/00028312111435229
- Schwartz, S. H., and Rubel, T. (2005). Sex differences in value priorities: cross-cultural and multi-method studies. *J. Pers. Soc. Psychol.* 89, 1010–1028. doi: 10.1037/0022-3514.89.6.1010
- Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., et al. (2007). Assessing media influences on middle school-aged children's perceptions of women in science using the draw-a-scientist test (DAST). *Sci. Commun.* 29, 35–64. doi: 10.1177/1075547007306508
- Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364
- Tai, R. H., Liu, C. Q., Maltese, A. V., and Fan, X. (2006). Planning early for careers in science. *Science* 312, 1143–1144. doi: 10.1126/science.1128690
- Tyson, W. (2011). Modeling engineering degree attainment using high school and college physics and calculus course taking and achievement. *J. Eng. Educ.* 100, 760–777. doi: 10.1002/j.2168-9830.2011.tb00035.x
- Wai, J., Lubinski, D., Benbow, C. P., and Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: a 25-year longitudinal study. *J. Educ. Psychol.* 102, 860–871. doi: 10.1037/a0019454
- Wai, J., Putallaz, M., and Makel, M. C. (2012). Studying intellectual outliers: are there sex differences, and are the smart getting smarter? *Curr. Dir. Psychol. Sci.* 21, 382–390. doi: 10.1177/0963721412455052
- Wang, M. T. (2012). Educational and career interests in math: a longitudinal examination of the links between perceived classroom environment, motivational beliefs, and interests. *Dev. Psychol.* 48, 1643–1657. doi: 10.1037/a0027247
- Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001
- Wang, M. T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in STEM career choice. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937
- Webb, R. M., Lubinski, D., and Benbow, C. P. (2002). Mathematically facile adolescents with math-science aspirations: new perspectives on their educational and vocational development. *J. Educ. Psychol.* 94, 785–794. doi: 10.1037/0022-0663.94.4.785
- Wigfield, A., and Eccles, J. S. (2002). “The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence,” in *Development of Achievement Motivation*, eds A. Wigfield and J. S. Eccles (San Diego, CA: Academic Press), 91–120.
- Wigfield, A., Eccles, J. S., Davis-Kean, P., Roeser, R., and Scheifele, U. (2006). “Motivation to succeed,” in *Handbook of Child Psychology. Social, Emotional, and Personality Development*, Vol. 3, 6th Edn, eds W. Damon and N. Eisenberg (New York: Wiley), 933–1002.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 November 2014; accepted: 08 January 2015; published online: 17 February 2015.

Citation: Wang M-T, Degol J and Ye F (2015) Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Front. Psychol.* 6:36. doi: 10.3389/fpsyg.2015.00036

This article was submitted to *Developmental Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Wang, Degol and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Women are underrepresented in fields where success is believed to require brilliance

Meredith Meyer^{1*}, Andrei Cimpian² and Sarah-Jane Leslie³

¹ Department of Psychology, Otterbein University, Westerville, OH, USA, ² Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL, USA, ³ Department of Philosophy, Princeton University, Princeton, NJ, USA

OPEN ACCESS

Edited by:

Stephen J. Ceci,
Cornell University, USA

Reviewed by:

Sapna Cheryan,
University of Washington, USA
Elizabeth A. Gunderson,
Temple University, USA

*Correspondence:

Meredith Meyer,
Otterbein University,
1 South Grove Street,
Westerville, OH 43081, USA
mmeyer@otterbein.edu

Specialty section:

This article was submitted to
Developmental Psychology, a section
of the journal *Frontiers in Psychology*

Received: 01 January 2015

Paper pending published:
25 January 2015

Accepted: 15 February 2015

Published: 11 March 2015

Citation:

Meyer M, Cimpian A and Leslie S-J
(2015) Women are underrepresented
in fields where success is believed to
require brilliance.
Front. Psychol. 6:235.
doi: 10.3389/fpsyg.2015.00235

Women's underrepresentation in science, technology, engineering, and mathematics (STEM) fields is a prominent concern in our society and many others. Closer inspection of this phenomenon reveals a more nuanced picture, however, with women achieving parity with men at the Ph.D. level in certain STEM fields, while also being underrepresented in some *non*-STEM fields. It is important to consider and provide an account of this field-by-field variability. The field-specific ability beliefs (FAB) hypothesis aims to provide such an account, proposing that women are likely to be underrepresented in fields thought to require raw intellectual talent—a sort of talent that women are stereotyped to possess less of than men. In two studies, we provide evidence for the FAB hypothesis, demonstrating that the academic fields believed by laypeople to require brilliance are also the fields with lower female representation. We also found that the FABs of participants with college-level exposure to a field were more predictive of its female representation than those of participants without college exposure, presumably because the former beliefs mirror more closely those of the field's practitioners (the direct “gatekeepers”). Moreover, the FABs of participants with college exposure to a field predicted the magnitude of the field's gender gap above and beyond their beliefs about the level of mathematical and verbal skills required. Finally, we found that beliefs about the importance of brilliance to success in a field may predict its female representation in part by fostering the impression that the field demands solitary work and competition with others. These results suggest new solutions for enhancing diversity within STEM and across the academic spectrum.

Keywords: gender, stem, lay theories of success, field-specific ability beliefs, diversity in academia

Introduction

A recent article in *Scientific American Mind* begins: “Try this simple thought experiment. Name 10 female geniuses from any period of history. Odds are you ran out of names pretty quickly” (Upson and Friedman, 2012, p. 63). The thought experiment can be adapted: try to name 10 female figures in popular culture who—like Sherlock Holmes, Dr. House, or Will Hunting—are characterized by their innate brilliance, their raw intellectual firepower. As before, one rapidly runs out of names. Whatever the cause, the message is clear: women are not culturally associated with such inherent gifts of genius (Bennett, 1996, 1997, 2000; Tiedemann, 2000; Rammstedt and Rammsayer, 2002; Furnham et al., 2006; Kirkcaldy et al., 2007; Upson and Friedman, 2012; Lecklider, 2013;

Stephens-Davidowitz, 2014). The consequences of this stereotype are likely wide-ranging. In the current study, we focus on one of these consequences, asking whether such a pervasive cultural message might have a role in shaping individuals' academic and career paths. Specifically, if it is widely believed that men tend to possess more intellectual ability than women, then women may be discouraged from entering into fields that are thought to require this ability. We call this the field-specific ability beliefs (FAB) hypothesis (**Figure 1**): *the more a field is believed to require raw brilliance, the fewer the women* (Leslie and Cimpian et al., 2015). We test this hypothesis in the context of gender gaps in academia, investigating whether these gaps are predicted by how much laypeople assume that success in various fields rests on raw ability.

Gender disparity in academia has been a generative topic of research for many years, with contemporary focus on this issue largely centering on men's and women's participation in (natural) sciences, technology, engineering, and mathematics (STEM). The general phenomenon is clear: on average, female representation in STEM fields (particularly those that are math-intensive) is lower than in the social sciences and humanities (SocSci/Hum). Though the magnitude of this gap has largely decreased across the last several decades, the difference is still reliable, prompting a number of efforts to explain it (for reviews, see Ceci and Williams, 2007; Ceci et al., 2009, 2014; Hill et al., 2010).

The low number of women in STEM is indeed of real concern. However, it is also important to observe that there is at least as much variation in female representation *within* STEM and SocSci/Hum as there is *between* them. For instance, when examining the number of recent doctoral degrees earned by women in the U.S. in 30 different fields (**Table 1**), STEM fields are characterized by female representation ranging from

just under 20% (physics) to over 50% (molecular biology; National Science Foundation [NSF], 2011). An even larger range is observed within SocSci/Hum fields, with women earning fewer than 35% of doctoral degrees in philosophy and economics, yet over 75% in art history. Indeed, the range of variation is so wide that many STEM fields feature *higher* female representation at the Ph.D. level than many SocSci/Hum fields. Given this large variation within STEM and SocSci/Hum considered separately, it is apparent that expanding the focus of inquiry to include gender gaps in both STEM and SocSci/Hum might provide new insights into the problem of female underrepresentation. In the current study, we adopt such a broad focus, examining whether the FAB hypothesis can account for the field-by-field variability observed across the entire academic spectrum.

Initial Evidence for the FAB Hypothesis

In a recent study, we sought to test whether the FABs held by *academics* could predict the wide field-by-field variability observed in female representation across both STEM and non-STEM fields (Leslie and Cimpian et al., 2015). We gathered responses from a sample of over 1800 professors, graduate students, and post-doctoral researchers from research-intensive universities across the U.S. in 30 different fields (12 STEM, 18 SocSci/Hum; **Table 1**). We first asked participants to report on their beliefs regarding what was required for success in their own field, focusing on assessing beliefs about the relative importance of intrinsic, stable ability vs. effort and practice (see Dweck, 2000, 2006). We then used these items to provide a metric of field-level ability beliefs; each field received a FAB score expressing average endorsement of ability vs. effort across individuals within a given field (with higher scores indicating more emphasis on raw ability). Results included three important findings. First, FABs

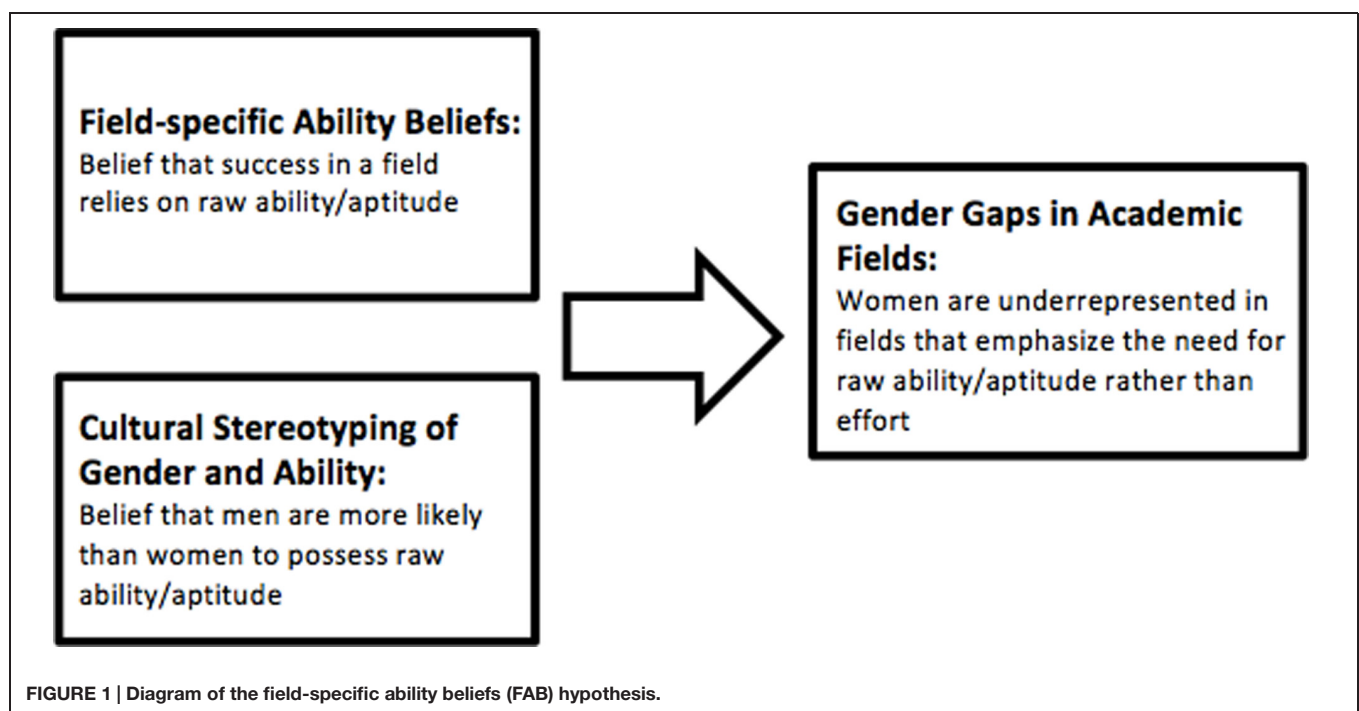


TABLE 1 | Percent of American Ph.D.'s earned by women in 2011* in science, technology, engineering, and mathematics (STEM) and Social science/Humanities fields.

STEM Field	% of Ph.D.'s who are Female	Social science/Humanities field	% of Ph.D.'s who are Female
Physics	18.0	Music theory and composition	15.8
Computer science	18.6	Philosophy	31.4
Engineering	22.2	Economics	34.4
Mathematics	28.6	Middle Eastern studies	38.1
Astronomy	29.2	Classics	41.8
Earth sciences	36.2	Political science	43.1
Chemistry	37.8	History	45.0
Statistics	41.6	Archeology	52.3
Biochemistry	45.4	Linguistics	59.2
Neuroscience	49.4	Anthropology	59.6
Evolutionary biology	49.8	Spanish/Spanish literature	59.9
Molecular biology	54.4	Comparative literature	60.9
		Sociology	61.3
		English literature	62.4
		Communication studies	64.2
		Education	69.3
		Psychology	72.1
		Art history	76.8

*Data from 2011 NSF Survey of Earned Doctorates.

were strongly negatively associated with female representation (as measured by proportion of U.S. Ph.D. degrees earned by women; National Science Foundation [NSF], 2011), providing initial broad support for the hypothesis: there were fewer women in fields believed to require stable, raw talent. Second, ability beliefs were predictive of female representation over and above whether a field was STEM or SocSci/Hum, suggesting that the FAB hypothesis can account well for the wide variability observed even within the two categories of fields. Finally, the FAB hypothesis outperformed a number of other constructs often theorized to contribute to gender gaps in academia, including field-specific variation in work-life balance (e.g., Ferriman et al., 2009), selectivity (e.g., Hedges and Nowell, 1995), and reliance on skills related to systemizing vs. empathizing (e.g., Baron-Cohen, 2002).

The Present Research

Do Laypeople's Beliefs Predict Female Representation?

Results from Leslie and Cimpian et al. (2015) suggest that women are underrepresented in fields whose practitioners consistently endorse the idea that success rests on brilliance. The current study extends existing work on the FAB hypothesis by exploring what *non-academics* believe is required for success in a variety of fields. We hypothesize that the FABs endorsed by academics will be shared, at least to some extent, by people outside of academia as well. This is an important extension of the FAB framework because, to the extent that these beliefs are shared by the general public, they could influence women's career choices in a much broader variety of contexts than the beliefs of academics *per se*. If these beliefs pervade our society, then—in combination with the stereotypes against women's intellectual abilities—they could lead a variety of individuals (parents, teachers, peers, etc.) to see women as somewhat unsuited for “brilliance-required”

domains. Even in the absence of such biased treatment, widely shared ability beliefs similar to those previously identified in academics could lead young women to doubt that they could succeed in brilliance-focused disciplines and thus to decide against pursuing careers in them. Our main prediction is thus that laypeople's beliefs, like those of academics, will predict female representation: the more a field is believed to require intellectual brilliance, the fewer the women.

We can also formulate a more detailed hypothesis here: people with more exposure to the fields in question (e.g., via college classes) will have FABs that predict female representation more precisely than the beliefs of people with less exposure. We expect this to be the case because the FABs of those with more exposure to a discipline will likely be more similar to those of the practitioners of that discipline, and are thus more likely to be similar to the kinds of beliefs that students will encounter and absorb as they start to consider higher education and careers in these fields (Leslie and Cimpian et al., 2015). In Study 1, we tested the prediction that college-exposed individuals' ability beliefs would better predict gender gaps in representation by first dividing our participants into those who had college exposure to a field and those who did not, and then exploring whether the beliefs of the college-exposed group predict female representation at a more fine-grained level.

A final aim of Study 1 was to address an alternative explanation for the hypothesized relationship between ability beliefs and female representation. As we have argued, underlying the main predictions described above is our claim that FABs influence women's academic and career choices. However, might laypeople's beliefs be simply inferred from their pre-existing knowledge about the proportion of women in the different fields? For instance, our participants—particularly those who have had

college experience in the relevant fields and thus have had the opportunity to witness gender disparities firsthand—might rely on stereotypes against women’s intellectual abilities to arrive at the conclusion that fields with few women must require high levels of such abilities, whereas fields with many women must not. To address this possibility, we asked participants to estimate the proportion of women in the fields under investigation, and assessed whether participants’ ability beliefs still predicted female representation independent from these estimates. If so, this would undermine the possibility that participants simply infer their ability beliefs from their estimates of the field’s diversity.

Assessing Beliefs Beyond FAB

There are surely many other dimensions that vary among fields and influence the gender breakdown of the people who participate in them, and we do not claim that FAB is the only factor in determining academic gender gaps. Indeed, as we observed at the outset, other such factors have been evaluated extensively in prior studies (for reviews, see Ceci and Williams, 2007; Ceci et al., 2009, 2014; Hill et al., 2010). Although an exhaustive evaluation of these additional factors is outside the scope of the current studies, we take up this issue within the framework of evaluating other *beliefs* about what is required for success in the fields under investigation. In particular, Study 2 examines two questions. First, do field-specific beliefs about the importance of intellectual brilliance reduce to beliefs about specific types of skills required for success? Specifically, do they reduce to beliefs about the degree to which mathematical and verbal skills are required for individual fields? Second, is the relationship between FABs and female representation mediated by beliefs about what kinds of work (solitary vs. collaborative; competitive vs. cooperative) are required?

Our first question addresses a potential alternative explanation for the predictive power of FABs. A critic might note that the extent to which mathematics is involved in a field appears to be particularly predictive of whether women are underrepresented or not: fields that are math-intensive attract and retain fewer women, with math-intensive STEM fields (e.g., engineering, math, or physics) characterized by the most extreme gender disparities (in comparison to STEM fields that are less math-intensive, like the life sciences, which often feature parity or even a predominance of women; National Science Foundation [NSF], 2011). The smaller number of women in math-intensive fields may be due in part to the cultural belief that math is “for” males, a belief that appears to emerge as early as elementary school and may contribute to women’s reduced interest in careers that require it (Fredericks and Eccles, 2002; Herbert and Stipek, 2005; Cvencek et al., 2011). In light of this evidence, one might ask: is it possible that the “intellectual brilliance” at the heart of the FAB hypothesis is just another way of referring to mathematical aptitude, which is also popularly conceived as a fixed, innate quantity? That is, might it be the case that people’s FABs simply reduce to their beliefs about how much individual fields require math over other kinds of skills (e.g., verbal skills)?

Results from Study 1 could bear on this question. If, as hypothesized we find that FABs are capable of predicting female representation across a variety of fields, including those unlikely

to be thought of as drawing on mathematical skills (such as most social sciences and humanities disciplines), it is unlikely that these beliefs are *merely* capturing people’s beliefs about field-specific mathematical requirements. However, it is important to more directly establish whether FABs are distinct from beliefs about the importance of mathematics. To do so, in Study 2 we tested whether beliefs about raw ability and brilliance predict unique variance in gender gaps, beyond that predicted by people’s beliefs about how much individual fields rely on mathematical and verbal ability.

Finally, it is worth noting that, as in Study 1, college exposure may matter. Participants with college experience likely have more nuanced, differentiated beliefs both about which fields require mathematical skills and about which fields require intrinsic ability. We thus hypothesized that when looking specifically at individuals with college exposure, FABs would independently predict female representation over and above beliefs about math and verbal skills, supporting the idea that FAB can account for female representation across the academic spectrum.

Next, we turn to the issue of potential beliefs that may mediate the relationship between FABs and female representation. In particular, we explore the possibility that people’s beliefs about the importance of brilliance vs. effort for success in a field give rise to differentiated perceptions of the kind of atmosphere that field promotes. We focused our exploration on two important aspects of a field’s atmosphere that (1) could be plausibly inferred based on the field’s presumed emphasis on brilliance, and that (2) men and women have diverging attitudes toward: namely, the extent to which the field requires *competition* (vs. collaboration) and *solitary work* (vs. group work; e.g., Lippa, 1998; Niederle and Vesterlund, 2007; Diekmann et al., 2010; Gupta et al., 2013; Lippa et al., 2014). There are several reasons why “brilliance-required” fields might also be presumed to require competition and solitary work. If a field values intellectual prowess, it is reasonable to expect that it would also encourage *displays* of that sort of ability, which might in turn encourage competition between individual practitioners. After all, it is only by comparing one’s ability against others (by participating in contests, engaging in aggressive debates, being harshly critical of others’ perceived mistakes, etc.) that one can reveal how brightly one’s intellectual ability shines. Working with others in cooperative contexts, on the other hand, would make it hard to assess whose talent was responsible for any ultimate success attained, so this type of collaborative work may be assumed to be rare within fields that prize brilliance. The inference that brilliance-requiring fields involve solitary, and often competitive, work is also likely to be supported by pervasive cultural tropes that portray brilliance and genius as qualities that a person possesses and displays in isolation rather than as part of a team of collaborators (e.g., Shenk, 2014). In turn, these inferences about the nature of the work environment in a field may influence whether young men and women consider careers in it because males and females are socialized to place different value on communal vs. agentic goals and on collaborative vs. competitive interactions. In other words, the downstream inferences licensed by FABs may be part of the reason why these

beliefs are predictive of gender gaps¹. We tested this hypothesis in Study 2.

Summary of Predictions

Study 1 examined two main predictions, one broad and one more specific. Broadly speaking, we expected that there would be a relationship between laypeople's FABs and female representation, such that fields believed to require brilliance would have fewer women. At a greater level of specificity, we expected that college exposure would differentiate the predictive power of FABs, such that the beliefs of those exposed to the fields during college would be particularly predictive. Finally, Study 1 also examined whether ability beliefs independently predicted female representation above and beyond people's *estimates* of female representation (suggesting that any observed relationship between ability beliefs and actual female representation did not emerge simply because individuals constructed FABs from their beliefs about female representation).

Study 2 was designed to replicate the main findings of Study 1, and to extend the inquiry into additional beliefs that might relate to gender disparities. We made two predictions. First, we predicted that FABs would *not* reduce to people's beliefs about mathematical skill, particularly when examining beliefs from individuals with college exposure in the field. Second, fields that are believed to require raw ability should also be perceived as requiring solo work and competition; in turn, these perceptions should predict gender gaps, with fewer women obtaining Ph.D.'s in fields assumed to demand high levels of solo work and competition. In other words, we expected that beliefs about solo work and competition would mediate (at least partially) the observed association between FAB and gender breakdowns.

Study 1

Method

Participants

Participants included 307 individuals recruited via Amazon's Mechanical Turk (MTurk), an online crowd-sourcing platform^{2,3}. Only participants reporting themselves as living in the U.S. and with prior MTurk approval rates of 90%

or above were included. Participants were compensated \$0.75 for survey completion. Data were excluded from an additional 48 individuals who (1) failed to complete the survey, (2) answered an attention-check question incorrectly, (3) had IP addresses indicating they were outside the U.S., and/or (4) had IP addresses indicating they had completed similar studies in the past.

Materials and Procedure

To avoid participant fatigue, we created three versions of the survey, each of which contained 10 of the 30 fields under investigation. (Fields were identical to those examined in our original study of academics (Leslie and Cimpian et al., 2015)). Fields were chosen to represent a broad spectrum of social sciences, humanities, and STEM disciplines. Approximately equal numbers of subjects participated in the three versions, and assignment was random (Version 1, $n = 103$; Version 2, $n = 101$; Version 3, $n = 103$). Each version included three humanities subjects, three social science subjects, and four STEM subjects. Each survey contained four questions assessing FABs about each of the 10 fields (from Leslie and Cimpian et al., 2015; **Table 2**). Questions were presented individually in random order with all 10 fields listed beneath each question. Participants indicated their agreement with the statement as it applied to each field using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree, with eight as an option to indicate "don't know"). Two attention-check questions were also included to ensure that participants were attending to the task.

Next, a series of questions asked about participants' academic exposure to the 10 fields, including whether they had had (1) a high school class, (2) a college class, and/or (3) a graduate-level class in each of them. Participants were also asked to estimate how many women had received American doctoral degrees in each field in the recent past, with 10 response options corresponding to 10% intervals ranging from 0 to 100%. A final set of questions asked about demographic information (gender, age, ethnicity, and race).

For each field, we calculated FAB scores by averaging scores across participants from the four ability belief questions. Higher scores indicated more emphasis on brilliance. Three separate FAB scores were calculated: (1) All Participants' FAB (using data from all participants except those with graduate level experience in the field)⁴, (2) College Exposure FAB (using data from participants who had taken college, but not graduate level, courses in the field) and (3) No College Exposure FAB (using data from participants who had taken neither college nor graduate courses in the field). The four items had high internal reliability (for all participants, $\alpha = 0.90$; for College Exposure, $\alpha = 0.93$; for No College Exposure, $\alpha = 0.89$).

¹We acknowledge that beliefs about solo/competitive work mediating FAB's relationship with women's representation represents only one possible causal pathway; it is also possible that people could perceive the solo/competitive nature of a field and then conclude that it requires raw ability. More generally, we also note that there are likely many more factors involved in the pathways that ultimately result in the observed field-by-field variation in women's representation. More comprehensive exploration of these factors, as well as experimental work, will be needed to definitively establish how FABs influence the observed gender gaps.

²All human subjects research reported in this paper was approved by the Institutional Review Board of the first author's home institution.

³Mechanical Turk offers a convenience sample rather than a fully nationally representative sample. Analyses of American MTurk workers have demonstrated that women are overrepresented, that workers are typically younger and more educated than average, and that Blacks and Hispanics are underrepresented (Berinsky et al., 2012; Paolacci and Chandler, 2014). Thus, we do not claim to be capturing beliefs that are fully representative of the U.S. public. Nevertheless, the diversity of an MTurk sample is arguably higher than that of most samples used in human subjects research (i.e., college samples), and it provides a good source of data for an examination of beliefs held by individuals outside of academia.

⁴We excluded data regarding individual fields if they were provided by people reporting graduate-level experience in that field. We did so because we wanted to exclude beliefs held by people with extensive familiarity with the field gained through graduate-level exposure, allowing the focus of the current study to be restricted only to individuals with no college experience vs. college experience with the field.

TABLE 2 | Survey items for Study 1 and Study 2.**Field-specific ability beliefs**

- Being a top scholar of [field] requires a special aptitude that just can't be taught.
- If you want to succeed in [field], hard work alone just won't cut it; you need to have an innate gift or talent.
- With the right amount of effort and dedication, anyone can become a top scholar in [field]. (R)
- When it comes to [field], the most important factors for success are motivation and sustained effort; raw ability is secondary. (R)
- To succeed in [field] you have to be a special kind of person; not just anyone can be successful in it. (in Study 2 only.)
- People who are successful in [field] are very different from ordinary people. (in Study 2 only.)

Estimate of female representation (Study 1)

Please provide your best guess or estimate to this question: in the recent past, what percentage of doctoral (Ph.D.) degrees from American universities do you think have been earned by women in [field]?

Verbal and mathematical ability (Study 2)

- Top-level success in [field] depends to a large extent on one's verbal ability.
- Top-level success in [field] depends to a large extent on one's mathematical ability.

Solo and competitive work (Study 2)

- [Field] is a field in which you spend a lot of time working by yourself rather than being around other people.
- [Field] is a field in which competition with others is much more common than collaboration.

(R) indicates items that were reverse scored.

Responses to all items except estimate of female representation were given on a 7-point scale (1 = strongly disagree to 7 = strongly agree), with an additional option for "don't know." Responses for estimate of female representation were given on a 10-point scale, with each point representing a 10% increment.

Results and Discussion

Study 1 tested two main predictions. First, we expected that participants' FABs would be correlated with female representation regardless of participants' level of direct prior exposure with the fields (via courses). Second, we predicted that beliefs held by individuals with college experience would nevertheless be predictive of female representation at a *finer-grained level* than those of people with no college experience. In particular, we expected that the College Exposure, but not the No College Exposure, FAB scores would predict female representation even after taking into account a gross STEM vs. non-STEM distinction between fields, which would speak to the ability of the College Exposure FAB scores to predict the complex field-by-field variability in female representation observed within these broad domains. Finally, we examined whether beliefs of college-exposed and non-college-exposed individuals predicted actual female representation independent of participants' *estimates* of female representation. If so, this would rule out the possibility that ability beliefs predicted female representation for the trivial reason that they were inferred from participants' pre-existing knowledge about gender disparities.

To assess our first prediction, we examined the correlation between FABs and female representation. Any fields for which we received fewer than 10 participants in either the no-college-experience or college-experience samples were removed from the analysis; estimates based on so few participants would likely be unreliable. This resulted in 29 fields being retained for analysis. (The single removed field was neuroscience; only seven individuals reported college experience with this field.) As predicted, fields believed to require brilliance had lower female representation, $r(27) = -0.59, p = 0.001$ (**Figure 2**).

To address the second prediction, we separately examined beliefs held by people with college exposure and those held by people without college exposure. Beliefs of both groups were significantly negatively associated with female representation:

College Exposure scores, $r(27) = -0.67, p < 0.001$, and No College Exposure scores, $r(27) = -0.51, p = 0.005$. Steiger's z score comparison (Lee and Preacher, 2013) indicated that College Exposure scores were more strongly associated with female representation than No College Exposure scores, $z = 2.09, p = 0.037$, providing initial support for the prediction that college-exposed individuals' beliefs would relate more strongly with representation. We then investigated whether the ability beliefs of these two groups predicted female representation above and beyond whether a field was STEM vs. a social science/humanities discipline (i.e., non-STEM). Two separate multiple regression analyses were performed with female representation as the dependent variable and two predictors: a STEM/non-STEM indicator variable and either (1) College Exposure FAB scores or (2) No College Exposure FAB scores. These analyses indicated that, as hypothesized, the FABs held by participants with college exposure to the fields were uniquely predictive of female representation, above and beyond whether the fields were in STEM or SocSci/Hum ($\beta = -0.44$, bootstrapped $p = 0.013$), whereas the beliefs of participants without college exposure were not ($\beta = -0.15$, bootstrapped $p = 0.449$).

Finally, we added college-exposed and non-college-exposed participants' estimates of female representation as predictors to the two regressions above. Consistent with our argument, the FABs of college-exposed participants remained a significant predictor of actual female representation even when adjusting for these participants' estimates of female representation ($\beta = -0.41$, bootstrapped $p = 0.043$; see **Table 3**). In contrast, the beliefs of participants without college exposure were not a significant predictor of female representation in this model ($\beta = -0.30$, bootstrapped $p = 0.257$; see **Table 3**). Thus, it is not the case that college-exposed participants' ability beliefs are predictive of gender gaps across academia simply because they are derived from prior knowledge of such gaps.

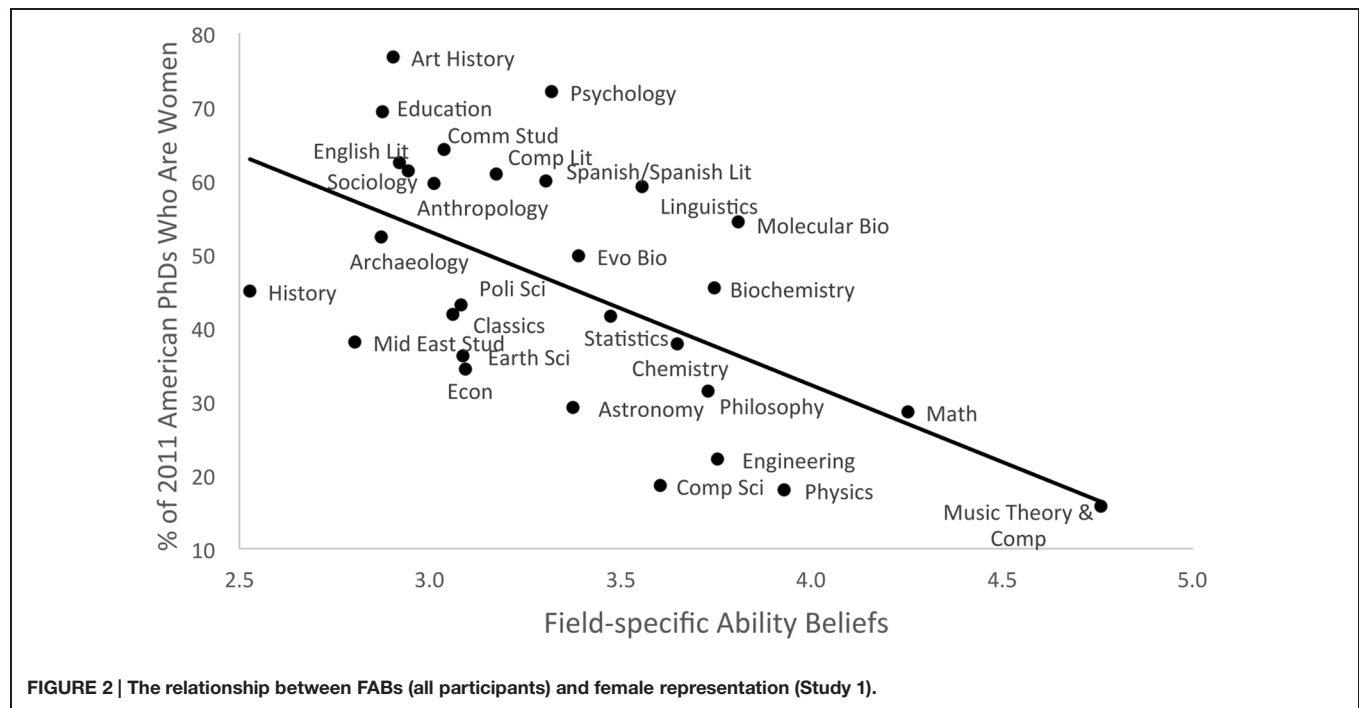


TABLE 3 | Regressions predicting female representation using field-specific ability beliefs and estimates of female representation of participants with college experience (CE; Upper) and with no college experience (NCE; Lower), Study 1.

Predictor			R^2	F	p
	β	p	0.686	15.87	< 0.001
STEM indicator	0.07	0.686			
Estimate of female representation (CE)	0.58	0.001			
Field-specific ability beliefs (CE)	−0.41	0.043			
	β	p	0.55	10.37	<0.001
STEM indicator	−0.22	0.340			
Estimate of female representation (NCE)	0.44	0.023			
Field-specific ability beliefs (NCE)	−0.30	0.257			

$n = 29$, $df (4,24)$.

We considered one final alternative interpretation, which applies particularly to the findings obtained with college-exposed individuals. Perhaps College Exposure FAB scores emphasize brilliance for fields where there are few women just because (1) men may be more likely than women to believe that brilliance is required for success, and (2) more men in the current sample may have taken college classes in disciplines where women are typically underrepresented. In other words, disciplines with lower female representation may have higher College Exposure FAB scores for the simple reason that male participants' brilliance-focused ability beliefs are overrepresented in

our sample for these disciplines. Consistent with this possibility, college-exposed men's scores ($M = 3.56$, $SD = 0.55$) were indeed higher than college-exposed women's scores ($M = 3.18$, $SD = 0.64$), $t(28) = 4.02$, $p < 0.001$, suggesting that men placed more emphasis on raw ability. In addition, our sample contained proportionately more college-exposed men in fields with lower female representation at the Ph.D. level, $r(27) = -0.66$, $p < 0.001$. To test whether these differences could explain our main result, we calculated a gender-balanced FAB score for each field by computing the average scores for men and women separately within fields, and then averaging these two gender-specific scores. This measure adjusts for the differential representation of college-exposed males and females across fields, giving the two groups an equal say in determining the FAB score for each field. If the current alternative explanation were correct, this gender-balanced score should no longer be predictive of female representation. However, when we entered the gender-balanced FAB score in place of the original FAB score in the regression including both STEM status and estimated female representation, it still predicted female representation ($\beta = -0.35$, bootstrapped $p = 0.069$). Thus, the main results described above were not merely a byproduct of men's brilliance-oriented beliefs inflating the College Exposure FAB scores of fields with fewer women.

In sum, the results of Study 1 lend clear support to the predictions we derived from the FAB model: women are less likely to be represented in fields believed to require stable, innate ability. Furthermore, as predicted, the field-specific beliefs of people with college experience in our fields were predictive of female representation at a more detailed level than were the beliefs of those without college experience. To speculate, perhaps initially people hold a global belief that disciplines in the STEM family require innate skill; as a result, the predictive power of these

initial, inchoate ability beliefs is mostly captured by the STEM vs. non-STEM distinction. It is only after exposure to the particularities of the fields and the beliefs of their practitioners that FABs take on independent predictive power in relation to female representation.

Study 2 provides an opportunity to replicate the above findings, and to further explore how gender breakdowns are related to field-differentiated beliefs about the types of skills and work that are required. Two predictions are central to Study 2. First, we expect that the FABs of participants with college experience will predict unique variance in female representation, above and beyond their beliefs about the role of mathematical or verbal skills. Second, we predict that participants' assumptions about how much solitary and competitive work is required by individual fields will mediate the relationship between FABs and female representation.

Study 2

Method

Participants

Participants included 302 individuals recruited via Amazon's MTurk, using the same inclusion criteria as in Study 1. Participants were compensated \$0.95 for survey completion. Data were excluded from an additional 53 individuals who met one or more of the exclusion criteria used in Study 1: (1) failing to complete the survey, (2) answering an attention check question incorrectly, (3) having an IP address suggesting residence outside the U.S., and/or (4) having IP addresses indicating completion of similar studies (including our Study 1) in the past.

Materials and Procedure

As in Study 1, three versions of the survey were created, each of which contained the same subsets of 10 of the 30 fields under investigation. Approximately equal numbers of subjects participated in the three versions, and assignment was random (Version 1, $n = 101$; Version 2, $n = 103$; Version 3, $n = 98$). Surveys included the same four FAB items as in Study 1, along with two additional, broader questions on this topic (Table 2). These questions were added to further assess FABs, with the goal of using more accessible language while still providing a sensitive index of participants' beliefs about innate ability vs. effort. Two items were also included to address people's beliefs about the extent to which verbal and mathematical skills are required, and two final items were included to assess beliefs about whether competition and solitary work are important for success in a field (Table 2).

As in Study 1, items were presented individually in random order with all 10 fields listed beneath each item. Participants again indicated their agreement with the statement as it applied to each of the 10 fields using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree, with eight as an option to indicate "don't know"). Two attention-check questions were also included. The survey then ended with questions assessing high school, college, and graduate level exposure to each of the 10 fields, along

with several demographic questions. (These questions were all identical to those in Study 1).

We calculated FAB scores by averaging scores across the six items, and then averaging within fields to create field-level scores. Three separate FAB scores were calculated reflecting (1) All Participants' FAB (using data from all participants except those with graduate level experience in the field), (2) College Exposure FAB (using data from participants who had taken college, but not graduate level, courses in the field), and (3) No College Exposure FAB (using data from participants who had taken neither college nor graduate courses in the field). Scores for the six ability beliefs questions had high internal reliability (for all participants, $\alpha = 0.89$; for College Exposure, $\alpha = 0.93$; for No College Exposure, $\alpha = 0.87$). Deletion of the last two items added for Study 2 did not improve scale reliability, indicating it was appropriate to include them as part of the FAB scale.

Results and Discussion

To explore whether Study 2 replicated the key finding that FABs predict female representation, we again examined correlations between FAB and percentage of female Ph.D. recipients. As before, fields with fewer than 10 participants reporting either college or no college exposure were removed. This resulted in 27 fields being retained for analysis. (Middle Eastern studies, neuroscience, and archeology were removed because they had College Exposure n s of 8, 4, and 3, respectively.) Replicating findings from Study 1, FAB scores were negatively associated with female representation when examining belief scores of all participants, $r(25) = -0.63$, $p < 0.001$, as well as when examining College Exposure scores, $r(25) = -0.65$, $p < 0.001$, and No College Exposure scores, $r(25) = -0.54$, $p = 0.004$. Although College Exposure scores were more strongly associated with female representation than No College Exposure scores, this difference was not significant according to Steiger's z score comparison (Lee and Preacher, 2013), $z = 1.13$, $p = 0.258$. However, again replicating Study 1, we found that the FABs of participants with college experience predicted female representation even when a STEM indicator variable was added to the regression model as a competitor ($\beta = -0.37$; bootstrapped $p = 0.048$); in contrast, the beliefs of those without college experience were not uniquely predictive when the STEM indicator was added ($\beta = -0.15$; bootstrapped $p = 0.48$). Thus, beliefs held by college-exposed individuals again predicted female representation better than those of non-college-exposed individuals.

We next examined the relationship between female representation and the extent to which a field is perceived as demanding verbal and mathematical skills. Beliefs about the need for verbal skills were positively associated with female representation: beliefs of all participants, $r(25) = 0.63$, $p < 0.001$; of participants with college exposure, $r(25) = 0.63$, $p = 0.001$; of participants with no college exposure, $r(25) = 0.65$, $p < 0.001$. Beliefs about the need for mathematical skills were negatively associated with female representation: beliefs of all participants, $r(25) = -0.64$, $p < 0.001$; of participants with college exposure, $r(25) = -0.60$, $p = 0.001$; of participants with no college exposure $r(25) = -0.64$, $p < 0.001$.

We then tested our prediction that FABs of individuals with college exposure would predict female representation independently from beliefs about the role of mathematical and verbal skills. If so, this would strengthen the claim that FABs tap into something distinct from people's beliefs about which fields require mathematical aptitude. To assess this prediction, we added perceptions of the need for verbal and mathematical skill as variables in the two regressions predicting female representation. For the regression testing beliefs of those *with* college experience (Table 4), FABs were uniquely predictive of women's representation, above and beyond STEM status and beliefs about the importance of mathematical and verbal skill, $\beta = -0.39$, bootstrapped $p = 0.085$, although this coefficient was only significant at the $\alpha = 0.10$ level. In contrast, beliefs about the importance of verbal and mathematical ability did not independently predict female representation in this model, $ps > 0.489$. For the regression testing the beliefs of those *without* college experience, no factor was significantly predictive of female representation, $ps > 0.454$ (Table 4).

As in Study 1, we also calculated a gender-balanced FAB score to examine the possibility that differences in male and female participants' ability beliefs and college experience were driving the effects observed for college-exposed participants. (To reiterate, the possibility being tested here is that College Exposure FAB scores in fields with fewer women are inflated simply because men may have ability beliefs that are more brilliance-oriented and may also be overrepresented in the college-exposure sample for these fields.) Again, the proportion of college-exposed male participants within each field was negatively related to female representation at the Ph.D. level, $r(25) = -0.41$, $p = 0.03$, indicating that college-exposed male participants were more numerous in fields with lower female representation. In this sample, however, college-exposed men's FAB scores ($M = 3.70$, $SD = 0.71$) were actually lower than college-exposed women's

scores ($M = 3.85$, $SD = 0.60$), though not significantly so, $t(26) = 1.69$, $p = 0.103$. Thus, this alternative explanation is unlikely: college-exposed male participants, though more numerous in fields with fewer women at the Ph.D. level, did *not* differ from college-exposed female participants in their ability beliefs. Nevertheless, we entered a gender-balanced FAB score in place of the original FAB score in the regression model that also included a STEM indicator, beliefs about mathematical ability, and beliefs about verbal ability. As before, ability beliefs were the sole predictor of female representation ($\beta = -0.44$, bootstrapped $p = 0.059$). These results strengthen the main claim that ability beliefs are predictive of female representation, above and beyond beliefs about mathematical and verbal skills.

Finally, we tested the prediction that beliefs about solo work and competitiveness would mediate the relationship between FABs and female representation. Consistent with our argument, a bootstrapped (1,000 replications) product-of-coefficients mediation analysis performed with the PROCESS procedure in SPSS 22 (Hayes, 2013) revealed that the relationship between college-exposed participants' ability beliefs about a discipline and the proportion of female Ph.D.'s in that discipline was significantly mediated by these participants' ideas about the amount of solo work and the level of competitiveness required by the discipline, $ab = -13.56$ (-26.74 , -2.91). Similar results were obtained when examining beliefs of non-college-exposed participants, $ab = -13.61$ (-24.65 , -5.94)⁵. (For full results of the mediation models, see Figures 3 and 4.) Results are thus consistent with the idea that FABs may influence women's participation in a field in part by influencing their beliefs about what it is like to be a member of that field—in particular, whether one works by oneself or with others, and whether success rests more on competition with colleagues rather than cooperation. Interestingly, this result was observed even within the group who had not had college exposure to the field, which may be because inferences about the nature of the work demanded by various fields are easily drawn from one's ability beliefs about these fields, no matter how much first-hand experience one has with them.

General Discussion

Women are underrepresented in many STEM fields, but the pattern of gender distribution is complex, and a substantial amount of variation also exists in non-STEM fields. An important aim of the current studies was to provide an account for the wide variability in female representation across the entire academic spectrum. We maintain that the FAB hypothesis provides such an account. This hypothesis predicts that women will be underrepresented in fields believed to emphasize brilliance and inherent ability as the key to success; this is because women are often stereotyped as lacking the same sort of innate intelligence as men, and thus women will be discouraged from participating in fields to the extent that these fields are perceived as requiring this type

TABLE 4 | Regressions predicting female representation using field-specific ability beliefs and beliefs about the importance of verbal and mathematical skill of participants with college experience (CE; Upper) and with no college experience (NCE; Lower), Study 2.

Predictor			R^2	F	p
	β	p	0.52	6.05	0.002
STEM indicator	-0.11	0.747			
Field-specific ability beliefs (CE)	-0.39	0.085			
Verbal skill beliefs (CE)	0.26	0.489			
Mathematical skill beliefs (CE)	-0.06	0.820			
	β	p	0.49	5.29	0.004
STEM indicator	-0.21	0.524			
Field-specific ability Beliefs (NCE)	-0.14	0.516			
Verbal skill beliefs (NCE)	0.17	0.614			
Mathematical skill beliefs (NCE)	-0.26	0.454			

$n = 27$, df (4,22).

⁵The indirect paths were again significant even when adjusting for beliefs about mathematical skills, both for college-exposed participants, $ab = -9.21$ (-22.11 , -1.73) and for non-college-exposed participants, $ab = -9.54$ (-19.77 , -2.89).

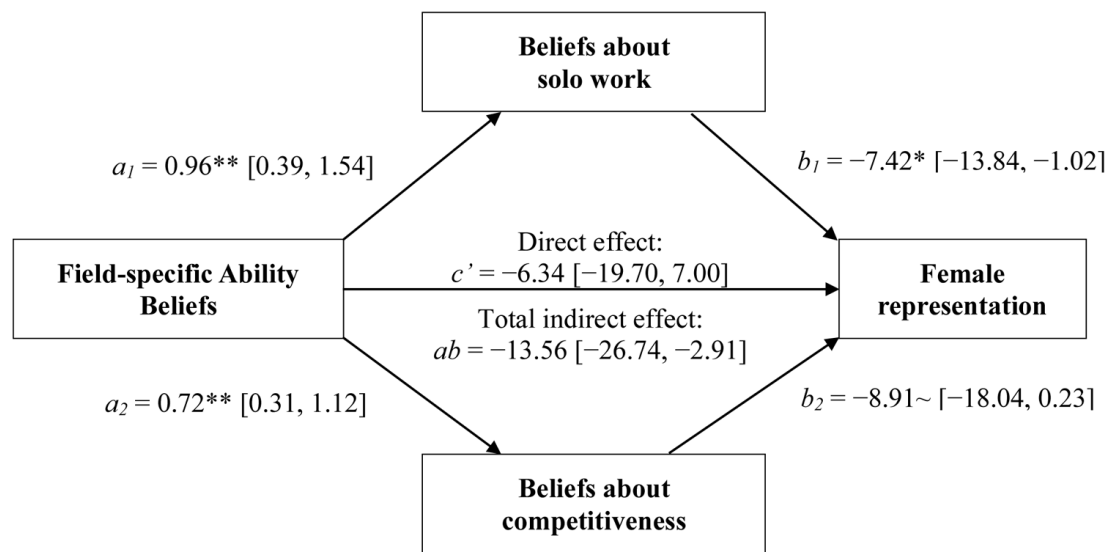


FIGURE 3 | The indirect pathways linking college-exposed participants' FABs with women's representation via participants' beliefs about the amount of solo work (top) and the level of competitiveness (bottom) required by a field (Study 2). $\sim p < 0.10$; $*p < 0.05$; $**p < 0.01$.

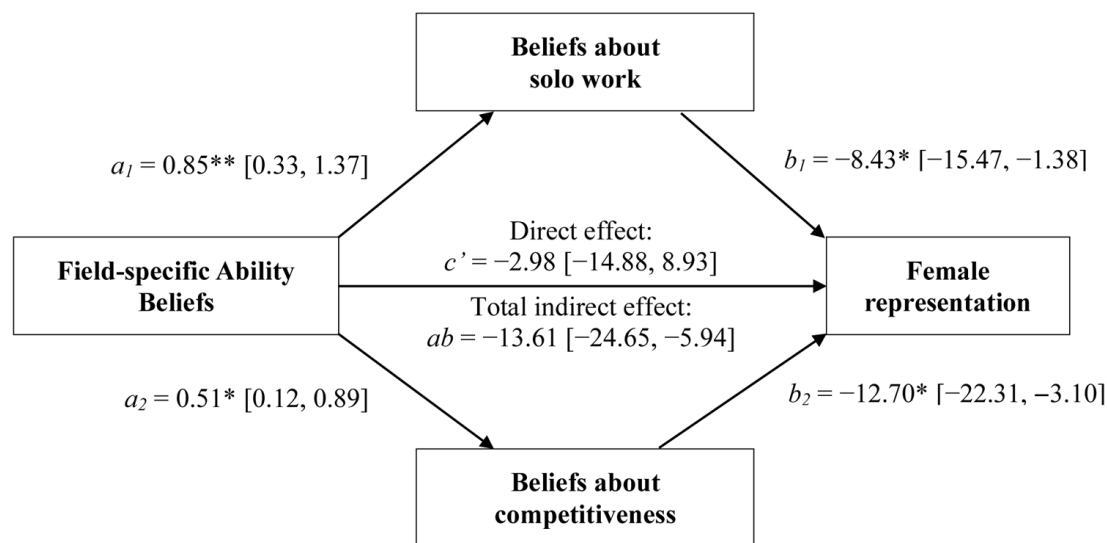


FIGURE 4 | The indirect pathways linking non-college-exposed participants' FABs with women's representation via participants' beliefs about the amount of solo work (top) and the level of competitiveness (bottom) required by a field (Study 2). $*p < 0.05$; $**p < 0.01$.

of intelligence. Prior research has provided support for the FAB hypothesis within higher academia (Leslie and Cimpian et al., 2015). The current studies extended the focus to an examination of beliefs held by individuals outside academia. The results of our two studies are consistent with the FABs hypothesis: fewer women are involved in fields that laypeople believe to require raw intellectual ability.

Several additional findings from the present studies are worth highlighting. The ability beliefs of individuals who had college-level exposure to the fields in question predicted female representation even when controlling for whether a field was in

STEM or not, indicating that college may provide a unique context for refinement and elaboration of beliefs about what fields require for success. Results also suggested that the ability beliefs of participants with college experience are not simply a byproduct of participants' inferring these beliefs based on their prior knowledge of female representation (Study 1). Further, college-exposed participants' ability beliefs capture something beyond perceptions of specific types of skills required for success, as FABs of college-exposed individuals did not reduce to beliefs about which fields require mathematical and/or verbal skills (Study 2).

Notably, these findings have important consequences for potential interventions to improve diversity, both in terms of timing and in terms of content. College may be a pivotal experience during which people's FABs become entrenched, and start to conform to those of their instructors. This highlights the crucial role that college educators play in communicating these maladaptive beliefs—but also suggests that they may be able to play an active role in changing the relevant messages. In particular, our data suggest that instructors who want to promote diversity might aim to minimize discussion of innate talent, regardless of the domain of skills with which it is associated, and instead highlight the importance of effort, practice, and persistence to success in a field. Prior work on individuals' achievement beliefs suggests that such growth-oriented messages can be relayed in a range of ways: by choice of adjectives (in particular by avoiding words like “brilliant,” “genius,” etc.; , Mueller and Dweck, 1998; Cimpian et al., 2007; Heyman, 2008), by focusing on what the person *has achieved* rather than on the person's *inherent traits* (Kamins and Dweck, 1999), and by explicitly stating that dedication and effort are paramount (Mueller and Dweck, 1998; Kamins and Dweck, 1999; Blackwell et al., 2007). We expect that practices such as these would be easily implementable by college educators across many fields. It should be borne in mind that the messages that college educators send may not only affect the participation of the women in their classes, but also have more far-reaching impact. As their students—both men and women—may go on to become parents, caregivers, school teachers, etc., they may subtly communicate their own ability beliefs to future generations (e.g., through their own choice of adjectives; Cimpian et al., 2007). This in turn may influence even very young girls' engagement and educational choices (Cimpian et al., 2014; Leslie et al., 2015).

The current studies also suggest that beliefs about solo and competitive work may mediate the relationship between ability beliefs and female representation. It is possible that this result reflects a process by which ability beliefs influence perceptions of what it is like to work in certain fields, which in turn may influence the participation of women in these fields. Of course, we acknowledge this is not the only possible pathway here; our mediation analyses were designed to test an a priori hypothesis regarding how ability beliefs relate to representation, but they cannot determine directionality. Similarly, causality regarding ability beliefs and female representation cannot be claimed from the current studies due to the correlational nature of the data. However, our theoretical model posits that ability beliefs do drive women's career and educational choices, and recent experimental manipulations in our lab have provided evidence consistent with this causal claim (e.g., Cimpian et al., 2014). For instance,

simply describing a novel educational or professional opportunity as requiring raw talent (vs. dedication) was sufficient to lower women's—and even young girls'—motivation to pursue it. Thus, there is some independent evidence suggesting that the relationship between ability beliefs and female representation is due to the causal influence of the ability beliefs.

Further investigating the precise pathways by which non-academics' ability beliefs influence participation is one important topic for future research. To begin, it is worth noting that young men and women often decide whether or not to pursue a field long before interacting with professors, graduate students, or any other active practitioners of that field (e.g., Watt and Eccles, 2008). Indeed, many fields with disproportionately high representation of men at the Ph.D. level see gender disparities in interest as early as elementary school (Lubinski and Benbow, 2006; Ceci and Williams, 2010; Cvencek et al., 2011). From our viewpoint, some of these early differences may be due to the ability beliefs of people outside of academia (teachers, parents, peers, etc.). For example, adults' FABs, in combination with the stereotype that females are less likely than males to be brilliant, could lead to small differences in the extent to which adults encourage girls' and boys' interest in fields believed to require this intellectual trait, the extent to which they provide boys and girls with opportunities to develop their skills in these fields, the extent to which they dwell on boys' and girls' achievements in these fields, and so on. Adults are also likely to convey their FABs to the children themselves. Once absorbed, these beliefs might make it more difficult for girls to consider careers in fields believed to require brilliance (again, since the ambient stereotypes portray them as being unsuited for these fields). As well, children might communicate these FABs to their peers, either via explicit statements or more subtly—say, by reacting with surprise to behaviors that are inconsistent with these beliefs. As a result of these multiple parallel processes, young women may be less likely to be interested in “brilliance-required” fields, and those who do pursue them may be less likely to persist and achieve at the same levels as men.

In summary, we have provided support for the FAB hypothesis, demonstrating that women tend to be underrepresented in fields believed to require innate intellectual talent for success. Our data also open up possibilities for future research on the pathways by which ability beliefs influence women's participation. Finally, these studies point to possibilities for effective interventions. If the practitioners of fields with gender gaps made a concerted effort to highlight the role of sustained, long-term effort in achievement, the gender gaps in these fields may correspondingly be diminished.

References

- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends Cogn. Sci.* 6, 248–254. doi: 10.1016/S1364-6613(02)01904-6
- Bennett, M. (1996). Men's and women's self-estimates of intelligence. *J. Soc. Psychol.* 136, 411–412. doi: 10.1080/0022454.1996.9714021
- Bennett, M. (1997). Self-estimates of ability in men and women. *J. Soc. Psychol.* 137, 540–541. doi: 10.1080/00224549709595475
- Bennett, M. (2000). Self-estimates and population estimates of ability in men and women. *Aust. J. Psychol.* 52, 23–28. doi: 10.1080/00049530008255363
- Berinsky, A. J., Huber, G. A., and Linz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis* 20, 351–368. doi: 10.1093/pan/mpr057
- Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Dev.* 78, 246–263. doi: 10.1111/j.1467-8624.2007.00995.x

- Ceci, S. J., Ginther, D., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest*. 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., and Williams, W. M. (2007). *Why aren't More Women in Science? Top Researchers Debate the Evidence*. Washington, DC: American Psychological Association.
- Ceci, S. J., and Williams, W. M. (2010). Sex differences in math-intensive fields. *Curr. Dir. Psychol. Sci.* 19, 275–279. doi: 10.1177/0963721410383241
- Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412
- Cimpian, A., Arce, H. C., Markman, E. M., and Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychol. Sci.* 18, 314–316. doi: 10.1111/j.1467-9280.2007.01896.x
- Cimpian, A., Bian, L., and Leslie, S.-J. (2014). Experimental and developmental evidence for the field-specific ability beliefs hypothesis. *Paper presented at the Fifteenth Annual Meeting of the Society for Personality and Social Psychology*, Austin, TX.
- Cvencek, D., Meltzoff, A., and Greenwald, A. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x
- Diekmann, A. B., Brown, E. R., Johnston, A. M., and Clark, E. K. (2010). Seeking congruity between roles and goals: a new look at why women opt out of STEM careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342
- Dweck, C. S. (2000). *Self-theories: Their Role in Motivation, Personality, and Development*. Philadelphia, PA: Psychology Press.
- Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. New York: Random House.
- Ferriman, K., Lubinski, D., and Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: developmental changes and sex differences during emerging adulthood and parenthood. *J. Pers. Soc. Psychol.* 97, 517–532. doi: 10.1037/a0016030
- Fredericks, J. A., and Eccles, J. E. (2002). Children's competence and value beliefs from childhood through adolescence. *Dev. Psychol.* 38, 519–533. doi: 10.1037/0012-1649.38.4.519
- Furnham, A., Crawshaw, J., and Rawles, R. (2006). Sex differences in self-estimates on two validated IQ test subscale scores. *J. Appl. Psychol.* 36, 417–440. doi: 10.1111/j.0021-9029.2006.00013.x
- Gupta, N. G., Poulsen, A., and Villeval, M. C. (2013). Gender matching and competitiveness: experimental evidence. *Econ. Inq.* 51, 816–835. doi: 10.1111/j.1465-7295.2011.00378.x
- Hayes, A. F. (2013). *Introduction to Mediation, Moderation and Conditional Process Analysis*. New York: Guilford Press.
- Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. doi: 10.1126/science.7604277
- Herbert, J., and Stipek, D. T. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Appl. Dev. Psychol.* 26, 276–295. doi: 10.1016/j.appdev.2005.02.007
- Heyman, G. D. (2008). Talking about success: implications for achievement motivation. *J. Appl. Dev. Psychol.* 29, 361–370. doi: 10.1016/j.appdev.2008.06.003
- Hill, C., Corbett, C., and St. Rose, A. (2010). *Why so few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: American Association of University Women.
- Kamins, M. L., and Dweck, C. S. (1999). Person versus process praise and criticism: implications for contingent self-worth and coping. *Dev. Psychol.* 35, 835–847. doi: 10.1037/0012-1649.35.3.835
- Kirkcaldy, B., Noack, P., Furnham, A., and Siefen, G. (2007). Parental estimates of their own and their children's intelligence. *Eur. Psychol.* 12, 173–180. doi: 10.1027/1016-9040.12.3.173
- Lecklider, A. (2013). *Inventing the Egghead*. Philadelphia, PA: University of Pennsylvania Press. doi: 10.9783/9780812207811
- Lee, I. A., and Preacher, K. J. (2013). *Calculation for the Test of the Difference Between Two Dependent Correlations with one Variable in Common [Computer software]*. Available at: <http://quantpsy.org>
- Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Women are underrepresented in disciplines that emphasize brilliance as the key to success. *Science* 347, 262–265.
- Leslie, S.-J., Saunders, K., and Rhodes, M. (2015). “Let's be scientists!": Hearing generic language about scientists lowers preschool-aged girls' motivation and engagement with a science task,” in *Proceedings of the Symposium Conducted at the Meeting of the Society for Personality and Social Psychology: Bringing Thought Experiments to Life: Perspectives on the Essence of Identity from Psychology and Philosophy*, eds L. Heiphetz and A. Waytz (Chairs), Long Beach, CA.
- Lippa, R. A. (1998). Gender-related individual differences and the structure of vocational interests: the importance of the people-things dimension. *J. Pers. Soc. Psychol.* 74, 996–1009. doi: 10.1037/0022-3514.74.4.996
- Lippa, R. A., Preston, K., and Penner, J. (2014). Women's representation in 60 occupations from 1972 to 2010: more women in high-status jobs, few women in things-oriented jobs. *PLoS ONE* 9:e95960. doi: 10.1371/journal.pone.0095960
- Lubinski, D., and Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: uncovering antecedents for math science expertise. *Perspect. Psychol. Sci.* 1, 316–345. doi: 10.1111/j.1745-6916.2006.00019.x
- Mueller, C. M., and Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *J. Pers. Soc. Psychol.* 75, 33–52. doi: 10.1037/0022-3515.75.1.33
- National Science Foundation [NSF]. (2011). *Survey of Earned Doctorates*. Available at: <http://www.nsf.gov/statistics/srvydoctorates/> (2011).
- Niederle, M., and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Q. Econ.* 122, 1067–1101. doi: 10.1162/qjec.122.3.1067
- Paolacci, J., and Chandler, J. (2014). Inside the turk: understanding mechanical turk as a participant pool. *Curr. Dir. Psychol. Sci.* 23, 184–188. doi: 10.1177/0963721414531598
- Rammstedt, B., and Ramsayer, T. H. (2002). Self-estimated intelligence: gender differences, relationship to psychometric intelligence and moderating effects of level of education. *Eur. Psychol.* 7, 275–284. doi: 10.1027/1016-9040.7.4.275
- Shenk, J. W. (2014). The end of 'genius.' *The New York Times*. Available at: <http://www.nytimes.com>
- Stephens-Davidowitz, S. (2014). *Google, Tell Me: Is My Son a Genius?* Available at: http://www.nytimes.com/2014/01/19/opinion/sunday/google-tell-me-is-my-son-a-genius.html?_r=0 [accessed October 16, 2014].
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Educ. Stud. Math.* 50, 40–62.
- Upson, S., and Friedman, L. F. (2012). Where are all the female geniuses? *Sci. Am. Mind* 23, 63–65. doi: 10.1038/scientificamericanmind1112-63
- Watt, H. M. G., and Eccles, J. S. (eds). (2008). *Gender and Occupational Outcomes: Longitudinal Assessments of Individual, Social, and Cultural Influences*. Washington, DC: American Psychological Association.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Meyer, Cimpian and Leslie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes

Sapna Cheryan^{1*}, Allison Master^{1,2} and Andrew N. Meltzoff^{1,2}

¹ Department of Psychology, University of Washington, Seattle, WA, USA

² Institute for Learning & Brain Sciences, University of Washington, Seattle, WA, USA

Edited by:

Stephen J. Ceci, Cornell University, USA

Reviewed by:

Andrei Cimpian, University of Illinois, USA

Toni Schmader, University of British Columbia, Canada

*Correspondence:

Sapna Cheryan, Department of Psychology, University of Washington, Box 351525, Seattle, WA 98195, USA
e-mail: scheryan@uw.edu

Despite having made significant inroads into many traditionally male-dominated fields (e.g., biology, chemistry), women continue to be underrepresented in computer science and engineering. We propose that students' stereotypes about the culture of these fields—including the kind of people, the work involved, and the values of the field—steer girls away from choosing to enter them. Computer science and engineering are stereotyped in modern American culture as male-oriented fields that involve social isolation, an intense focus on machinery, and inborn brilliance. These stereotypes are compatible with qualities that are typically more valued in men than women in American culture. As a result, when computer science and engineering stereotypes are salient, girls report less interest in these fields than their male peers. However, altering these stereotypes—by broadening the representation of the people who do this work, the work itself, and the environments in which it occurs—significantly increases girls' sense of belonging and interest in the field. Academic stereotypes thus serve as gatekeepers, driving girls away from certain fields and constraining their learning opportunities and career aspirations.

Keywords: science, underrepresentation, belonging, gender, stereotypes

In 2010, Mattel let girls vote online for which career they wanted Barbie to have next. They gave girls a choice of one of five careers: news anchor, architect, surgeon, environmentalist, and computer engineer. Computer Engineer Barbie ended up winning by a landslide after female engineers and others in technology launched online campaigns in technology communities to get out the vote. Their hope was that future generations of girls would play with Computer Engineer Barbie and be inspired to pursue careers in computer science and engineering (Martincic and Bhatnagar, 2012). After the voting closed, Mattel announced the simultaneous release of two of the Barbies: Computer Engineer and News Anchor. Although Computer Engineer Barbie had won the “popular vote,” Mattel's empirical research showed that the “girls' vote” went to News Anchor Barbie (Zimmerman, 2010). This anecdote is symbolic of a broader trend in our society: despite efforts by people in education, technology, government, and non-profits to get girls interested in a future in computer science and engineering, girls are choosing other fields.

Women currently make up 48% of medical school graduates and 47% of law school graduates (Jolliff et al., 2012; American Bar Association, 2014). Even within STEM (science, technology, engineering, and math), women obtain the majority of the U.S. undergraduate degrees (59%) in biology and nearly half in chemistry and math (National Science Foundation, 2013). However, in computer science and engineering, women earn less than 20% of undergraduate degrees (National Science Foundation, 2013). Gender disparities in computer science and engineering are problematic for at least three reasons. First, jobs in these fields are often high-status, lucrative, and flexible (Kalwarski et al., 2007), and thus

women are missing out on jobs that are potentially beneficial for them. Second, computer scientists and engineers design tools that shape modern society, and diversifying the field can help to ensure that these fields are creating designs appropriate for a broad population (Margolis and Fisher, 2002). Third, the U.S. is currently not training enough computer scientists and engineers to keep up with demand (Soper, 2014). Attracting more women and people of color would be an effective way of reducing this gap.

Women have entered many other previously male-dominated fields, including other STEM fields, but not computer science and engineering. Why the differential? According to Gelernter (1999), professor of computer science at Yale, the explanation for women's underrepresentation is obvious, “Women... must be choosing not to enter, presumably because they don't want to; presumably because they (by and large) don't like these fields.” His statement assumes that women's choices are freely made and not constrained. If women are freely choosing not to pursue computer science, perhaps nothing can or should be done about it—after all, it is their choice. However, it is clear from a large body of scientific research that there are significant social barriers to women's entry into computer science and engineering that preclude women from being able to make a truly “free” choice (Ceci et al., 2009). Here we analyze those barriers and what can be done about them.

In what ways are girls' educational choices constrained? First, girls may be steered away from computer science and engineering by parents, teachers, and others who think that these careers are better suited for boys (Eccles et al., 1990; Sadker and Sadker, 1994). Second, the mere fact of having underrepresentation can perpetuate future underrepresentation (Murphy et al., 2007). If

girls do not see computer scientists and engineers as people with whom they feel similar, they may be more reluctant to enter these fields (Dasgupta, 2011; Meltzoff, 2013). Third, girls systematically underestimate how well they will do in these fields, and this predicts their lower interest in entering them (Correll, 2001; Ehrlinger and Dunning, 2003). Fourth, girls may anticipate encountering greater work-family conflicts in these fields (Ceci et al., 2009). Fifth, there is discrimination in these fields that prevents qualified women from receiving the same opportunities as their male counterparts (Moss-Racusin et al., 2012). Sixth, women who enter traditionally masculine domains can be socially and professionally penalized for exhibiting competence and leadership qualities (Rudman, 1998). These are all barriers that contribute to why some women choose not to enter and persist in fields like computer science and engineering. Note, however, that these barriers previously existed (and continue to exist) in other male-dominated fields that women have entered. A key question remains: *what has allowed other fields to welcome more women while computer science and engineering continue to lag behind?*

In this paper, we present evidence for a novel and powerful social factor perpetuating the underrepresentation of women and girls: stereotypes about the culture of these fields. We begin by differentiating *stereotypes about the culture* from the large body of useful work on *stereotype threat*. Then, we describe the content of students' stereotypes about the culture of computer science and engineering and document their pervasiveness in the minds of American students. Third, we describe three ways that these stereotypes about the culture are transmitted: through environments, the media, and the people in the fields, and why these stereotypes are a more powerful deterrent for girls than boys. Fourth, we present empirical evidence that these stereotypes cause gender disparities in interest in entering computer science and engineering not only in college but earlier in the pipeline, including among high-school students. Finally, we show that these stereotypes, while powerful, are nonetheless highly malleable and that changing them encourages girls and women to enter these fields (without dissuading boys and men). Note that research on different populations, at different ages, and asking different questions (e.g., why are women underrepresented in the STEM workforce?) may discover different factors responsible (e.g., Eagly and Carli, 2007; Hewlett et al., 2008; Ceci et al., 2009, 2014). Our argument is that stereotypes of the field act as educational gatekeepers, constraining who enters these fields, and that interventions to broaden the cultural representation of these fields can help to draw more diversity into them.

DUAL STEREOTYPES AND GENDER DISPARITIES

By elementary school, indeed as early as second grade, girls already hold stereotypes associating boys with math (Cvencek et al., 2011). A large body of research on stereotype threat has investigated the consequences of concerns about being judged through the lens of a negative stereotype (Steele, 1997). This research has shown that negative stereotypes about girls' math abilities hinder their math performance (Huguet and Regner, 2007; see also Spencer et al., 1999; Master et al., 2014). There are three ways in which the

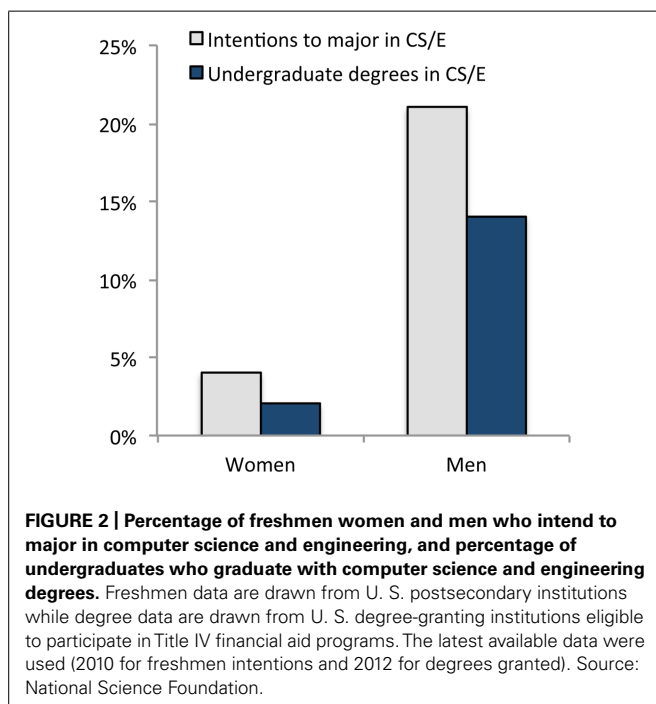
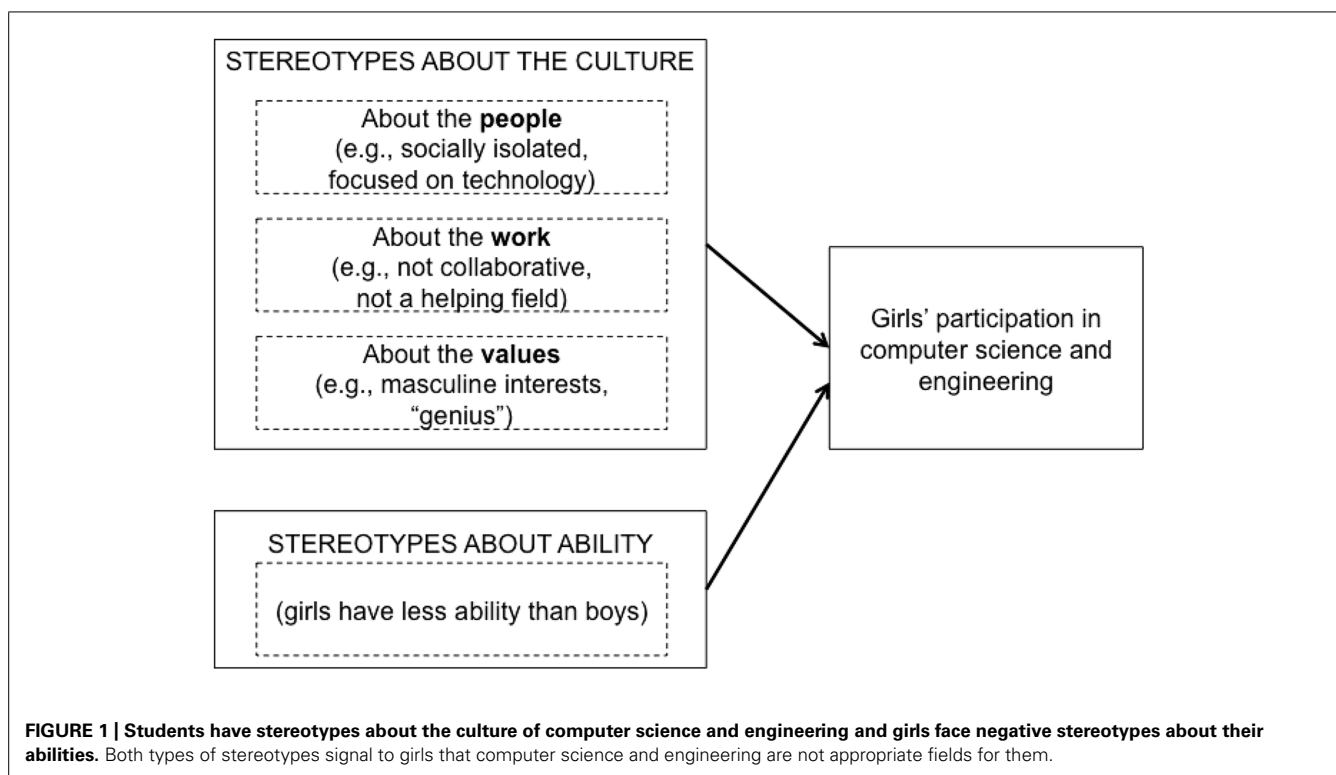
work presented here differs from this established work on stereotype threat. First, work on stereotype threat focuses on stereotypes about girls and women whereas our focus is on students' stereotypes about the culture of the fields. Both sets of stereotypes – stereotypes about girls themselves and girls' stereotypes about the culture – may be operating simultaneously to make girls feel like they do not belong in computer science and engineering (see **Figure 1**).

Second, whereas stereotypes about girls' math abilities ("girls are not good at math") are negative, we investigate stereotypes that are not always negative (Cheryan et al., 2009). Indeed, stereotypes of computer scientists and engineers can be a source of pride, identification, and belonging for some in the field (e.g., the Geek Girl Dinners organization). This lack of objective negativity can make diversifying how the fields are portrayed more challenging because these stereotypes might not be seen as problematic, even in the face of evidence that many students find them incompatible with how they see themselves. Third, stereotype threat effects are most prominent among women who are already highly identified and invested with STEM, such as STEM majors (Schmader et al., 2008). In contrast, we suggest that stereotypes about the culture preclude many girls from even considering the fields in the first place, and thus deter a larger number of girls from STEM.

THE ROLE OF STEREOTYPES EARLY IN THE PIPELINE

At what juncture in the pipeline are girls and women opting out of computer science and engineering? Although many highly qualified women *leave* these fields (Hewlett et al., 2008), a much larger contributor to the gender gap is that girls are much less likely than boys to *choose them in the first place* (de Cohen and Deterding, 2009). Among high-school students, girls are significantly less likely to take a computer programming class than boys (Shashaani, 1994; Schumacher and Morahan-Martin, 2001), less likely to take the computer science Advanced Placement (AP) test than boys (College Board, 2013), and express less interest in pursuing careers in computer science and engineering than boys (Weisgram and Bigler, 2006). By the time they enter college, men are already more than four times more likely to have an intention to major in computer science and engineering than women (National Science Foundation, 2012). Even if every woman who intended to major in computer science and engineering upon entering college was retained in these fields, men would still be significantly more likely to earn a computer science and engineering degree than women (see **Figure 2**).

Though there is debate on whether biological factors play a role in women's underrepresentation in STEM (Benbow and Stanley, 1982; Spelke, 2005), differences in interest in computer science and engineering between boys and girls are evident even among students with the highest math abilities. Among the top scorers on a standardized math test administered in the 10th grade, girls relative to boys were more likely to choose social science and health-related majors in college over majors in computer science, engineering, physical sciences, and mathematics (Perez-Felkner et al., 2012). Computer science and engineering are missing out on an entire population of talented girls who are not entering these fields to begin with.



Intervening early in the pipeline (i.e., before college) is important to remedying disparities in computer science and engineering. Societal change will occur only to the extent that the students who are initially drawn into the field are able to remain in it, thus research on retention is, of course, important and useful.

However, closing the gender gap in computer science and engineering participation will initially require convincing more girls to join these fields. As we will argue, stereotypes of the culture affect girls' choices and interest, and do so early in the pipeline.

WHAT IS THE CONTENT OF COMPUTER SCIENCE AND ENGINEERING STEREOTYPES?

When students think of computer scientists, they often think of "geeky" guys who are socially awkward and infatuated with technology (Mercier et al., 2006; Rommes et al., 2007). The work in computer science and engineering is seen as isolating and relatively dissociated from communal goals such as helping society and working with others (Hoh, 2009; Diekmann et al., 2010). Computer scientists and engineers are also perceived as having masculine interests (e.g., playing video games; Cheryan et al., 2011b), and their faculty are more likely than faculty in other fields (e.g., biology, psychology) to believe that an inborn brilliance or genius is required to be successful (Leslie et al., 2015). Of course, many computer scientists do not fit these stereotypes (Borg, 1999). But people's beliefs have a tremendous power to determine their attitudes, behaviors, and choices, even if these perceptions are completely disconnected from reality (Hasdorf and Cantril, 1954; Ross and Nisbett, 1991). In the words of one female computer science major at Carnegie Mellon, "Oh my gosh, this isn't for me.'... I don't dream in code like they do" (Margolis et al., 2000, p. 17).

Computer science and engineering stereotypes are pervasive in modern American society and even young students frequently endorse them. When high-school students described computer scientists, the majority (84%) mentioned at least one measurable stereotype, including being technically oriented, singularly

focused on technology, socially awkward, masculine, intelligent, or having particular physical traits such as glasses or pale skin (Master et al., unpublished). College students reported similar stereotypes, with 67% mentioning at least one of these stereotypes about computer scientists (Cheryan et al., 2013b). College students were also less likely to believe that computer science and engineering were fields that could be used to help people or work with others than fields such as medicine and law (Hoh, 2009; Diekmann et al., 2010).

In today's society, computer science and engineering stereotypes are perceived as incompatible with qualities that are valued in women, such as being feminine, people-oriented, and modest about one's abilities (Diekmann et al., 2011; Cheryan, 2012; Leslie et al., 2015). As a result, when these stereotypes are prominent, girls and women, but not boys and men, believe that they are dissimilar from those in the field and report a lower "sense of belonging," or feeling of fit with the culture of the field (Cheryan et al., 2009; Master et al., under review). The less that students feel a sense of belonging in a field, the less likely they are to pursue that field (Good et al., 2012; Smith et al., 2013; Master et al., under review). Changing these stereotypes may allow more girls and women to believe they are welcome in computer science and engineering.

TRANSMISSION CHANNELS FOR STEREOTYPES ABOUT COMPUTER SCIENCE AND ENGINEERING

Below we review three ways in which students may be exposed to computer science and engineering stereotypes – through media, people in the fields, and environments. Because computer science and engineering are not mandatory and often not even offered in U.S. high schools (Stephenson et al., 2005), many students do not have direct experience with these fields. As a result, students often rely on cultural stereotypes about computer scientists and engineers for knowledge about these fields. However, these stereotype transmission channels have an upside as well: they are particularly well-suited mechanisms of cultural change if interventions are designed appropriately.

STEREOTYPES TRANSMITTED THROUGH THE MEDIA

Popular movies and television shows like *Real Genius*, *The Big Bang Theory*, and *Silicon Valley* depict computer scientists and engineers as mostly White (and more recently Asian) males, socially unskilled, and singularly obsessed with technology. Similarly, portrayals of technology companies in popular newspapers and books often depict the "startup culture" that infuses some technology and engineering jobs (e.g., Guo, 2014; Miller, 2014). This is unfortunate because in reality such portrayals depict at best only a small percentage of the jobs in computer science and engineering (Bureau of Labor Statistics, 2014). Yet high-school students report that their ideas about what scientists are like are influenced more by the media than by any other source (Steinke et al., 2007). Even brief exposures to television portrayals can influence attitudes toward the group portrayed (Weisbuch et al., 2009).

To examine the extent to which exposure to stereotypical and non-stereotypical media representations influence women's interest in computer science, women undergraduates read one of two fabricated newspaper articles. One article stated that computer

scientists fit the current stereotypes, while the other stated that computer scientists were diversifying and no longer fit the stereotypes. Women who read the stereotypical article expressed less interest in majoring in computer science than women who read the non-stereotypical article. Furthermore, women who read the non-stereotypical article were significantly more interested in computer science than women who read no article (Cheryan et al., 2013b). Changing the images of computer science and engineering in the media may increase women's interest in these fields.

STEREOTYPES TRANSMITTED BY NARROW CHARACTERIZATIONS OF PEOPLE IN THE FIELDS

Faculty, students, and industry professionals embody certain characteristics, habits, and belief systems that can signal what is normative and valued in the field. For instance, the National Academy of Engineering's engineeryourlife.org website features a female computer engineer who appears to fit the definition of a role model for girls: she is successful, competent, and shares their gender (Marx et al., 2005; Stout et al., 2011). However, her profile also describes how she embodies stereotypes of computer scientists and engineers: she started programming at age 11 and works as a Star Wars video game designer. Computer scientists and engineers who embody these stereotypes may discourage women from entering these fields.

To investigate whether encountering a stereotypical computer science student can deter women, undergraduate women were brought into a room to have a conversation with a participant who was actually an actor. Three male and three female actors were used. The conversation was brief – less than 2 min on average – and consisted of the participant and the actor exchanging basic information about themselves (e.g., year, major, hobbies, favorite movie). The actor always stated that he or she was a junior and a computer science major, but half of the participants were randomly assigned to interact with an actor who fit current stereotypes in appearance and preferences (e.g., glasses, t-shirt that said "I code therefore I am," hobbies that included playing videogames) or one who did not fit these stereotypes (e.g., solid colored t-shirt, hobbies that included hanging out with friends). After the interaction was complete, participants were asked about their interest in their partner's major and then asked the same questions again 2 weeks later.

Results revealed that women who interacted with the stereotypical student were significantly less interested in majoring in computer science than those who interacted with the non-stereotypical student, and this effect was equally strong regardless of whether the actor was male or female. Moreover, negative effects of stereotypes endured for 2 weeks after the interaction (Cheryan et al., 2013a). The computer science major's gender mattered less in influencing women's interest in computer science than the extent to which he or she fit current computer science stereotypes.

Follow-up experiments (a) revealed similar effects of peer stereotypicality on anticipated success in computer science (Cheryan et al., 2011b) and also (b) investigated why people in the field who embody computer science stereotypes may be steering women away from the field. Interacting with a stereotypical computer science major reduced women's anticipated success in

computer science but did not affect men's anticipated success (Cheryan et al., 2011b). Why? Women felt less similar to the stereotypical student than to the non-stereotypical student, suggesting students may look to other characteristics besides gender when determining with whom they feel similar (see also Cheryan et al., 2011b; Meltzoff, 2013). When the people in computer science depict themselves in a manner consistent with the stereotypes, it can convey to other students that one must fit the stereotypes to be successful in these fields.

Computer scientists and engineers who depict the work in their fields as highly independent may also discourage women from entering their fields. College women who read about an entry-level scientist who spent a typical day doing independent tasks reported less positive attitudes about science careers than college women who read about an entry-level scientist who spent a typical day doing collaborative tasks (Diekmann et al., 2011). Moreover, fewer female students are present in fields whose faculty believe that success in their field requires innate brilliance, a belief that is prominent in computer science and engineering (Leslie et al., 2015). Changing stereotypes about the work being isolating and requiring an innate brilliance may draw more women into computer science and engineering.

STEREOTYPES TRANSMITTED THROUGH ENVIRONMENTS

Objects and environments are powerful because they are seen as providing clues about the dominant culture within that environment, including information about the values, beliefs, norms, and practices (Whiting, 1980; Cialdini et al., 1990; Markus and Kitayama, 1991). Environments that depict computer science and engineering as more compatible with characteristics, interests, and values associated with men and boys are likely to draw fewer girls than boys into them. However, exposing students to computer science and engineering environments that do not fit current male-oriented stereotypes may reduce gender disparities in interest in these fields.

College undergraduates who were not computer science majors (in order to focus on recruitment) entered a classroom in the computer science department at Stanford University, which was decorated in one of two ways (Cheryan et al., 2009). For half the participants, the room had objects that other undergraduates associated highly with computer science majors—Star Trek posters, science fiction books, and stacked soda cans. For the other half of participants, the room contained objects that other undergraduates did not associate with computer science majors—nature posters, neutral books, and water bottles. Women in the room that did not contain the stereotypical objects expressed significantly more interest in majoring in computer science than those in the room that did fit the stereotypes. For men, the environment did not affect their interest in computer science (Cheryan et al., 2009).

Online educational environments are becoming an increasingly important presence in students' lives as universities use them as tools for education. To test whether the design of virtual classrooms influences educational outcomes, undergraduates virtually entered two classrooms in Second Life, an online 3D interactive virtual environment. Both were introductory computer science classrooms, but one contained stereotypical objects while the other

contained non-stereotypical objects. Whereas only 18% of women chose to take the course in the stereotypical classroom, more than half of men (60%) chose that classroom. Furthermore, women expected to perform worse in the class with the stereotypical objects than men, but in the non-stereotypical classroom, women's expectations rose, so that women and men expected to do equally well (Cheryan et al., 2011a).

Why did the stereotypical environment deter women more than men? Women reported a lower sense of ambient belonging in the stereotypical environment, or sense of fit with the material components and with the people assumed to inhabit the environment. In contrast, men reported an equal, and sometimes greater, sense of ambient belonging in the stereotypical environment than the non-stereotypical environment (Cheryan et al., 2009, 2011a). Women were less likely than men to associate themselves with the stereotypical objects, and the more that women perceived the stereotypical environment as masculine, the less interest they expressed in being in that environment (Cheryan et al., 2009).

Earlier in the pipeline, high-school students also show similar effects on their interest in taking introductory computer science in a classroom that fits or does not fit current computer science stereotypes (Master et al., under review). Girls were more likely to choose a non-stereotypical classroom (68% of girls) over a stereotypical one, while boys showed no preference for a non-stereotypical classroom (48%). Moreover, girls' baseline interest in a computer science course in which the classroom was not described was no different from their interest in a stereotypical course (and both were lower than the non-stereotypical course), suggesting that a stereotypical classroom was consistent with girls' default assumptions about introductory computer science courses. However, a non-stereotypical environment provided a new image of computer science and increased their interest over baseline. Like their college counterparts, high-school girls felt a lower sense of fit with current computer science stereotypes than did boys. The less that girls reported a sense of fit with the current stereotypes, the more likely they were to be deterred from a stereotypical (but not a non-stereotypical) computer science environment (Master et al., under review). The observed variability between girls is striking and suggests that current stereotypes should be diversified rather than eliminated, a point we discuss in more detail in the next section.

Thus, women and girls may be choosing fields other than computer science and engineering in part due to the constraining power of current stereotypes that portray the culture of the field in a manner that is incompatible with the way that women see themselves. When the constraint is lifted by presenting a non-stereotypical image, girls' sense of belonging and interest in the field can increase, without reducing boys' interest.

THE IMPORTANCE OF VARIABILITY AND DIVERSIFYING PORTRAYALS OF COMPUTER SCIENCE AND ENGINEERING

In all studies investigating effects of stereotypes, there is a sizable portion of students who may be drawn to these fields *because* of these stereotypes. In the studies on environments, some women (typically 20–25% of women) preferred the stereotypical

environment over the non-stereotypical environment. Rather than attempting to overhaul current stereotypes, which may deter some men and women, a more effective strategy may be to diversify the image of these fields so that students interested in these fields do not think that they must fit a specific mold to be a successful computer scientist or engineer.

Diversifying the image of computer scientists and engineers may not only attract more women to the field, but also make some men feel more welcome in these fields. Indeed, in the studies on environments, some men (typically 25–30% of men) preferred the non-stereotypical environment over the stereotypical environment. In addition, many men also highly value opportunities to work with and help others (Diekmann et al., 2011). Attracting more non-stereotypical men to the field is a way to further stretch stereotypes and diversify a field (Drury et al., 2011).

A question that our readers may have is whether it is fair to present girls with a non-stereotypical image of the fields of computer science and engineering if they will then enter these fields and be unprepared for the male-oriented culture that they encounter there. We believe it is necessary and useful to prepare girls and women for the obstacles they may encounter in male-dominated fields and how to overcome them. We also believe that the cultures of these fields should be changed to be more welcoming of a diversity of people. However, our viewpoint is that girls are currently exposed to an unrealistic image of these fields that depicts all computer science and engineering cultures as fitting a narrow profile. A broader image that shows many different types of people and working environments in computer science and engineering actually represents a more realistic portrayal. Furthermore, once we start the process of welcoming more women and girls into these fields, the process of culture change will likely build on itself and contribute to further improving the actual and perceived culture of these fields for women.

The computer science departments at Carnegie Mellon and Harvey Mudd provide two real-world examples of the power of changing cultural stereotypes to reduce gender disparities in participation. Both increased the proportion of women majoring in computer science from ~10 to 40% in 5 years (Margolis and Fisher, 2002; Hafner, 2012). In addition to structural changes (e.g., changes in recruiting procedures), both programs changed stereotypes of computer science by using diverse role models, exposing students to a wide range of applications of computer science, and revamping their introductory course so that it was no longer seen as a field only for “geeky know-it-alls” (Margolis and Fisher, 2002; Hafner, 2012). These examples show that efforts to reduce gender disparities in computer science and engineering benefit from actively working to change the culture of these fields, so that they are seen as places where *all* students are valued and have the potential to be successful.

CONCLUSION: CONTRIBUTIONS TO THEORY AND PRACTICE

Why are girls, even those who grew up with technology in their homes and took advanced math classes in high school, less likely than boys to pursue computer science and engineering? Our central thesis is that girls’ underrepresentation in these fields is not due to their intractable lack of interest in choosing these fields.

Instead, we argue that women’s choices are constrained by societal factors, particularly their stereotypes about of the kind of *people*, the *work involved*, and the *values* of these fields (see **Figure 1**). These perceptions, even if they are not accurate, shape the academic choices that girls make by communicating to them where they belong.

We also argue that we can change students’ stereotypes of the culture using relatively simple interventions to environments, the media, and by diversifying the type of people representing these fields. Rather than “de-geeking” the fields, a more successful approach involves creating inclusive cultures so that those who are considering these fields do not necessarily have to embody the stereotypes to believe that they fit there. One concrete way to create inclusive cultures is to consider who is selected to represent the field (e.g., who teaches the introductory courses) and what messages he or she signals about the kind of student who belongs in the field. If all representatives are similar to one another, it can signal that one has to fit that mold in order to be successful in that environment. If there is diversity in who is presented, it sends the message that a variety of people can be successful. Physical spaces are another effective way to signal who belongs. We have shown that it is possible and feasible to create physical spaces within the larger environment that allow both men and women to feel welcome there. Finally, it is also important to change the stories told in the media about these fields and who is found in them.

The main message of this research is that variability is key. Instead of portraying computer science and engineering as narrow fields that are easily stereotyped—and which therefore steer a large number of students away because they “do not belong”—we can alter how the culture of these fields is represented in the minds of youth. By broadening the mental picture of what it means to be a computer scientist or engineer, we may not only attract more women to these fields, but also be more accurate about what computer science and engineering are like and what they have the potential to become.

ACKNOWLEDGMENTS

This work was supported by grants from the National Science Foundation, DRL-1420351 to SC and SMA-0835854 to ANM.

REFERENCES

- American Bar Association. (2014). *JD and LLB Degrees Awarded, 1981-2011*. Available at: http://www.americanbar.org/content/dam/aba/administrative/legal_education_and_admissions_to_the_bar/statistics/jd_llb_degrees_awarded.auth-checkdam.pdf
- Benbow, C. P., and Stanley, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: a longitudinal perspective. *Am. Educ. Res. J.* 19, 598–622. doi: 10.3102/00028312019004598
- Borg, A. (1999). What draws women to and keeps women in computing? *Annu. N. Y. Acad. Sci.* 869, 102–105. doi: 10.1111/j.1749-6632.1999.tb08362.x
- Bureau of Labor Statistics. (2014). *Computer and Information Research Scientist*. Available at: <http://www.bls.gov/ooh/computer-and-information-technology/computer-and-information-research-scientists.htm>
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women’s underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412

- Cheryan, S. (2012). Understanding the paradox in math-related fields: why do some gender gaps remain while others do not? *Sex Roles* 66, 184–190. doi: 10.1007/s11199-011-0060-z
- Cheryan, S., Drury, B., and Vichayapai, M. (2013a). Enduring influence of stereotypical computer science role models on women's academic aspirations. *Psychol. Women Q.* 37, 72–79. doi: 10.1177/0361684312459328
- Cheryan, S., Plaut, V. C., Handron, C., and Hudson, L. (2013b). The stereotypical computer scientist: gendered media representations as a barrier to inclusion for women. *Sex Roles* 69, 58–71. doi: 10.1007/s11199-013-0296-x
- Cheryan, S., Meltzoff, A. N., and Kim, S. (2011a). Classrooms matter: the design of virtual classrooms influences gender disparities in computer science classes. *Comput. Educ.* 57, 1825–1835. doi: 10.1016/j.compedu.2011.02.004
- Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B., and Kim, S. (2011b). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Soc. Psychol. Pers. Sci.* 2, 656–664. doi: 10.1177/1948550611405218
- Cheryan, S., Plaut, V. C., Davies, P. G., and Steele, C. M. (2009). Ambient belonging: how stereotypical cues impact gender participation in computer science. *J. Pers. Soc. Psychol.* 97, 1045–1060. doi: 10.1037/a0016239
- Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* 58, 1015–1026. doi: 10.1037/0022-3514.58.6.1015
- College Board. (2013). *AP Program Participation and Performance Data 2013*. Available at: <http://research.collegeboard.org/programs/ap/data/participation/2013>
- Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. Soc.* 106, 1691–1730. doi: 10.1086/321299
- Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x
- Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: the stereotype inoculation model. *Psychol. Inq.* 22, 231–246. doi: 10.1080/1047840X.2011.607313
- de Cohen, C. C., and Deterding, N. (2009). Widening the net: national estimates of gender disparities in engineering. *J. Eng. Educ.* 98, 211–226. doi: 10.1002/j.2168-9830.2009.tb01020.x
- Diekmann, A. B., Brown, E., Johnston, A., and Clark, E. (2010). Seeking congruity between goals and roles: a new look at why women opt out of STEM careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342
- Diekmann, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., and Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to STEM careers: evidence for a goal congruity perspective. *J. Pers. Soc. Psychol.* 101, 902–918. doi: 10.1037/a0025199
- Drury, B. J., Siy, J. O., and Cheryan, S. (2011). When do female role models benefit women? The importance of differentiating recruitment from retention in STEM. *Psychol. Inq.* 22, 265–269. doi: 10.1080/1047840X.2011.620935
- Eagly, A. H., and Carli, L. L. (2007). *Through the Labyrinth: The Truth about How Women Become Leaders*. Boston, MA: Harvard Business Press.
- Eccles, J. S., Jacobs, J. E., and Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *J. Soc. Issues* 46, 183–201. doi: 10.1111/j.1540-4560.1990.tb01929.x
- Ehrlinger, J., and Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *J. Pers. Soc. Psychol.* 84, 5–17. doi: 10.1037/0022-3514.84.1.5
- Gelernter, D. (1999). Women and science at Yale. *The Weekly Standard*. Available at: <http://www.weeklystandard.com/Content/Protected/Articles/000/000/010/277/gnvgd.asp>
- Good, C., Rattan, A., and Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *J. Pers. Soc. Psychol.* 102, 700–717. doi: 10.1037/a0026659
- Guo, P. (2014). Silent technical privilege. *Slate*. Available at: http://www.slate.com/articles/technology/technology/2014/01/programmer_privilege_as_an_asian_male_computer_science_major_everyone_gave.html
- Hafner, K. (2012). Giving women the access code. *The New York Times*. Available at: <http://www.nytimes.com/2012/04/03/science/giving-women-the-access-code.html?pagewanted=all&r=0>
- Hasdorf, A. H., and Cantrell, H. (1954). They saw a game. *J. Abnorm. Soc. Psychol.* 49, 129–134. doi: 10.1037/h0057880
- Hewlett, S. A., Luce, C. B., Servon, L. J., Sherbin, L., Shiller, P., Sosnovich, E., et al. (2008). *The Athena Factor: Reversing the Brain Drain in Science, Engineering, and Technology*. Harvard Business Review Research Report. Boston, MA.
- Hoh, Y. (2009). Using notable women in environmental engineering to dispel misperceptions of engineers. *Int. J. Environ. Sci. Educ.* 4, 117–131.
- Huguet, P., and Regner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *J. Educ. Psychol.* 99, 545–560. doi: 10.1037/0022-0663.99.3.545
- Jolliffe, L., Leadley, J., Coakley, E., and Sloane, R. A. (2012). *Women in U. S. Academic Medicine and Science: Statistics and Benchmarking Report 2011–2012*. Washington, DC: Association of American Medical Colleges. Available at: <https://www.aamc.org/download/415556/data/2011-2012wimsstatsreport.pdf>
- Kalwarski, T., Mosher, D., Paskin, J., and Rosato, D. (2007). Best jobs in America. *CNNMoney.com*. Available at: <http://money.cnn.com/magazines/moneymag/bestjobs/index.html>
- Leslie, S. J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375
- Margolis, J., and Fisher, A. (2002). Unlocking the clubhouse: women in computing. *Psychol. Women Q.* 26, 381–382.
- Margolis, J., Fisher, A., and Miller, F. (2000). Anatomy of interest: women in undergraduate computer science. *Womens Stud. Q. Spec. Issue Women Sci.* 28, 104–126.
- Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224
- Martincic, C. J., and Bhatnagar, N. (2012). Will Computer Engineer Barbie® impact young women's career choices? *Inform. Syst. Educ. J.* 10, 4–14.
- Marx, D. M., Stapel, D. A., and Muller, D. (2005). We can do it: the interplay of construal orientation and social comparisons under threat. *J. Pers. Soc. Psychol.* 88, 432–446. doi: 10.1037/0022-3514.88.3.432
- Master, A., Cheryan, S., and Meltzoff, A. N. (2014). Reducing adolescent girls' concerns about STEM stereotypes: when do female teachers matter? *Int. Rev. Soc. Psychol.* 27, 79–102.
- Meltzoff, A. N. (2013). “Origins of social cognition: bidirectional self-other mapping and the ‘Like-Me’ hypothesis,” in *Navigating the Social World: What Infants, Children, and Other Species Can Teach us*, eds M. Banaji and S. Gelman (New York, NY: Oxford University Press).
- Mercier, E. M., Barron, B., and O'Connor, K. M. (2006). Images of self and others as computer users: the role of gender and experience. *J. Comput. Assist. Learn.* 22, 335–348. doi: 10.1111/j.1365-2729.2006.00182.x
- Miller, C. C. (2014). Technology's man problem. *The New York Times*. Available at: <http://www.nytimes.com>
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16464–16479. doi: 10.1073/pnas.1211286109
- Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat: how situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x
- National Science Foundation. (2012). *Appendix Table 2-12. Freshmen Intending S&E Major, by Field, Sex, and Race/Ethnicity: 1995–2010*. Available at: <http://www.nsf.gov/statistics/seind12/append/c2/at02-12.xls>
- National Science Foundation. (2013). *TABLE 5-2. Bachelor's Degrees Awarded, by Field and Sex: 2002–2012*. Available at: http://www.nsf.gov/statistics/wmpd/2013/pdf/tab5-2_updated_2014_05.pdf
- Perez-Felkner, L., McDonald, S., Scheider, B., and Grogan, E. (2012). Female and male adolescents' subjective orientations to mathematics and the influence of those orientations on postsecondary majors. *Dev. Psychol.* 48, 1658–1673. doi: 10.1037/a0027020
- Rommers, E., Overbeek, G., Scholte, R., Engels, R., and De Kemp, R. (2007). ‘I'm not interested in computers’: gender-based occupational choices of adolescents. *Inform. Commun. Soc.* 10, 299–319. doi: 10.1080/13691180701409838
- Ross, L., and Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology*. Philadelphia, PA: Temple University Press.
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *J. Pers. Soc. Psychol.* 74, 629–645. doi: 10.1037/0022-3514.74.3.629

- Sadker, M., and Sadker, D. (1994). *Failing at Fairness: How America's Schools Cheat Girls*. New York: Scribner.
- Schmader, T., Johns, M., and Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychol. Rev.* 115, 336–356. doi: 10.1037/0033-295X.115.2.336
- Schumacher, P., and Morahan-Martin, J. (2001). Gender, internet and computer attitudes and experiences. *Comput. Hum. Behav.* 17, 95–110. doi: 10.1016/S0747-5632(00)00032-7
- Shashaani, L. (1994). Gender differences in computer experience and its influence on computer attitudes. *J. Educ. Comput. Res.* 11, 347–367. doi: 10.2190/64MD-HTKW-PDXV-RD62
- Smith, J. L., Lewis, K. L., Hawthorne, L., and Hodges, S. D. (2013). When trying hard isn't natural: women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Pers. Soc. Psychol. Bull.* 39, 131–143. doi: 10.1177/0146167212468332
- Soper, T. (2014). Analysis: the exploding demand for computer science education, and why America needs to keep up. *Geekwire*. Available at: <http://www.geekwire.com/2014/analysis-examining-computer-science-education-explosion/>
- Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *Am. Psychol.* 60, 950–958. doi: 10.1037/0003-066X.60.9.950
- Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373
- Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037/0003-066X.52.6.613
- Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., et al. (2007). Assessing media influences on middle school-aged children's perceptions of women in science using the Draw-A-Scientist Test (DAST). *Sci. Commun.* 29, 35–64. doi: 10.1177/1075547007306508
- Stephenson, C., Gal-Ezer, J., Haberman, B., and Verno, A. (2005). *The New Educational Imperative: Improving High School Computer Science Education*. Final Report of the Computer Science Teachers Association, Curriculum Improvement Task Force. New York, NY: Computer Science Teachers Association.
- Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. (2011). STEMing the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 100, 255–270. doi: 10.1037/a0021385
- Weisbuch, M., Pauker, K., and Ambady, N. (2009). The subtle transmission of race bias via televised nonverbal behavior. *Science* 326, 1711–1714. doi: 10.1126/science.1178358
- Weisgram, E. S., and Bigler, R. S. (2006). Girls and science careers: the role of altruistic values and attitudes about scientific tasks. *J. Appl. Dev. Psychol.* 27, 326–348. doi: 10.1016/j.appdev.2006.04.004
- Whiting, B. B. (1980). Culture and social behavior: a model for the development of social behavior. *Ethos* 8, 95–116. doi: 10.1525/eth.1980.8.2.02a00010
- Zimmerman, A. (2010). Revenge of the nerds: How Barbie got her geek on. *The Wall Street Journal*. Available at: <http://www.wsj.com/articles/SB10001424052702304198004575171791681002592>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 October 2014; paper pending published: 01 December 2014; accepted: 10 January 2015; published online: 11 February 2015.

Citation: Cheryan S, Master A and Meltzoff AN (2015) Cultural Stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes. *Front. Psychol.* 6:49. doi: 10.3389/fpsyg.2015.00049

This article was submitted to *Developmental Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Cheryan, Master and Meltzoff. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

On the gender–science stereotypes held by scientists: explicit accord with gender-ratios, implicit accord with scientific identity

Frederick L. Smyth^{1*} and Brian A. Nosek^{1,2}

¹ Department of Psychology, College and Graduate School of Arts and Sciences, Charlottesville, VA, USA, ² The Center for Open Science, University of Virginia, Charlottesville, VA, USA

OPEN ACCESS

Edited by:

Stephen J. Ceci,
Cornell University, USA

Reviewed by:

David I. Miller,
Northwestern University, USA
Meredith Meyer,
Otterbein University, USA
Sarah-Jane Leslie,
Princeton University, USA

*Correspondence:

Frederick L. Smyth,
Department of Psychology, College
and Graduate School of Arts and
Sciences, 485 McCormick Road, 102
Gilmer Hall, PO Box 400400,
Charlottesville, VA, USA
fsmyth@virginia.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 November 2014

Accepted: 23 March 2015

Published: 27 April 2015

Citation:

Smyth FL and Nosek BA (2015) On
the gender–science stereotypes held
by scientists: explicit accord with
gender-ratios, implicit accord with
scientific identity.
Front. Psychol. 6:415.
doi: 10.3389/fpsyg.2015.00415

Women's representation in science has changed substantially, but unevenly, over the past 40 years. In health and biological sciences, for example, women's representation among U.S. scientists is now on par with or greater than men's, while in physical sciences and engineering they remain a clear minority. We investigated whether variation in proportions of women in scientific disciplines is related to differing levels of male-favoring explicit or implicit stereotypes held by students and scientists in each discipline. We hypothesized that science-is-male stereotypes would be weaker in disciplines where women are better represented. This prediction was tested with a sample of 176,935 college-educated participants (70% female), including thousands of engineers, physicians, and scientists. The prediction was supported for the explicit stereotype, but not for the implicit stereotype. Implicit stereotype strength did not correspond with disciplines' gender ratios, but, rather, correlated with two indicators of disciplines' scientific intensity, positively for men and negatively for women. From age 18 on, women who majored or worked in disciplines perceived as more scientific had substantially weaker science-is-male stereotypes than did men in the same disciplines, with gender differences larger than 0.8 standard deviations in the most scientifically-perceived disciplines. Further, particularly for women, differences in the strength of implicit stereotypes across scientific disciplines corresponded with the strength of scientific values held by women in the disciplines. These results are discussed in the context of dual process theory of mental operation and balanced identity theory. The findings point to the need for longitudinal study of the factors' affecting development of adults' and, especially, children's implicit gender stereotypes and scientific identity.

Keywords: diversity, gender, science education, science workforce, stereotypes

Introduction

In 1966 just 7% of undergraduate women took their bachelor's degrees in STEM (Science, Technology, Engineering, Math, excluding health and social science; National Science Foundation, National Center for Science and Engineering Statistics, 2011a, Table 9). More than four decades later, despite passage in 1972 of Title IX of the Civil Rights Act and an ensuing great expansion of higher education opportunities for women, the

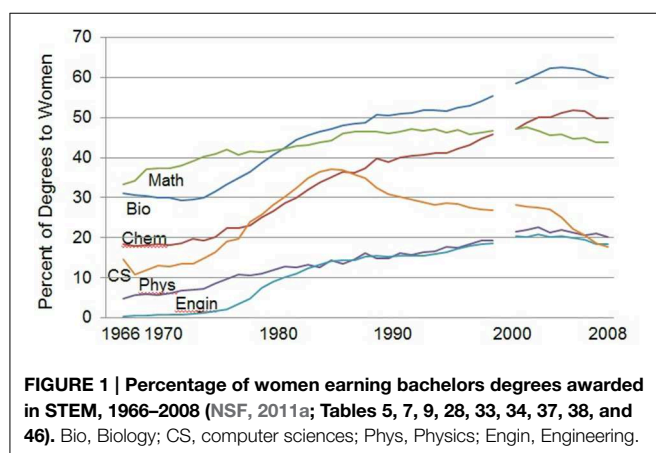
figure moved only to 10% (National Science Foundation, National Center for Science and Engineering Statistics, 2011a, data for 2008). This is slightly down from the high water mark of 12%, first reached during the mid-1980s and achieved again in 2000 and 2003. Meanwhile, men's likelihood of majoring in STEM disciplines decreased, from 29% in 1966 to 23% in 2008 (National Science Foundation, National Center for Science and Engineering Statistics, 2011a, Table 7). These trends—slow, halting progress into STEM for women and declining interest for men—may explain why leaders in STEM fields are concerned with recruitment and retention of everyone, regardless of sex, but also draw attention to the persisting sex difference in pursuit of STEM-related careers. Even with men's sagging interest, in 2008 they were still more than twice as likely as women to pursue and earn an undergraduate STEM degree.

Eliminating the apparent ceiling on women's STEM interest has long been a national priority, its causes and possible remedies the focus of extensive research and debate (e.g., Gallagher and Kaufman, 2005; Summers, 2005; Ceci and Williams, 2007, 2010, 2011; Halpern et al., 2007; National Academies of Science, 2007). Increasing attention has been paid to the variation in women's representation across different STEM domains (e.g., greater in biology than in engineering; National Science Foundation, National Center for Science and Engineering Statistics, 2011a) as a clue to understanding the causes of their underrepresentation (Su et al., 2009; Robertson et al., 2010; Cheryan, 2012).

Variation in STEM Gender Ratios

Undergraduate Degrees

Beneath the overall sex difference in STEM pursuit there is wide variation in STEM gender ratios across disciplines. **Figure 1** plots the percentage of women earning the bachelor's degrees awarded in various major STEM fields from 1966 to 2008. By 2008 women earned between 44 and 60% of the degrees in biology, chemistry and mathematics, but only about 20% in physics, engineering, and computer science (National Science Foundation, National Center for Science and Engineering Statistics, 2011a).



Occupations

Variation in women's STEM representation is also apparent among practicing scientists. In 2006, among employed U.S. scientists (defined by NSF as anyone in a STEM field with at least a bachelor's degree; National Science Foundation, 2011b) women constituted nearly two-thirds (65%) of those in health sciences, 50% of those in biological sciences, 27% in physical sciences (including 33% in chemistry, 21% in earth sciences, and 16% in physics/astronomy), and 13% in engineering. Women now constitute roughly half of all new physicians (Association of American Medical Colleges, 2013), but among employed PhD-level scientists, women comprise just 16% in the physical sciences and 9% in engineering (National Science Foundation, 2011b). Thus, in the professional scientific ranks, biological and health sciences are characterized by relatively high female-male ratios, at 1:1 or higher, while physical sciences and engineering are characterized by low female-male ratios, at 0.33:1 or lower.

The Influence of Stereotypes

Recent studies indicate that variability in women's engagement across STEM fields reflects patterns of early-developing childhood interests, and that these interests may be influenced by stereotypes and by inadequate information about the nature of opportunities in different scientific domains (Ceci and Williams, 2011; Cheryan, 2012; Eccles, 2007; Kaminski and Geisler, 2012). Although stereotypes about gender and STEM (e.g., more naturally the domain of boys and men) are now usually explicitly disavowed as a rationale for choosing courses to take, majors to enter, or persons to hire, evidence suggests that they nevertheless affect perceptions, performance and decisions, primarily without intention or awareness (e.g., Crowley et al., 2001; Moss-Racusin et al., 2012; Nosek et al., 2012; Galdi et al., 2014). Stereotypes, generally defined as associations of an attribute with members of a group, can operate at an "explicit" level, i.e., conscious perceptions of, or beliefs about, group-attribute covariation, and also at an "implicit" level, as automatic, possibly unwanted, group-attribute associations that operate outside of conscious awareness (Greenwald and Banaji, 1995; Lee et al., 1995; Gawronski and Bodenhausen, 2006; Nosek et al., 2012).

A naturalistic observational study of families at science museums seems to illustrate the independence of explicit and implicit gender-science stereotypes. Crowley et al. (2001) found that parents who brought their children to science museum exhibits spontaneously offered more explanations of phenomena to their sons than to their daughters. Here were parents that, ostensibly, were working to expose both their girls and boys to science, yet, unknowingly, were engaging more, teaching more, with the boys. If asked, these parents would doubtless say (and believe)—explicitly—that they are equally committed to the best possible science education for their child of either sex; that's why they were visiting the science museum! But Crowley et al.'s observations of many families belie differential treatment according to sex of the children, an implicit bias impossible for any of these parents to observe on their own. Such unconscious sex-differentiated patterns in adults' interaction with children in the science domain are the sort that Galdi et al. (2014) speculated

may spur the early development of girls' implicit math-gender stereotypes, which they found operating for six-year-olds prior to the emergence of explicit stereotyping. Galdi et al. (2014) experimentally exposed six-year-old boys and girls to either stereotype-congruent or incongruent images of children and math accomplishment and observed corresponding influence on the girls', but not the boys', implicit math(language)-gender stereotypes. The induced implicit stereotyping differences, in turn, were found to mediate stereotype-consistent effects on the girls' math performance, while there were no effects of explicit endorsement of math-gender stereotypes. If parents' and teachers' unconscious behaviors systematically suggest that certain STEM disciplines are more fitting for one sex than the other, the effect on children's implicit stereotypes, accumulating from a very young age, may differentially influence interest, accomplishment and persistence in particular sciences.

Relations between Gender Ratios and Stereotypes

Our data allowed investigation of whether variation in female representation across scientific disciplines is associated with differences in the strength of gender-science stereotypes, explicit and implicit, held by men and women in these fields. Current theory and evidence suggests that both explicit and implicit gender-science stereotypes should change as conditions in local environments change, including gender ratios. In perhaps the most relevant work supporting this idea, Miller et al. (in press), using the same stereotype measures we will analyze, found that average explicit and implicit stereotypes across 66 countries correlated negatively with countries' female proportion of college science majors; that is, higher proportions of women in collegiate science predicted weaker country-level science-is-male stereotyping.

Gawronski and Bodenhausen (2006), in their associative-propositional evaluation (APE) model, argue that explicit evaluations, such as stereotypes, ultimately depend on weighing the truth and importance of propositions that come to mind, e.g., "When I look around in my physics class I see mostly men." or "I'm a woman doing very well in physics." When answering a question about degree of association between science and gender, if women in physics take stock of gender ratios, they will see, on average, fewer women than will be seen by women in biology. Thus, other factors being equal, physics women should explicitly report a stronger science-is-male association than should biology women. This is consistent with Eagly and colleagues' social role theory (Eagly and Steffens, 1984; Eagly et al., 2000) which posits that varying distributions of men and women in certain activities and occupations drive explicit stereotyping and promote a cycle of corresponding skill and interest development. Consistent with such a cycle, Inzlicht and Ben-Zeev (2000), studying a sample of students from a highly selective private university (though not from any particular academic major) experimentally demonstrated a connection between gender ratios, stereotypes and academic performance. They found that women's math, but not verbal, test scores suffered as a function of increased proportion of men in the immediate environment. Diekmann

and Eagly (2000) demonstrated that explicit stereotypes are responsive to changes in women's representation; if gender distributions change, explicit stereotypes follow suit.

Implicit stereotyping, too, should vary with gender ratios. Ratliff and Nosek (2010) demonstrated that implicit associations quickly form in accord with environmental stimuli. Gawronski and Bodenhausen's (2006) APE model specifies that change in implicit evaluation will follow from either a changed structure of mental associations (actual strengthening of the associative link between a group and an attribute) or from the differential activation of existing structures (e.g., science-male associations may be more likely to be activated if one is routinely surrounded by men in a scientific context). Thus, for both men and women studying or working in scientific environments with higher male-to-female ratios, we can expect either route to result, on average, in stronger implicit science-is-male stereotyping. Miller et al. (in press), found that the negative country-level relation between female proportion of science majors and implicit science-is-male stereotyping was stronger for participants with college experience than for those without, suggesting that greater associative exposure to particular collegiate gender-science ratios may be the difference.

Results of studies of *change* of implicit stereotypes as a function of gender representation in the environment, however, are mixed. Stout et al. (2011) found no change in the math-gender stereotype evidenced by female calculus students as a function of the sex of their professor, even though strong positive change was observed for these women's implicit math attitude and identity. Consequently, Dasgupta (2011) argues that implicit STEM-gender stereotypes are rather intractable, but that their effects can be neutralized to the extent that implicit STEM identity is strong, and that the latter strengthens with increased exposure to female faculty and competent STEM peers. Smyth and colleagues (Martin et al., 2013; Smyth unpublished data), studying the math-gender stereotypes of students in university differential equations courses with female professors, found that implicit stereotype change depended on the sex of the student. Women's stereotypes, relatively weak to begin with, did not change, but men evidenced statistically significant weakening of their initially strong stereotype. Perhaps the strongest evidence for implicit stereotype change as a function of gender ratios in the local environment comes from change in a leadership-is-male stereotype (Dasgupta and Asgari, 2004). The strength of women's stereotypes changed across the first year of college depending on their degree of contact with female faculty, weakening with greater contact. Their results imply, Dasgupta and Asgari concluded, that increased female representation in local environments in previously male-dominated fields can, even in the short space of a year, "...have a powerful impact on stereotype change" (p. 654).

Greenwald and colleagues' balanced identity theory (BIT) of implicit social cognition (Greenwald et al., 2002) is grounded in principles of cognitive consistency and balanced identity (Heider, 1958). BIT anticipates that change in any one of these three sets of associations—group identity (e.g., self-female), attribute identity (e.g., self-math), or stereotype (e.g., math-male)—will induce balancing change in at least one of the others.

Thus, if women's self-identification strengthens with the male-stereotyped field of math, as found by Stout et al. (2011), BIT predicts weakening of either their implicit female gender identity or their implicit math-is-male stereotype, or both, to maintain congruence or cognitive balance among the associations. If girls' and women's science identity is strengthened by increased opportunity to interact with female peers and mentors in scientific endeavors (as suggested by Dasgupta, 2011), then according to BIT we should find weaker science-is-male implicit stereotypes among women in high female-male ratio science fields than among those in low female-male ratio science fields. In other words, if their self-science associations strengthen, and their self-female associations hold constant, then their counter-stereotypical female-science associations will strengthen—and their stereotypical male-science associations will weaken.

There is abundant evidence that implicit STEM-gender stereotypes are not monolithic, but vary predictably with interest, persistence, and performance in math and science (Nosek et al., 2002a; Kiefer and Sekaquaptewa, 2007; Nosek and Smyth, 2011; Lane et al., 2012). As predicted by BIT, men and women who identify with science differ substantially in the strength of their implicit gender stereotypes about science and math (Nosek et al., 2002a; Nosek and Smyth, 2011; Lane et al., 2012). For men, stronger science self-concepts are associated with stronger implicit science-is-male bias, while for women stronger science self-concepts coincide with weaker implicit science-is-male bias. Nosek and Smyth (2011), studying data from other online volunteers, found a trend of weaker implicit math-is-male stereotyping for both men and women who pursued graduate work in STEM compared to those with only undergraduate training (between 0.1 and 0.2 standard deviations weaker). The much larger current data set, which includes more detailed reports of degree-level and specifications of different STEM disciplines, allows testing of "dosage" effects within particular fields. Does prolonged exposure to a particular gender-ratio correlate with stereotype strength differences within given fields? That is, do scientists in low-female fields evidence stronger science-is-male stereotypes, and scientists in high-female fields evidence weaker ones, the longer they practice in that field?

Hypotheses about Variation in the Strength of Science-is-male Stereotyping

1. Stereotype differences between female and male scientists:

- (a) Implicit: Women who are strongly identified with science will have relatively weak implicit stereotypes, while men who are strongly science-identified will have relatively strong ones. This pattern, already well-established in the literature based on broad classifications like STEM vs. not-STEM, should yield large sex differences in implicit stereotyping among the scientists of our sample. Our data, which includes more detailed distinctions of academic major and profession than collected in other studies of implicit stereotyping in STEM, allows a more fine-grained replication of this well-established pattern.

- (b) Explicit: The same pattern of sex difference will hold for scientists' explicit stereotyping, weaker for women, stronger for men, if only because self-science propositions that may come to mind will differ ("I'm successful in science and a woman." vs. "I'm successful in science and a man."). However, owing to conscious endorsement of egalitarian values and social approbation against stereotyping, we expected the explicit stereotyping gender gap to be smaller than for implicit stereotypes.

2. Stereotype differences as a function of gender ratios in science environments:

- (a) Implicit: Science-is-male stereotypes will be stronger for both women and men in low-female STEM fields than in high-female fields, though sex differences should remain. For example, women in physics (low-female) should evidence stronger implicit science-is-male stereotyping than women in biology (high-female). The pattern for men in these majors should be similar, even if the means are higher than women's.
- (b) Explicit stereotypes will also reflect gender ratios (stronger stereotypic associations reported in low-female than in high-female science disciplines). Again, however, group variation on explicit stereotype means should be somewhat constrained by conscious values and motivations to respond without bias.

3. Stereotype differences as a function of "dosage" of exposure to given gender-ratios:

- (a) Implicit: Prolonged exposure to STEM environments characterized by particular gender ratios will strengthen the corresponding implicit stereotype. That is, prolonged exposure to low-female environments should strengthen science-is-male stereotyping, while prolonged exposure to high-female environments should weaken it. This hypothesis derives from theory and empirical findings concerning the formation of new implicit associations, and some cross-sectional data that are consistent with dosage effects. Nosek and Smyth (2011) found a slight diminution of stereotyping for both men and women reporting graduate study in STEM, compared to those with only undergraduate study, and Miller et al. (in press) found a college vs. no-college effect on stereotyping as a function of collegiate gender ratios in STEM, generally. Neither of these analyses distinguished between types of STEM fields. In the current data, we expect increasing stereotype-strength from age 18 to age 22 among science-declared college students in low-female fields, and a declining trend in high-female fields. Similar patterns should be found across increasing levels of training (e.g., from BS to MA to PhD).
- (b) Explicit stereotypes, when measured for scientists in a given field with roughly constant gender-ratio, will not be systematically responsive to dosage because the general propositions being consciously weighed may not change very much. That is, whether for an undergraduate woman majoring in physics or a female professor of physics, the

propositions to weigh will likely involve, on average (1) the fact of the majority-male field and (2) assessments of personal, or other women's, accomplishments. If noteworthy scientific accomplishments by women come to mind easier for women who have been in the field longer, we might expect a diminution of the explicit stereotype. But if the intractability of the gender-ratio in the field is more salient for these women, their stereotype self-reports might strengthen. To the extent that these different framings are idiosyncratically applied by individuals, systematic change across cohorts would seem unlikely.

Study Overview

We tested these predictions with over 176,000 visitors to a publicly accessible educational website (<https://implicit.harvard.edu/>) who reported U.S. citizenship, at least some college experience and an academic major. Explicit "science-is-male" stereotyping was defined simply as verbally associating the term "science" more with "male" than with "female," and implicit stereotyping by performance on an Implicit Association Test, fully described in the methods section. Our data are cross-sectional, so differences across age or level of training can only be considered suggestive of change. A particular strength of our sample is inclusion of thousands of STEM majors, whereas most other research on implicit STEM associations has been conducted with small samples.

A public website, known as *Project Implicit*, was launched in September 1998 with the purpose of heightening public awareness of implicit social cognition, and alerting participants to the possibility that mental associations outside of their awareness or control might differ from their consciously held attitudes (Nosek et al., 2002b, 2007). Visitors to the site may choose from a variety of Implicit Association Tests (IATs; Greenwald et al., 1998) and "Gender-science" has been a long-standing and popular topic (for summary of the topics and data, see Nosek et al., 2007). Though the sample is not representative of a definable population other than that of visitors to the *Project Implicit* site, it reflects greater age and education variation than the samples of college students that characterize much research. Study protocol was reviewed and authorized by the University of Virginia Institutional Review Board for Social and Behavioral Sciences.

Methods

Participants

Our analyses are of 176,935 *Project Implicit* volunteers from May 2004 through January 2012 who reported U.S. citizenship, their sex, at least some college experience, an academic major, and completed the implicit and explicit academic stereotype measures. Seventy percent of participants were female, and racial identifications, in order of proportion, were White, 81.1%; Black, 5.5%; More than one race, 5.1%; Other or unknown, 3.1%; East Asian, 2.6%; South Asian, 1.5%; American Indian/Alaska Native, 0.6%; Native Hawaiian/Pacific Islander, 0.5%. An Hispanic ethnicity was reported by 6.7%, not-Hispanic

by 89.3%, and unknown ethnicity by 4.0%¹. The median age was 25 ($M = 29$, $SD = 12$, range 17–90), and 59% were older than the typical college age range of 17–22. Fifty-one percent of participants reported some college experience short of a bachelor's degree (most of these were aged 18–22), 30% reported a bachelor's degree as their highest level, and 19% reported a graduate degree.

Explicit Science Identity: Academic Major

Participants could select from the following list of 13 categories of majors to indicate their "Major field of study or that of your highest degree." Underlined categories were coded as STEM majors in our analyses.

Biological sciences/life sciences

Business

Communications

Computer and information sciences

Education

Engineering, mathematics, or physical sciences/science technologies

Health professions or related sciences

Humanities/liberal arts

Law or legal studies

Psychology

Social sciences or history

Visual or performing arts

Other

Following other researchers in the STEM achievement literature (e.g., Elliott et al., 1995; Xie and Shauman, 2003; Smyth and McArdle, 2004; Tai et al., 2006), we defined STEM majors as those in biological, physical, computer, or health² sciences (those choosing the "Other" option, about 6% of respondents, were excluded from analyses). For the purpose of displaying the dozen categories of majors in **Figure 2A** through **Figure 4B**, from least to most science-intensive, we asked 19 psychology graduate students who were blind to our hypotheses and analytic plan to rank the categories in order of their perception of the amount of scientific course work required (Cronbach's $\alpha = 0.985$; see Supplementary Materials for details).

Explicit Science Identity: Scientific Profession

A question about occupation was added to the *Project Implicit* survey in December 2006. Our analyses focus on comparisons between respondents who identified, by both occupation and corresponding education level, as engineers,

¹Race and ethnicity proportions are based on participants since December 2006 ($N = 106,515$, or 60% of the sample) when new U.S. Office of Management and Budget reporting formats were adopted in the study. Racial/ethnic classification of earlier participants can be observed in the raw data available at <https://osf.io/y7a3n/>

²Note that some of our illustrative STEM statistics in the Introduction did not include Health sciences. Precise definitions of STEM vary, sometimes defined narrowly as natural science, technology, engineering, and mathematics fields, excluding Health and Social Sciences, and sometimes including all of these. Consistent with many researchers, our empirical analyses treat the health sciences as a STEM field.

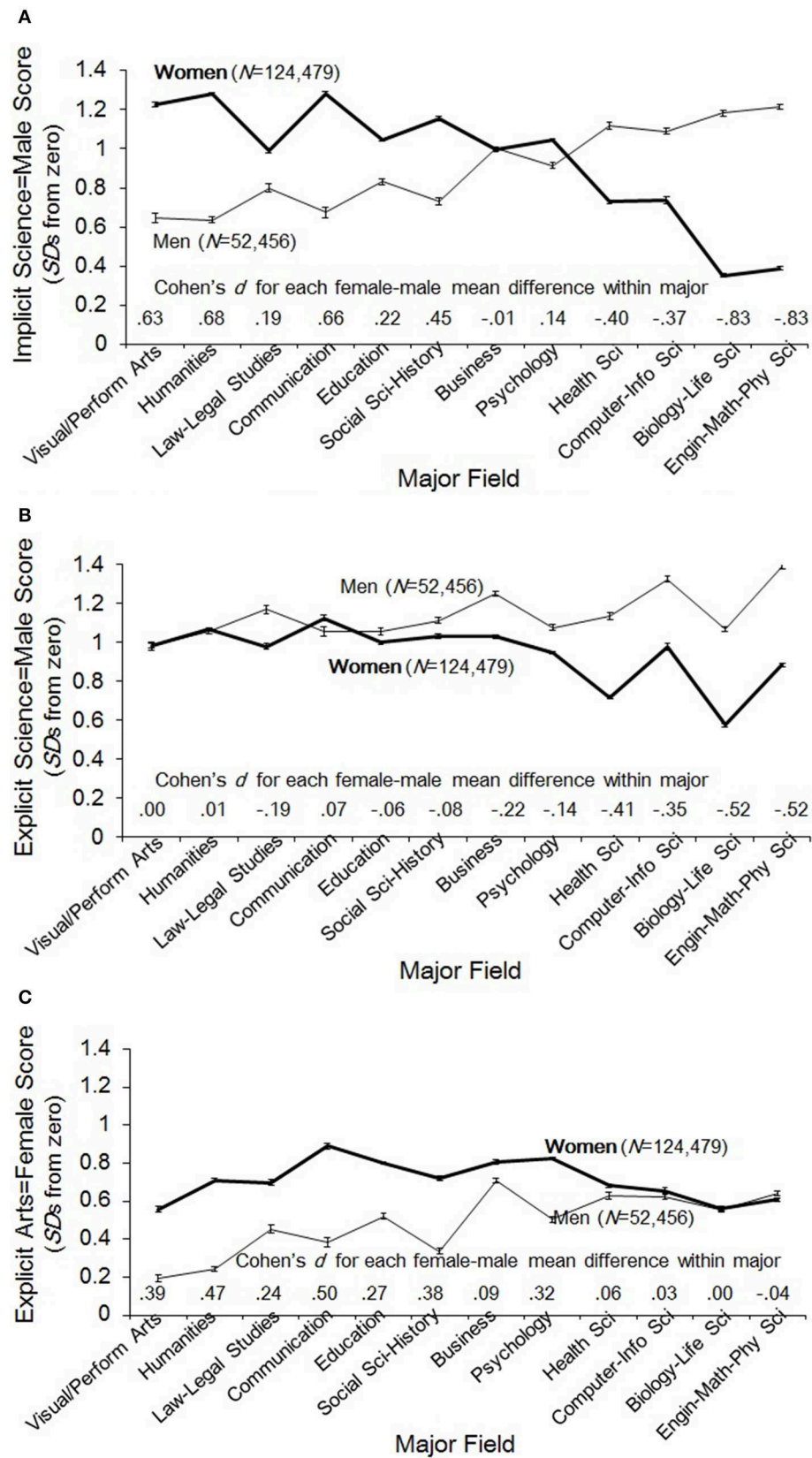


FIGURE 2 | Continued

FIGURE 2 | (A) Mean implicit science = male IAT score (± 1 se) by sex and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias. **(B)** Mean explicit science = male score (± 1 se) by sex and major field. Majors are ordered, left to right, by ratings of science

content (method described in Supplementary Material). A score of zero indicates no academic gender bias. **(C)** Mean explicit arts = female score (± 1 se) by sex and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias.

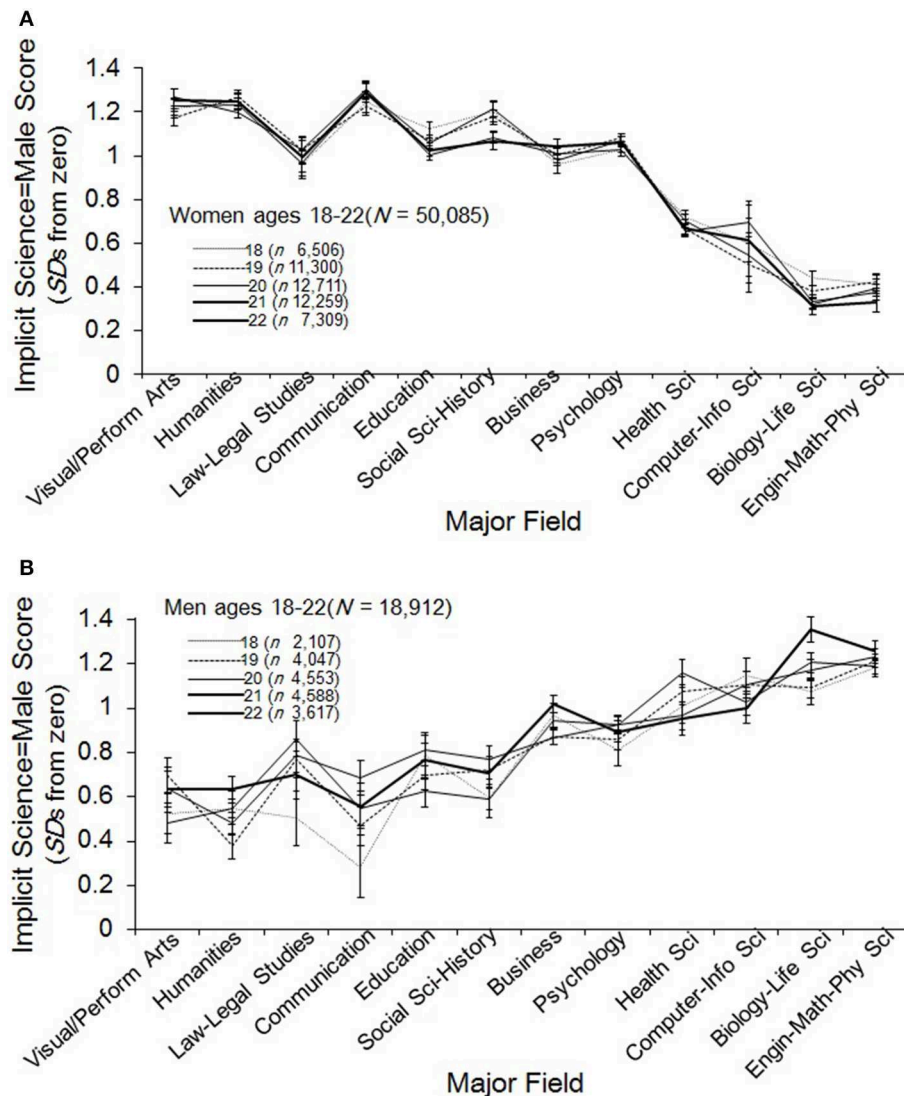


FIGURE 3 | (A) Mean implicit science = male IAT score (± 1 se) for women ages 18–22 by age and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias. **(B)** Mean

implicit science = male score (± 1 se) for men ages 18–22 by age and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias.

physicians, biological scientists, or physical scientists and reported an age of 26 or older ($N = 4593$, which constituted 12% of occupations reported by participants in that age range). Age 26 was used as a threshold to roughly control for the youngest typical age of attaining an MD degree in the U.S. (Association of American Medical Colleges, 2014).

Explicit Science Identity: Importance of being Personally Knowledgeable about Science

Also added to the survey in December 2006 were questions about personal knowledge goals in three broad domains, liberal arts, math, and science. Specifically, each participant was asked about a random two of the three, as follows: “Rate the following personal-goal-statements on their importance to you”:

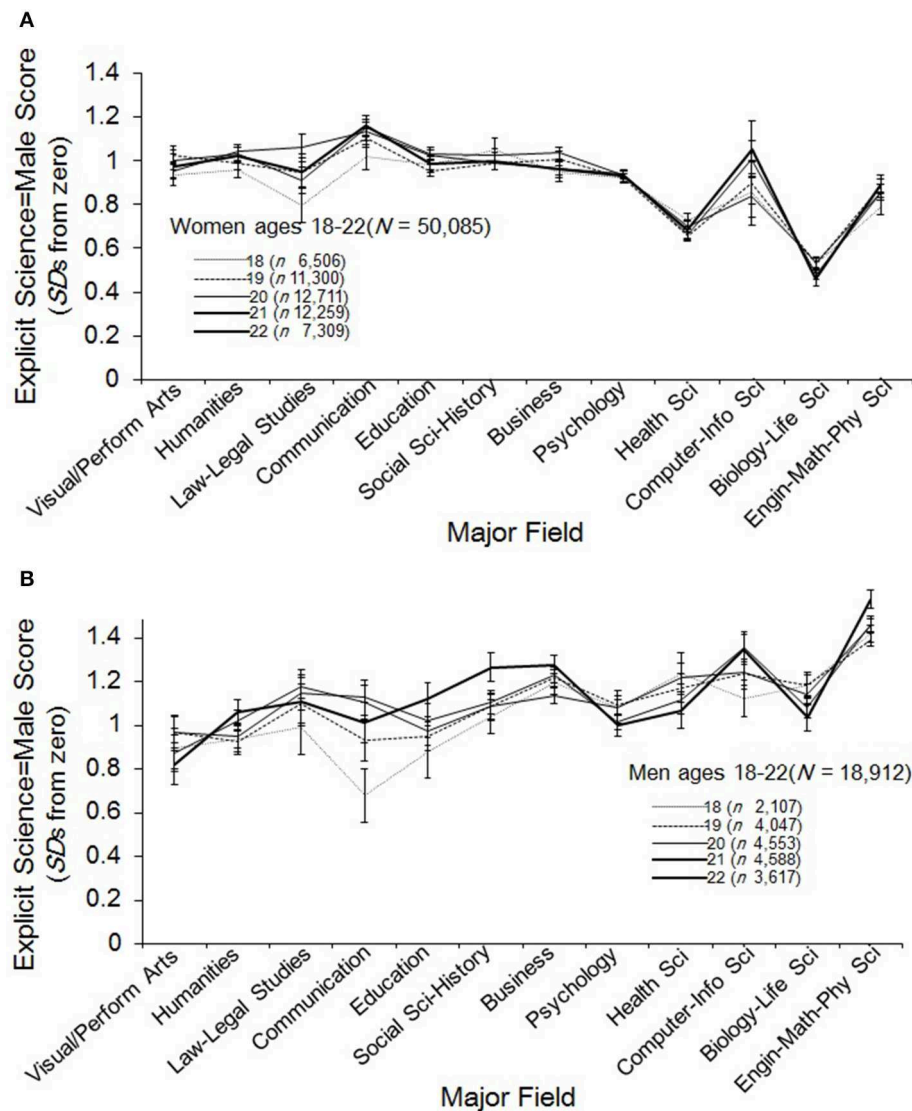


FIGURE 4 | (A) Mean explicit science = male score (± 1 se) for women ages 18–22 by age and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias. **(B)** Mean

explicit science = male score (± 1 se) for men ages 18–22 by age and major field. Majors are ordered, left to right, by ratings of science content (method described in Supplementary Material). A score of zero indicates no academic gender bias.

“Being knowledgeable about liberal arts.” “...about math.”
“...about science.”

The five rating options, with our coding in parentheses, were: Not at all important (0), Slightly (1), Moderately (2), Very (3), and Extremely important (4). The science question was answered by $N = 69,929$ participants.

Mapping our Four STEM Academic Major Categories to Collegiate and Professional STEM Gender Ratios

Based on our review of NSF data (National Science Foundation, National Center for Science and Engineering Statistics, 2011a), we classified as either relatively high-female or low-female the

gender ratios of the four STEM major categories from which our participants could choose:

High-female:

Biological sciences/life sciences.
 Health professions or related sciences.

Low-female:

Computer and information sciences.
 Engineering, mathematics, or physical sciences/science technologies.

The biological and health science fields are classified as high-female because women have recently constituted half or more of college graduates and employed scientists, while engineering,

physical and computer science fields are classified as low-female because women tend to constitute less than one-third of undergraduates and scientists in these areas. While there is considerable variation of gender ratios within the disciplines of our category *Engineering, mathematics, or physical sciences/science technologies* (e.g., considering bachelor's degrees awarded in 2008, female proportions were 50% in chemistry, 44% in mathematics, 40% in earth sciences, 20% in physics, and 19% in engineering), nearly half of women earning degrees in these areas are in the particularly low-female ratio fields of engineering and physics (National Science Foundation, National Center for Science and Engineering Statistics, 2011a). If the proportions of women recently graduating in these subfields match the underlying proportions among our volunteers choosing the physical sciences category, then a reasonable approximation of the upper limit for the average percentage of women encountered by those reporting a physical science major is 32%.

Explicit Academic Gender Stereotypes

Explicit academic gender stereotypes were assessed separately for both "Liberal Arts" and "Science" by asking participants to "Please rate how much you associate the following domains with males or females." Five response options were provided on the questionnaire until December of 2006 (strongly male, somewhat male, neither male nor female, somewhat female, strongly female) and seven options were provided thereafter (replacing the "somewhat" options with "moderately" and "slightly" options). Thus, 40% of participants answered with a 5-point scale and 60% with a 7-point scale. Regardless of scale type, a "neither male nor female" response was coded zero, stereotype-congruent responses were coded with positive integers and stereotype-incongruent responses were coded with negative integers (i.e., for the science-gender item, coding was either -2 to 2 or -3 to 3, with positive scores indicating stronger science-male associations; while for the arts-gender item, positive scores indicate stronger arts-female associations). To facilitate comparison of scores across the 5- and 7-point scales, scores were standardized within scale-type relative to a score of zero (means for 5- and 7-point standardized scores for science-male stereotype were 0.99 and 1.01, respectively, and 0.66 and 0.67, respectively, for the 5- and 7-point arts-female stereotype).

Implicit Academic Gender Stereotypes

The IAT assesses the relative strengths of cognitive associations and was administered according to the recommendations of Nosek et al. (2005). The gender-science IAT required quickly sorting words into one of four designated categories—*Female*, *Male*, *Liberal Arts*, or *Science*—using two computer keys. Training established the proper category for four corresponding sets of words: for example, Woman, Mother and Wife with "Female"; Man, Father and Husband with "Male"; Arts, Literature and Philosophy with "Liberal Arts"; and Biology, Chemistry and Physics with "Science" (complete list can be seen in the Supplementary Materials). Each participant sorted under two conditions: (1) stereotype-congruent, in which *science* and *male* words were sorted with one key, *liberal arts* and

female words with the other; and (2) stereotype-incongruent, in which *science* and *female* words were sorted with one key, *liberal arts* and *male* words with the other. The order of the conditions was randomized. Faster correct sorting in the stereotype-congruent condition than in the stereotype-incongruent condition indicates greater strength of science-male (and liberal arts-female) associations relative to science-female (and liberal arts-male) associations. Data were cleaned according to guidelines recommended by Nosek et al. (2005) and used by Nosek et al. (2009) to guard against careless responding. These procedures resulted in disqualification of IAT data for 11% of respondents (see Supplementary Materials for details). An IAT *D* score was computed for each participant by taking the difference in mean response latency between the conditions and scaling it by the overall variation (*SD*) of the participant's response latencies (Greenwald et al., 2003). Raw *D* scores were then standardized for the entire sample relative to a score of zero, thus allowing standard-deviation-unit comparisons with the explicit stereotype scores. For the sake of simplicity we refer to this measure as the "gender-science" IAT and say, for example, that positive scores indicate science-male stereotyping. However, it is important to note that arts-female associations are an integral, inseparable component of this IAT (Nosek et al., 2005).

Procedure

Upon entering the online *Project Implicit Demonstration* portal, participants were presented, in randomized order, with a list of topics from which to choose. Those who selected "Gender-Science," were presented with three study components in randomized order: (1) a questionnaire about academic attitudes, goals and stereotypes, (2) the gender-science IAT, and (3) a brief demographic questionnaire.

Results

Given the large sample sizes, even very small differences between means are significant at $p < 0.0001$. Therefore, we focus our reporting on the effect sizes, mostly Cohen's *ds*, and the reader can assume that if *p*-values are not given, they are less than 0.0001. Following Halpern et al.'s (2007) report on sex differences in science and math, we use the term *sex* when distinguishing men's and women's cognitions. To facilitate comparability, both implicit (*Istd*) and explicit (*Estd*) stereotype scores are expressed in standard deviation units relative to a zero, or no bias, score. There are two sets of analyses for each of our hypotheses, one with participants grouped by academic major, and another focused on those reporting scientific professions. Descriptive statistics are listed in **Table 1** and all data and materials are available at the Open Science Framework (<https://osf.io/y7a3n>).

Hypothesis 1a: Implicit Stereotype Differences between Female and Male Scientists

Women who are strongly identified with science will have relatively weak implicit stereotypes, while men who are strongly science-identified will have relatively strong ones.

TABLE 1 | Descriptive statistics for stereotypes and importance of scientific knowledge by sex and academic major of highest degree.

	Gender stereotypes, <i>M</i> (<i>SD</i>)					Goal: science knowledge		
	<i>N</i>	<i>N</i> -pct	<i>I</i> -std	<i>E</i> -Scistd	<i>E</i> -Artstd	<i>N</i>	<i>M</i> (<i>SD</i>)	Extreme%
WOMEN								
Visual or performing arts	4348	3.5	1.23 (0.88)	0.98 (0.94)	0.55 (0.96)	1662	2.44 (1.01)	15
Humanities/liberal arts	12,522	10.1	1.28 (0.91)	1.07 (0.97)	0.71 (1.02)	4549	2.48 (0.97)	16
Law or legal studies	4694	3.8	0.99 (0.96)	0.98 (1.01)	0.70 (1.02)	1654	2.52 (0.98)	17
Communications	4495	3.6	1.28 (0.89)	1.12 (0.97)	0.89 (1.03)	1665	2.30 (0.99)	11
Education	16,808	13.5	1.05 (0.95)	1.00 (0.99)	0.80 (1.02)	6519	2.62 (0.95)	19
Social sciences or history	8760	7.0	1.15 (0.93)	1.03 (0.97)	0.72 (1.01)	3332	2.55 (0.95)	16
Business	12,520	10.1	0.99 (0.92)	1.03 (1.03)	0.81 (1.04)	4878	2.43 (0.98)	13
Psychology	20,547	16.5	1.05 (0.94)	0.95 (0.96)	0.82 (0.99)	8414	2.73 (0.92)	21
Health or related sciences	14,403	11.6	0.73 (0.98)	0.72 (1.01)	0.68 (0.99)	6337	3.22 (0.80)	42
Computer and info sciences	2607	2.1	0.74 (0.99)	0.98 (1.01)	0.65 (0.99)	919	3.01 (0.90)	34
Bio/life sciences	12,755	10.2	0.35 (1.02)	0.58 (0.95)	0.56 (0.91)	5558	3.61 (0.59)	66
Engin, math, phys sciences	10,020	8.0	0.39 (1.03)	0.88 (0.97)	0.61 (0.92)	4269	3.46 (0.71)	57
All	124,479		0.92 (1.00)	0.92 (0.99)	0.72 (1.00)	49,756	2.85 (0.98)	30
MEN								
Visual or performing arts	1954	3.7	0.65 (0.99)	0.98 (0.98)	0.19 (0.88)	635	2.59 (0.95)	17
Humanities/liberal arts	5254	10.0	0.64 (1.03)	1.06 (1.00)	0.24 (0.90)	1837	2.53 (0.97)	17
Law or legal studies	2440	4.7	0.80 (1.01)	1.17 (1.04)	0.45 (1.02)	976	2.56 (0.96)	17
Communications	1522	2.9	0.67 (1.00)	1.06 (1.02)	0.38 (0.97)	551	2.40 (0.99)	13
Education	3179	6.1	0.83 (1.00)	1.06 (0.99)	0.52 (1.00)	1254	2.67 (0.95)	20
Social sciences or history	3393	6.5	0.73 (1.00)	1.11 (0.99)	0.34 (0.97)	1251	2.65 (0.95)	20
Business	7853	15.0	1.00 (0.95)	1.25 (1.00)	0.71 (1.07)	3102	2.57 (0.92)	15
Psychology	4779	9.1	0.91 (0.99)	1.08 (0.96)	0.50 (0.97)	1888	2.83 (0.92)	26
Health or related sciences	2899	5.5	1.12 (0.94)	1.14 (0.98)	0.63 (1.01)	1282	3.28 (0.77)	45
Computer and info sciences	4257	8.1	1.09 (0.94)	1.32 (0.99)	0.62 (0.99)	1480	3.09 (0.81)	33
Bio/life sciences	4433	8.5	1.18 (0.95)	1.07 (0.95)	0.56 (0.96)	1834	3.53 (0.63)	60
Engin, math, phys sciences	10,493	20.0	1.21 (0.96)	1.39 (0.96)	0.64 (0.97)	4083	3.39 (0.73)	51
All	52,456		0.97 (0.99)	1.19 (0.99)	0.53 (0.99)	20,173	2.93 (0.94)	32

Categories of majors are ordered, top to bottom, from lowest to highest science-content ratings made by psychology graduate students blind to our study hypotheses. Stereotype scores (suffix std) are standardized across the full sample relative to a score of zero. Stereotype labels: *I*, Implicit science-male; *E*-Sci, Explicit science-male; *E*-Art, Explicit arts-female. Response scale for the "Goal: Science Knowledge" variable was Not at all important, Slightly important, Moderately important, Very important, and Extremely important, and is coded 0–4. Extreme% indicates the percentage of respondents choosing Extremely important.

Identification by Academic Major

Averaged across the entire sample, implicit science–male stereotyping was strong, nearly a full standard deviation above zero, $Istd = 0.93$. Overall, participant sex made a trivial difference, men averaging higher by just 0.05 standard deviations (i.e., Cohen's $d = 0.05$). However, as predicted, substantial sex differences were observed when participants were grouped by their academic major (**Figure 2A**), the direction of the difference varying systematically with rankings of the majors' degree of science content. The largest differences between men and women ($ds \sim 0.8$) came in the fields rated highest in scientific content (biological and physical sciences), where men's stereotyping was the strongest among all men and women's was the weakest among all women. This pattern conforms to Greenwald et al.'s (2002) cognitive consistency principles. That is, the strongest science–male (liberal arts–female) stereotypes are observed among those whose sex is aligned with their major in a stereotype-congruent fashion (e.g., among women identified

with strongly non-STEM majors like arts and humanities, and among men identified with STEM majors), while the weakest stereotypes are seen among those with stereotype-incongruent combinations (men in arts and women in STEM). This pattern makes clear that this implicit stereotype is not simply a socially-shared association acquired through cultural exposure. Rather, it reveals important dependencies with combinations of gender identity and science/arts identities.

While women in STEM have the weakest science–male stereotypes, suggesting an important relation with their scientific identity, it is notable that they do not evidence a counter-stereotypical implicit association. On average, women majoring in STEM of any kind (the four right-most groups of majors in **Figure 2A**) still implicitly stereotyped science as male by half a standard deviation above the zero-point of no stereotyping ($Istd = 0.53$), and even those in the two categories of majors with the lowest average stereotypes, biological and physical sciences, still evidenced stereotypes of at least a third of a standard

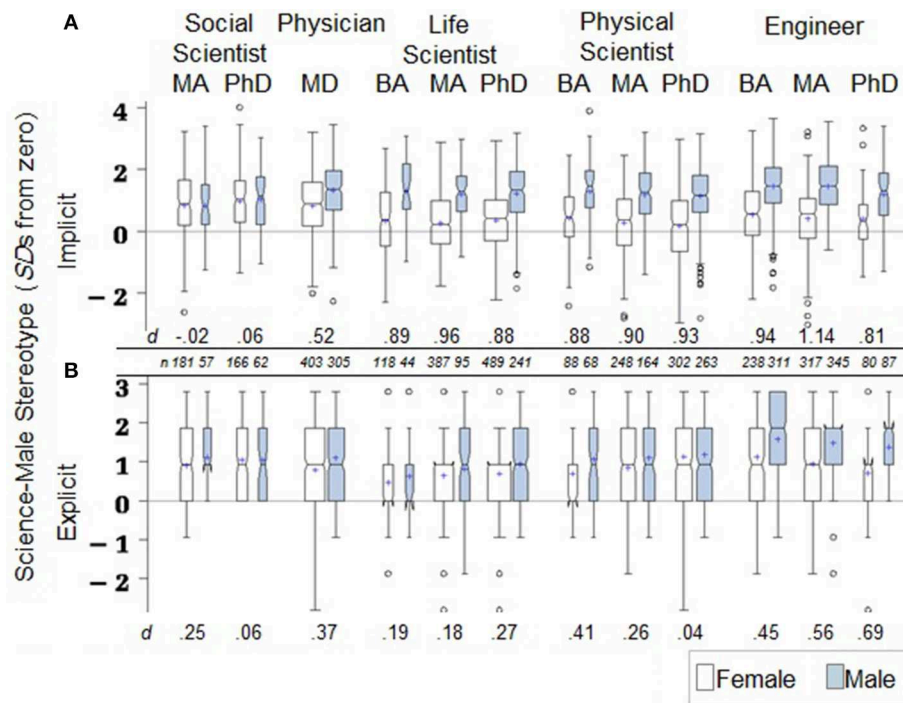


FIGURE 5 | (A,B) Boxplots of stereotype scores by sex and science profession for participants age 26 and older. Box width is proportional to subgroup n and notches indicate standard errors around the median. d scores are standardized sex effects, male mean minus female mean, within each profession/degree category.

deviation in the science-is-male direction, $Istd$ 0.33 and 0.39, respectively.

Identification by Scientific Profession

When participants are classified by scientific profession (see **Figure 5A**), we find the same general sex-effect pattern that was observed within academic major classifications—stronger stereotyping by men than women (this is not the case among social scientists, shown only for comparison, who stereotyped at a robust level but without sex differences). Male physicians, life scientists, physical scientists, and engineers all evidence much stronger levels of implicit science–male stereotype than the women in these professions, with a median sex-difference effect size of d of 0.89. Indeed, the smallest d , 0.52 for the physicians, is something of an outlier, with the next smallest effect being $d = 0.81$. The smaller sex effect for the physicians is driven by the relatively stronger average stereotype evidenced by the women, which is higher than that of the next highest female group in **Figure 5A**—engineers with bachelor's degrees—by a d of 0.29.

Hypothesis 1b: Explicit Stereotype Differences between Female and Male Scientists

Women who are strongly identified with science will have relatively weak explicit science-is-male stereotypes, while men who are strongly science-identified will have relatively strong ones. Conscious motivations to respond in accord with both personally-

and societally-endorsed values, however, are expected to constrain the magnitude of effects relative to those for the implicit stereotype.

Identification by Academic Major

Responses to the two explicit stereotype measures, science–gender and arts–gender, are shown as a function of major in **Figures 2B,C**. As was the case for implicit stereotypes, overall means averaged strongly in stereotypical directions (science–male and arts–female), though associations of science with male were stronger than those of arts with female, $Estd = 1.01$ and 0.66, respectively. Sixty-three percent of participants reported associating science at least “slightly” with male, compared with 46% who associated arts at least “slightly” with female. Science–male and arts–female ratings were not highly correlated ($r = 0.28$).

Unlike for implicit stereotyping, there were overall sex differences in explicit stereotyping, with men more likely than women to associate science with male (70 vs. 60%), and women more likely than men to associate arts with female (49 vs. 41%). As noticeable in **Figures 2B,C**, these sex differences are driven by participants in the corresponding STEM and non-STEM majors, respectively. That is, the sex difference in science stereotyping is primarily among STEM majors and the sex difference in arts stereotyping is among non-STEM majors. Supporting our hypothesis, the sex differences in explicit science stereotyping were largest among the STEM majors, men stereotyping more strongly than women (ds ranging 0.35–0.52), but smaller than the sex differences observed for implicit stereotyping. The

sex difference in arts stereotyping, however (**Figure 2C**), owed virtually nothing to STEM majors (a d of 0.06 is the largest sex difference in any of the STEM fields). It came primarily, instead, from those in the eight *non*-STEM fields (median $d = 0.35$), women stereotyping more than men.

Identification by Scientific Profession

The same, expected pattern of stronger stereotyping by men in STEM was also observed for the explicit gender stereotypes of science professionals aged 26 and older (see **Figure 5B**). Male physicians, biological and physical scientists, and engineers ($N = 1923$) averaged $Estd = 1.23$, $SD = 0.99$, while women in these fields ($N = 2670$) averaged $Estd = 0.82$, $SD = 0.98$, for a $d = 0.41$. The magnitude of the sex difference in explicit stereotyping is, again, less than for implicit stereotyping.

Hypothesis 2a: Implicit Stereotype Differences as a Function of Gender Ratios in Science Environments

Implicit science-is-male stereotypes will be stronger for both women and men in low-female STEM fields than in high-female fields, though sex differences should remain.

Identification by Academic Major

The primary question of our study is whether science-is-male stereotypes vary as a function of gender ratio differences across scientific disciplines. First, it is apparent from **Figure 2A** that variation in average implicit stereotyping across the four science domains (at right) is greater for women than for men. Within-sex analysis of variance (ANOVA) modeling for participants in these domains yielded an extremely small effect of scientific category for men, $R^2 = 0.003$, $F_{(3, 22078)} = 20.9$ ($Istd = 1.17$, $SD = 0.95$), but a larger one for women, $R^2 = 0.031$, $F_{(3, 39781)} = 421.6$ ($Istd = 0.52$, $SD = 1.02$). The noticeable stereotyping difference for women does not align with differences in gender ratios across the scientific disciplines. If stereotypes covaried with gender ratios, we would expect differences in the strength of stereotypes evidenced by women in the health vs. computer science fields (which are high- and low-female, respectively) and in biological vs. physical sciences (also high- and low-female). We find, however, that the stereotype strengths for each of these comparisons differ very little, $F_{(1, 17008)} = 0.2$, $p = 0.68$, for health vs. computer science, and $F_{(1, 22773)} = 7.4$, $p = 0.007$, $d = 0.037$, for biological vs. physical sciences. The noticeable difference falls, instead, between health and computer sciences combined, on one hand (average $Istd = 0.73$, $SD = 0.98$), and biological and physical sciences combined on the other (average $Istd = 0.37$, $SD = 1.03$). The implicit stereotyping difference between these two combinations was more than a third of a standard deviation, $d = 0.37$, $R^2 = 0.031$, $F_{(1, 39783)} = 1257$.

These patterns do not support our hypothesis that strength of implicit stereotyping among those in science domains will vary with gender ratios. Rather, the noticeable difference for women tracked with differences in scientific values as

indicated by their ratings of the personal importance of scientific knowledge. Those with the weakest stereotypes—the women in biological and physical sciences—assigned the greatest importance to the personal goal of being knowledgeable in science. As seen in **Table 1**, 66% of biological sciences women and 57% of physical sciences women rated knowledge of science as an extremely important personal goal, compared with only 42 and 34%, respectively, of women with health and computer sciences majors. Following Baron and Kenny's recommended steps (Baron and Kenny, 1986; Kenny, 2014), we used three regression models to evaluate the science-knowledge-importance variable as a potential mediator of the difference in implicit stereotyping between biological/physical sciences women and health/computer sciences women. Model 1 estimated the bivariate regression between group membership (X , dummy-coded 0 for health/computer science majors and 1 for biological/physical science majors) and implicit stereotyping (Y); Model 2 estimated the bivariate regression between X and the proposed mediator (M , the science-knowledge-importance variable, coded 0–4); and Model 3 estimated the multiple regression of implicit stereotyping on both X and M . The baseline effect (Model 1) of X (being a biological/physical science major) on implicit stereotyping was estimated as $b = -0.36$, model $R^2 = 0.031$. Model 2 demonstrated that type of science field (X) predicts science-knowledge-importance score (M), $b = 0.35$, model $R^2 = 0.054$. When both X and M were included in the multiple regression Model 3, the effect of being a biological/physical science major was reduced to $b = -0.29$, a reduction of about one-fifth compared to the estimate from Model 1, and model R^2 nearly doubled to 0.057. Thus, an indicator of the strength of personal scientific values provided some traction in accounting for stereotyping differences among women in the different science groups.

Identification by Scientific Profession

Since the sex effects (Hypothesis 1a; **Figure 5A**) were fairly uniformly large and not our critical question, we fit separate within-sex bivariate regression models to estimate gender-ratio effects. First, as previously noted, female physicians had stronger implicit stereotypes than women in any other STEM professional group. This is incongruent with the gender-ratios hypothesis that predicts weaker stereotyping among physicians, relative to physical scientists or engineers, given the relatively high-female ratio in medicine.

Among the remaining types of scientists, we estimated gender-ratio effects by contrasting the implicit stereotypes of life scientists (coded 0), where high-female ratios are more likely, with those of physical scientists and engineers (coded 1), where low-female ratios are more likely. Results of regression analyses predicting implicit stereotyping from this contrast of disciplines were non-significant for both women and men (for women, $b = 0.000$, $t = -0.01$, $p = 0.99$, $R^2 = 0.000$; for men, $b = 0.135$, $t = 2.67$, $p = 0.008$, $R^2 = 0.004$). Thus, our gender-ratio hypothesis for implicit stereotyping is not supported when tested among professional scientists.

Hypothesis 2b: Explicit Stereotype Differences as a Function of Gender Ratios in Science Environments

Explicit science-is-male stereotypes will be stronger for both women and men in low-female STEM fields than in high-female fields, though, again, group variation on explicit stereotype means should be somewhat constrained by conscious values and motivations to respond without bias.

Identification by Academic Major

Unlike for implicit stereotyping, patterns of explicit science-male stereotyping conformed to our gender-ratio hypothesis (see **Figure 2B**). For both men and women in sciences, the weakest explicit stereotypes were in the domains where women are more strongly represented, i.e., in health and biological sciences, and the strongest were seen where women are least represented, in computer and physical sciences. Notably, for scientific men who are in high-female fields (health and biological sciences) stereotype levels are rather generic, i.e., similar to those among the non-STEM men and women. For such men, while their identity ("I'm scientific and I'm male") maps onto the stereotype, their environment, on average, belies the stereotype (not clearly male-majority). It is only the men in majority-male environments, computer and physical sciences, who deviate (upward) from the generic level of stereotyping. For scientific women, the generic level is seen for those in the low-female fields, where, again, there is mismatch between their identity and the stereotype manifested in gender-ratios. In their case, however, personal identities, scientific and female, belie the stereotype and the environment supports it. The women whose stereotypes deviate (downward) from the generic tend to be in the high-female fields in which gender ratios complement their identities in undermining the stereotypical propositions they may consider when explicitly reporting gender-science associations.

Identification by Scientific Profession

The explicit stereotypes of scientific professionals were also congruent with the gender-ratio predictions of hypothesis 2b. Physical scientists and engineers, together, had stronger stereotypes than life scientists, d effect sizes of 0.37 for women and 0.52 for men (the estimated stereotyping effect of being a physical scientist or engineer was $b = 0.35$, $t = 9.64$, $p < 0.0001$, $R^2 = 0.032$ for women; $b = 0.50$, $t = 9.81$, $p < 0.0001$, $R^2 = 0.047$, for men).

Since physicians' explicit science-male stereotypes did not obviously differ from those of the other STEM scientists, as female physicians' did for implicit stereotype, we included physicians in another set of regression analyses with PhD-level participants in the other STEM domains. That is, we contrasted physicians/life scientists at the MD or PhD-level (coded 0) with physical scientists/engineers at the PhD-level (coded 1). We relaxed alpha to 0.01 because of the smaller cell sizes (e.g., $N = 350$ male physical scientists/engineers with PhDs). Effects again supported our gender-ratio hypotheses, albeit less strongly among these MD/PhDs (for women, $b = 0.30$, $t = 5.26$, $p <$

0.0001 , $R^2 = 0.022$; for men, $b = 0.20$, $t = 3.09$, $p = 0.0021$, $R^2 = 0.011$).

Thus, unlike results for the implicit stereotype, the patterns of explicit science-is-male stereotypes generally conform to our gender ratios hypothesis (2b) for STEM professionals. Physicians and life scientists, who are more likely to work in high-female ratio environments, explicitly stereotype science as male less strongly than do engineers and physical scientists, who are more likely to find themselves in low-female ratio work settings.

Hypothesis 3a: Implicit Stereotype Differences as a Function of "Dosage" of Exposure to Given Gender-Ratios

Implicit science-is-male stereotyping should increase with prolonged exposure to low-female STEM environments and decrease with ongoing exposure to high-female STEM environments.

Identification by Academic Major

Women's means are plotted in **Figure 3A** for each year of age, 18–22, across all 12 academic categories. If length of exposure to collegiate science environments with skewed gender ratios has an effect on implicit stereotypes, then in the low-female computer and physical science domains we should see stereotype-strengthening across these ages, and weakening in the high-female health and biological science domains. We tested these expectations with ANOVAs contrasting stereotype means across the five age groups for each sex within each of our four categories of STEM majors. Given the smaller cell sizes in these models for the effect of age within sex-by-major groups (the smallest being the samples for women in computer science, ranging from $N = 30$ –74), alpha for significance testing was reduced to 0.01.

Each of the ANOVA models for women was non-significant, bearing out the impression of stability suggested by the overlapping standard error bars around these means in **Figure 3A**³. Thus, for women in STEM fields, there was not statistically significant variation in implicit science stereotyping across groups spanning the traditional age range of college study.

For men, age effects were non-significant for all but one domain, biological science, $R^2 = 0.02$, $F_{(4, 1152)} = 6.23$, $p < 0.0001$ ($Istd = 1.23$, $SD = 0.90$). The pattern among the men in biology was of increasing stereotype strength (see **Figure 3B**), despite the majority proportion of women among biology majors nationally. Eighteen-year-old men in biology averaged $Istd = 1.13$ ($SD = 0.89$), compared with $Istd = 1.48$ ($SD = 0.84$) for the 22-year-olds, an effect of $d = 0.31$. Thus, a lack of difference across years of age was the dominant finding for women and men, and the one instance of significant age effect is in a direction opposite to the gender-ratio dosage hypothesis.

³To insure that results were not affected by extreme observations, models were also run without the lowest and highest 1% of IAT scores within ageXgenderXmajor categories. ANOVA result patterns were unchanged under these conditions.

Identification by Scientific Profession

To test whether implicit stereotyping varies with increasing duration and intensity of training, indexed by degree level, we estimated regression models using scientist-type (life scientist, coded 0, vs. physical scientist/engineer, coded 1) and degree-level (bachelor's, master's, or PhD) as predictors. Alpha was set at 0.05 because of the relatively small numbers in some categories, e.g., $N = 68$ for male physical scientists with a bachelor's degree. Two orthogonal contrast codes were used to index degree-level effects, code-1 for Masters vs. PhD and code-2 for Bachelors vs. Masters and PhD, together. Because degree-level is confounded with age, age was included as a covariate in all models (and yielded a positive main effect, but no interactions, for both sexes).

No effects of degree-level were found for men, but for women a significant interaction was observed between scientist-type and Masters- vs. PhD-level degrees, $t = -2.58$, $p = 0.01$. Implicit stereotyping strengthened with higher degrees among the life scientist women, but weakened with higher degrees among physical scientists and engineers, a pattern precisely opposite to our hypothesis (3a) that greater tenure in a field would correlate with stronger or weaker implicit stereotype depending on female-male ratios. Specifically, among female life scientists, implicit stereotypes were $Istd = 0.25$ ($SD = 0.97$) at the Masters level and $Istd = 0.37$ ($SD = 0.99$) at the PhD level, compared with those of female physical scientists and engineers that were $Istd = 0.31$ ($SD = 1.02$) at the Masters level and $Istd = 0.19$ ($SD = 1.05$) at the PhD level. We hypothesized the opposite, that weaker stereotyping would occur with higher degree-attainment in the relatively high-female life sciences, and stronger stereotyping would occur with higher degree-level in the low-female physical and engineering sciences.

Hypothesis 3b: Explicit Stereotype Differences as a Function of "Dosage" of Exposure to given Gender-Ratios

Explicit stereotypes, when measured for scientists in a given field with roughly constant gender-ratio, will not be systematically responsive to dosage because the general propositions being weighed may not change (systematically) very much.

Identification by Academic Major

Figures 4A,B show plots for women and men, respectively, of explicit science-male stereotype means for each age across the academic major categories. The only instance of age differences in stereotyping at the 0.01 alpha level was a very small effect for men in physical sciences, $R^2 = 0.007$, $F_{(4, 2211)} = 3.76$, $p = 0.005$ ($Estd = 1.46$, $SD = 0.90$). The peculiarity of this finding, the one significant test out of eight, warrants circumspection. Overall, the lack of variation in explicit stereotyping among STEM majors across the college years, supports our hypothesis of no systematic change for either men or women when the gender ratio of the given field is assumed constant.

Identification by Scientific Profession

Using the same regression estimation approach as was described for the implicit stereotype analysis (contrast-coded predictors

of scientist-type, life scientist vs. physical scientist/engineer, and degree-level, alpha 0.05), we found no dosage effect of degree level for explicit science-male stereotype among women, and a significant, but unpredicted interactive pattern for men similar to that observed for the implicit stereotyping of women, $t = -2.11$, $p = 0.035$. Among male life scientists, explicit stereotypes were stronger at the PhD level ($Istd = 0.95$, $SD = 0.94$) than at the Masters level ($Istd = 0.75$, $SD = 0.95$), but the opposite held for male physical scientists and engineers, who were weaker at the PhD level ($Istd = 1.25$, $SD = 1.00$) than at the Masters level ($Istd = 1.41$, $SD = 0.96$). This pattern supports our hypothesis that evidence of systematic change of explicit stereotyping was not expected within environments of particular gender ratios.

Discussion

With a sample of 176,935, including thousands of engineers, physicians and scientists, we examined science-is-male stereotypic associations as a function of sex, scientific identity, and gender ratios in scientific disciplines. Stereotyping science as male was normative, implicitly and explicitly, as both types of scores averaged roughly a standard deviation above the zero-level of stereotyping on the respective scales. However, both types were marked by considerable variation depending on sex and academic or career identity, demonstrating that these gender associations are not simple reflections of a common cultural stereotype in the air. As expected, consistent with a well-established literature, we observed a positive relationship between stereotyping and science identity for men and a negative relationship for women. Men in STEM evidenced stronger science-is-male stereotypes than their non-STEM brethren, especially implicitly, while women in STEM evidenced the opposite pattern, much weaker implicit stereotyping than non-STEM women. As a result, in biological and physical sciences and engineering (the categories of science majors in our study that were rated as most scientific), the sex difference in implicit stereotyping was large, more than 0.8 standard deviations, ranking among the largest sex differences seen in cognitive research (Miller and Halpern, 2014).

Our primary question, however, was whether strength of science-male stereotyping would vary across scientific disciplines as a function of gender ratios in the disciplines. This hypothesis was supported for explicit stereotypes, but not for implicit ones. As expected, relatively stronger *explicit* stereotypes were evidenced by scientists studying and practicing in fields where women continue to be distinct minorities, and weaker ones were expressed by scientists in fields where women are better represented. *Implicit* stereotyping differences between scientists in different disciplines, however, did not correspond with gender ratios. For men there was little variation in implicit stereotype strength across four classifications of academic science concentration. Women, in contrast, evidenced considerable variation across these classifications, but it did not coincide with gender ratio differences. Rather, it coincided with differences in an indicator of the women's scientific identity. Specifically, implicit stereotyping varied with the value the women assigned to being personally knowledgeable about science. Women reporting

that personal knowledge of science was “extremely important” had weaker implicit stereotypes than women reporting less personal priority on scientific knowledge. Though biological and physical science fields vary greatly in typical gender ratios, women in these disciplines were similar in the degree to which they placed extreme importance on personal scientific knowledge and in having the weakest implicit stereotypes of all women, whereas women in computer and health sciences, disciplines that also differ markedly in gender ratio, placed less importance on scientific knowledge and stereotyped more strongly. We found, furthermore, little evidence of difference in implicit stereotype strength corresponding to “dosage” of exposure to particular gender ratios. That is, within a particular field of whatever typical gender ratio, greater duration and intensity of exposure (whether operationalized by the cross-sectional proxy of traditional college ages from 18 to 22, or by levels of training among practicing scientists, BA, MA, or PhD) did not correspond to different implicit stereotype strengths as expected.

Why Didn't the Implicit Science-male Stereotype Vary with Gender Ratio Differences in Science Fields?

While our analyses make clear that implicit gender-science stereotype strength varies greatly, primarily for women, across different scientific disciplines, gender ratios were not found to be a factor. How can this be if implicit stereotypes are sensitive to environmental inputs (Gawronski and Bodenhausen, 2006; Ratliff and Nosek, 2010; Miller et al., in press)? The answer may lie in Greenwald et al.'s (2002) assertion that the *self* is the power-center of automatic associative processes. Once strong self-concept bonds are formed (e.g., me-woman; me-science), the resulting, secondary, stereotypical associations (women-science) may be fairly impervious to local environmental conditions, like a preponderance of men in the lab, that would otherwise change them. Miller et al.'s (in press) country-level analysis identified precisely the relation between collegiate science gender-ratios and implicit stereotyping that we expected—but that we did not find—at a scientific discipline-level of analysis, i.e., higher female proportions in science associated with lower science-is-male stereotypes. We suspect that the apparent incongruence between their finding and ours hinges on scientific self-concept. That is, their analysis took into account respondents' country of citizenship and gender, but did not distinguish between levels of personal scientific identification, while ours controlled for self-reported academic major and priority on personal scientific knowledge. Our finding leads us to expect that the implicit stereotypes held by strongly science-identified women, like majors or scientists in biological or physical sciences, will be similar across countries, regardless of country proportion of women in collegiate science. That is, we would now expect science identity to trump the influence of local conditions.

Ratliff and Nosek (2010) note that, while implicit associative processes do a good job of accounting for covariation

of events in the environment—like female proportions in science settings—they are also influenced by the frequency of association activations. Thus, if self-associations enjoy a leverage advantage in cognitive evaluative networks to begin with, as postulated by Greenwald et al. (2002), and self-associations also are more frequently activated than more abstract group-associations, then once a strong positive implicit science-self association is established (me-science), it may overpower potentially conflicting science-gender associations conveyed by the environment. We did not measure implicit science-self associations, but research indicates that they are strongly positively correlated with explicit indicators of science identity and favorability like ones we measured (Nosek and Smyth, 2011). Dasgupta's (2011) stereotype inoculation model hinges on developing a strong implicit STEM self-concept as a protection against pervasive cultural stereotypes and the vagaries of local conditions. Our data suggest that women in the most scientifically demanding fields, regardless of gender-ratios, are anchored at low levels of implicit stereotyping by their weighty scientific self-concepts and values.

Our “dosage” inference—that women's implicit stereotyping, within any particular academic major, is largely stable from age 18 on and across increasing levels of training and professional attainment—suggests that women's implicit stereotypes about gender and science may be fairly stable once strong scientific self-concepts are established. These cross-sectional data, however, can only be suggestive. Longitudinal research across the adult age range we studied is necessary for confidence in this pattern. Yet even if stability of adult implicit science associations was well-established, a more pressing longitudinal question would remain: How do children's and adolescents' implicit scientific associations develop and to what extent do they influence consequential STEM behaviors and choices? Galdi et al.'s (2014) demonstration that brief exposure to a stereotypical gender-math image vs. a counter-stereotypical one influenced both the implicit gender-math stereotyping and math performance of six-year-old girls should be a clarion call to such research. Longitudinal data on the development of implicit scientific self-concepts and stereotypes would help shed light on our “self-as-power-center” explanation for later stereotyping differences across scientific disciplines.

Based on their cross-sectional findings with elementary school students, that implicit math-gender stereotypes were already in force and were stronger than implicit math self-concepts, Cvencek et al. (2011) speculate that the implicit stereotype precedes, and may influence formation of, the implicit self-concept. It may be that stereotypes influence the early formation of self-concepts, but that once self-concepts are strong they are no longer easily influenced by stereotypes or stereotypical environmental conditions. Whatever the early trajectory and leading causal influences, Tai et al. (2006) found that by eighth grade, scientific goals—explicit values—were predictive of earning science degrees, especially in the physical sciences and engineering. It is time to add understanding of how implicit science stereotypes and self-concepts relate to such critical formative trajectories.

Why Did the Explicit Science–male Stereotype Vary with Gender Ratio Differences in Science Fields?

According to Gawronski and Bodenhausen (2006), explicit associations are an amalgam of both automatic, associative processes, and controlled, propositional processes. The latter can be applied in deliberate attempts to adjust responses for the “truth-value” of evaluations or stereotypes. So in formulating their responses to the questionnaire item, “Please rate how much you associate science with males or females,” participants were able to exercise choice about how to weight possibly varying components of this association. A female physicist may have thought, for example, “Well, I love physics and am highly accomplished in the field, but I think this question is less about my personal experience and more about what I see as gender proportions in science, generally.” If most of her physics colleagues are male, such an interpretation of the question might have led her to select the “strongly male” gender-science association answer. Conversely, a female biology professor might have reasoned, “Most of my undergraduate students are female, and now a third of my faculty colleagues are female—with even higher proportions of women among the young stars—so I’ll pick the ‘neither male nor female’ answer.” Though each of these hypothetical women would explicitly report a strong self-identification with science, their reports of gender associations with science can be made relatively independently of their self-concept. Our data suggest, however, that their implicit gender–science associations are a function of their self-associations with science, resulting in similarly weak implicit science–male stereotypes regardless of the different truth-values the gender ratios of their environments might have suggested.

Limitations

Though the sample is large and there is more age and occupational variation than found in most studies of STEM stereotypes, it is not representative of any definable population. Participants are self-selected volunteers and their responses have no experimenter oversight. Generalizability to highly STEM-identified people is suggested, however, by our findings of comparable patterns of implicit stereotype strength and gender differences for University of Virginia undergraduates in engineering and advanced mathematics courses (Smyth unpublished manuscript; Martin et al., 2013; Smyth unpublished data). These students were not self-selected (their participation was a course requirement), yet the sex differences found in their stereotyping were similar in magnitude to those of the STEM-identified participants in the current study.

One reviewer expressed concern that the different methods of defining “Science” in our explicit and implicit stereotype instruments posed a potential confound for our results. Specifically, it was argued that the explicit stereotype instrument, asking participants to consider how strongly they associate

males or females with the general concept, “Science,” presents an amorphous target that is likely to be interpreted through the lens of respondents’ particular scientific discipline and experience, and so is prone to a correspondence between gender-ratios and the explicit stereotype. We agree with this interpretation and predicted that respondents’ local experience would, indeed, inform their rating of the stereotype strength. “Science” in the *Implicit Association Test*, on the other hand, is ostensibly defined by all of the exemplars that are sorted into the category, including, for example, Biology and Chemistry, both relatively high-female fields, and Engineering and Physics, both low-female fields. Thus, the reviewer argues, the science construct used in the implicit measurement is more clearly defined as all-encompassing, and participants may be less likely to frame it in light of their particular disciplinary experience. We agree that this is possible, but believe that it is unlikely. A measurement property of the IAT is that the category labels dominate assessment over the exemplars (De Houwer, 2001; Nosek et al., 2007). The IAT in this study used the category labels “Science” and “Liberal Arts.” The individual exemplars may only have a small effect in as much as they change the construal of those category labels (Nosek et al., 2005). We think it is likely, therefore, that this implicit measure, like the explicit one, invokes a rather general science construct that is also subject to framing by respondents’ particular experiences. Even so, replication with other explicit and implicit measurement techniques would be a useful check on this question.

Another impetus for replication with a different implicit science-is-male instrument is to avoid the IAT’s structural requirement of a contrasting category, in this case, the fairly distinct other academic stereotype of “gender-arts.” Our explicit stereotype measurement approach, which allowed separate measurement of gender–science and gender–arts stereotypes, underscored their distinctiveness, especially for participants identifying as STEM majors (see Results for Hypothesis 1 and **Figures 2B,C**). Further evidence from our study suggests, however, that the science–gender construct may, indeed, be driving performance on this IAT more than the arts–gender construct. Among the STEM majors in our sample, explicit science–gender stereotype was a better predictor of IAT score than was the explicit arts–gender stereotype (regression R^2 s of 0.063 and 0.013, respectively, and increasing only to 0.067 when both factors and their interaction were included in a multiple regression model).

Finally, our cross-sectional data require that inferences about the lack of environmental “dosage” effects be held cautiously until longitudinal studies are brought to bear. Of our two proxies for dosage (1) gradations of experience across the five years of traditional college age and (2) bachelors, masters and doctoral levels of scientific achievement, we agree with one of our reviewers that the latter is likely the more reliable. The dominant finding of scant evidence for dosage effects with either method, however, lends credence to the general conclusion that ongoing exposure to particular gender ratios, once strong scientific identities are established, may have little effect on personally-held stereotypes.

Conclusion

Male scientists, on average, hold substantially stronger explicit and implicit science-is-male stereotypes than do female scientists. The gender difference is greatest, exceeding 0.8 standard deviations, for the implicit stereotypes held by men and women in either biological/life sciences or engineering/physical sciences (about twice the size of the differences in health and computer science fields). Average stereotype strengths differ across scientific disciplines, but in different patterns for explicit and implicit stereotypes. Differences in explicit stereotype strength correspond to gender ratios. That is, lower proportions of women in a field predict stronger explicit science-is-male associations. Implicit stereotype differences, in contrast, do not track with gender ratios. The implicit stereotyping levels for female and male scientists in life sciences, for example, where women are strongly represented, are similar to the levels in physical sciences and engineering, where women remain distinct minorities. Regardless of gender ratio, implicit stereotype differences align with indicators of individuals' scientific identity, such that disciplines with higher proportions of extremely science-identified people are characterized by more extreme implicit stereotype averages, extremely high for men and extremely low for women.

For scientifically-identified adults within a given discipline (assuming a generally constant gender-ratio), neither explicit nor implicit stereotype levels vary much as a function of cross-sectional proxies for "dosage" of the exposure to that gender ratio. That is, within disciplines, stereotype strengths are comparable between newly-declared STEM majors at age 18, bachelors, masters and PhD STEM degree-holders, and practicing scientists. Though stereotype change was not measured, these cross-sectional data suggest that, once a scientific identity is established, implicit stereotype strength remains fairly constant at a low level for women and at a high level for men, regardless of immediate gender ratio or duration and intensity of training and practice. They further suggest that neither sex differences in implicit stereotyping, nor individual differences in implicit stereotyping, are likely to account for women's differential representation across scientific disciplines *once a major is declared*.

This is not to suggest that adults' STEM interest and pursuit is not influenced by implicit stereotypes and self-concepts. There is much to learn, for instance, about implicit influences for

the many who begin college without a clear major direction, as well as for the substantial number who intend a major in STEM at the start, but do not persist (Chang et al., 2014; Higher Education Research Institute, 2014). Still, it makes sense that research resources be focused at understanding the influences on children's self-concepts and stereotypes, as they are certainly more malleable and there is still time for interventions to work ahead of the coalescing of academic interests and goals during adolescence (Tai et al., 2006; Galdi et al., 2014). As noted by Ceci and Williams (2010), it is likely that a lion's share of the STEM sex difference derives from choices made prior to taking college courses.

Dasgupta (2011) emphasizes critical periods for inoculating girls' and women's implicit stereotype-incongruent self-concepts through increased exposure to same-sex peers and experts in the given domain. These critical inoculation periods are theorized to include youth, when self-concepts are forming, and times of academic or professional transition for adults, when decisions about persisting may be influenced unconsciously by feelings of belonging. Our finding that the implicit gender-science stereotypes of adults in science, while quite variable, do not depend on proportion of same-sex peers in the environment suggests that adults' implicit science self-concepts may also have little to do with gender ratios. That is, if Greenwald et al.'s (2002) balanced identity theory of implicit cognition is correct, the pattern of stability we have found for implicit science-is-male stereotypes should also hold for science self-concepts. Women in science, whether first-year collegians with a STEM major or PhD scientists, tend to have relatively weak implicit science-is-male stereotypes and can be expected to have strong implicit science self-concepts, regardless of gender proportions in the environment. Long-term longitudinal research is still lacking on these questions, but our results, combined with other evidence about critical junctures in the STEM pipeline, suggest that resources will likely be most fruitfully invested in studies beginning with children.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00415/abstract>

References

- Association of American Medical Colleges. (2013). *Total Graduates by U.S. Medical School and Sex, 2009–2013*. Available online at: <https://www.aamc.org/download/321532/data/2013factstable27-2.pdf>
- Association of American Medical Colleges. (2014). *Age of Applicants to U.S. Medical Schools at Anticipated Matriculation by Sex and Race/Ethnicity, 2013–2014 and 2014–2015*. Available online at: <https://www.aamc.org/download/321468/data/factstable6.pdf>
- Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173
- Ceci, S. J., and Williams, W. M. (2007). *Why Aren't More Women in Science? Top Researchers Debate the Evidence*. Washington, DC: American Psychological Association. doi: 10.1037/11546-000
- Ceci, S. J., and Williams, W. M. (2010). Sex differences in math-intensive fields. *Curr. Dir. Psychol. Sci.* 19, 275–279. doi: 10.1177/0963721410383241
- Ceci, S. J., and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3157–3162. doi: 10.1073/pnas.1014871108
- Chang, M. J., Sharkness, J., Hurtado, S., and Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *J. Res. Teach. Sci.* 51, 555–580. doi: 10.1002/tea.21146

- Cheryan, S. (2012). Understanding the paradox in math-related fields: why do some gender gaps remain while others do not? *Sex Roles* 66, 184–190. doi: 10.1007/s11199-011-0060-z
- Crowley, K., Callanan, M. A., Tenenbaum, H. R., and Allen, E. (2001). Parents explain more often to boys than to girls during shared scientific thinking. *Psychol. Sci.* 12, 258–261. doi: 10.1111/1467-9280.00347
- Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x
- Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: the stereotype inoculation model. *Psychol. Inq.* 22, 231–246. doi: 10.1080/1047840X.2011.607313
- Dasgupta, N., and Asgari, S. (2004). Seeing is believing: exposure to counterstereotypic women leaders and its effect on automatic gender stereotyping. *J. Exp. Soc. Psychol.* 40, 642–658. doi: 10.1016/j.jesp.2004.02.003
- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *J. Exp. Soc. Psychol.* 37, 443–451. doi: 10.1006/jesp.2000.1464
- Diekmann, A. B., and Eagly, A. H. (2000). Stereotypes as dynamic constructs: women and men of the past, present, and future. *Pers. Soc. Psychol. Bull.* 26, 1171–1188. doi: 10.1177/0146167200262001
- Eagly, A. H., and Steffens, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *J. Pers. Soc. Psychol.* 46, 735–754. doi: 10.1037/0022-3514.46.4.735
- Eagly, A. H., Wood, W., and Diekmann, A. (2000). “Social role theory of sex differences and similarities: a current appraisal,” in *The Developmental Social Psychology of Gender*, eds T. Eckes and H. M. Trautner (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 123–174.
- Eccles, J. S. (2007). “Where are all the women? Gender differences in participation in physical science and engineering,” in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, eds S. J. Ceci and M. W. Williams (Washington, DC: American Psychological Association), 199–210.
- Elliott, R., Strenta, A. C., Adair, R., Matier, M., and Scott, J. (1995). *Non-Asian Minority Students in the Science Pipeline at Highly Selective Institutions*. Final Report, NSF Grant RED 93 53 821.
- Galdi, S., Cadinu, M., and Tomasello, C. (2014). The roots of stereotype threat: when automatic associations disrupt girls' math performance. *Child Dev.* 85, 250–263. doi: 10.1111/cdev.12128
- Gallagher, A. M., and Kaufman, J. C. (2005). *Gender Differences in Mathematics: An Integrative Psychological Approach*. New York, NY: Cambridge University Press.
- Gawronski, B., and Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychol. Bull.* 132, 692–731. doi: 10.1037/0033-2909.132.5.692
- Greenwald, A. G., and Banaji, M. R. (1995). A unified theory of implicit attitudes, beliefs, self-esteem and self-concept. *Psychol. Rev.* 102, 4–27. doi: 10.1037/0033-295X.102.1.4
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., and Mellott, D. S. (2002). A unified theory of implicit attitudes, beliefs, self-esteem and self-concept. *Psychol. Rev.* 109, 3–25. doi: 10.1037/0033-295X.109.1.3
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216. doi: 10.1037/0022-3514.85.2.197
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., and Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychol. Sci. Public Interest* 8, 1–51. doi: 10.1111/j.1529-1006.2007.00032.x
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York, NY: Wiley. doi: 10.1037/10628-000
- Higher Education Research Institute. (2014). *The American Freshman: National Norms 2014*. Available online at: <http://heri.ucla.edu/monographs/TheAmericanFreshman2014.pdf>
- Inzlicht, M., and Ben-Zeev, T. (2000). A threatening intellectual environment: why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychol. Sci.* 11, 365–371. doi: 10.1111/1467-9280.00272
- Kaminski, D., and Geisler, C. (2012). Survival analysis of faculty retention in science and engineering by gender. *Science* 335, 864–866. doi: 10.1126/science.1214844
- Kenny, D. A. (2014). *Mediation*. Available online at: <http://davidakenny.net/cm/mediate.htm#DIA> (Accessed Jan 18, 2015).
- Kiefer, A. K., and Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes: a prospective study of female college students. *Psychol. Sci.* 18, 13–18. doi: 10.1111/j.1467-9280.2007.01841.x
- Lane, K. A., Goh, J. X., and Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles* 66, 220–234. doi: 10.1007/s11199-011-0036-z
- Lee, Y., Jussim, L., and McCauley, C. R. (1995). *Stereotype Accuracy: Toward Appreciating Group Differences*. Washington, DC: American Psychological Association.
- Martin, D., von Oertzen, T., Smyth, F., Melcher, T., and Mitrea, I. (2013). “Implicit math gender stereotypes changing over time: An application of latent growth curve modeling,” in *Poster Presented at the 25th Annual Meeting of the Association for Psychological Science* (Washington, DC).
- Miller, D. I., Eagly, A. H., and Linn, M. C. (in press). Women's representation in science predicts national gender-science stereotypes: evidence from 66 nations. *J. Educ. Psychol.* doi: 10.1037/edu0000005
- Miller, D. I., and Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends Cogn. Sci.* 18, 37–45. doi: 10.1016/j.tics.2013.10.011
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16474–16479. doi: 10.1073/pnas.1211286109
- National Academies of Science. (2007). *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Washington, DC: The National Academies Press.
- National Science Foundation, National Center for Science and Engineering Statistics. (2011a). *Science and Engineering Degrees: 1966–2008*. Detailed Statistical Tables NSF 11-316, Arlington, VA. Available online at: <http://www.nsf.gov/statistics/nsf11316/>
- National Science Foundation, National Center for Science and Engineering Statistics. (2011b). *Women, Minorities and Persons with Disabilities in Science and Engineering: 2011*. NSF 11-309, Arlington, VA.
- Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002a). Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. doi: 10.1037/0022-3514.83.1.44
- Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002b). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dyn.* 6, 101–115. doi: 10.1037/1089-2699.6.1.101
- Nosek, B. A., Greenwald, A. G., and Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Pers. Soc. Psychol. Bull.* 31, 166–180. doi: 10.1177/0146167204271418
- Nosek, B. A., Hawkins, C. B., and Frazier, R. S. (2012). “Implicit social cognition,” in *Handbook of Social Cognition*, eds S. Fiske and C. N. Macrae (New York, NY: Sage), 31–53.
- Nosek, B. A., and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *Am. Educ. Res. J.* 48, 1124–1154. doi: 10.3102/0002831211410683
- Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., et al. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *Eur. Rev. Soc. Psychol.* 18, 36–88. doi: 10.1080/10463280701489053
- Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10593–10597. doi: 10.1073/pnas.0809921106
- Ratliff, K. A., and Nosek, B. A. (2010). Creating distinct implicit and explicit attitudes with an illusory correlation paradigm. *J. Exp. Soc. Psychol.* 46, 721–728. doi: 10.1016/j.jesp.2010.04.011
- Robertson, K. F., Smeets, S., Lubinski, D., and Benbow, C. P. (2010). Beyond the threshold hypothesis: even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences

- matter for career choice, performance, and persistence. *Curr. Dir. Psychol. Sci.* 19, 346–351. doi: 10.1177/0963721410391442
- Smyth, F. L., and McArdle, J. J. (2004). Ethnic and gender differences in science graduation at selective colleges with implications for admission policy and college choice. *Res. High. Educ.* 45, 353–381. doi: 10.1023/B:RIHE.0000027391.05986.79
- Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. A. (2011). STEMing the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 100, 255–270. doi: 10.1037/a0021385
- Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a metaanalysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364
- Summers, L. H. (2005). *Remarks Presented at NBER Conference on Diversifying the Science and Engineering Workforce*. Cambridge, MA. Available online at: <http://www.president.harvard.edu/speeches/2005/nber.html> (Accessed April 21, 2005).
- Tai, R. H., Liu, C. Q., Maltese, A. V., and Fan, X. T. (2006). Planning early for careers in science. *Science* 312, 1143–1144. doi: 10.1126/science.1128690
- Xie, Y., and Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.
- Conflict of Interest Statement:** Nosek is an officer and Smyth is a consultant of Project Implicit, Inc., a non-profit organization that includes in its mission “To develop and deliver methods for investigating and applying phenomena of implicit social cognition, including especially phenomena of implicit bias based on age, race, gender or other factors.”
- Copyright © 2015 Smyth and Nosek. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: a 30-year analysis

David I. Miller^{1*} and Jonathan Wai²

¹ Department of Psychology, Northwestern University, Evanston, IL, USA

² Talent Identification Program, Duke University, Durham, NC, USA

Edited by:

Stephen J. Ceci, Cornell University, USA

Reviewed by:

Catherine Rieggle-Crumb, University of Texas, USA

Matthew A. Cannady, University of California, Berkeley, USA

*Correspondence:

David I. Miller, Department of Psychology, Northwestern University, Annenberg Hall, Suite 162, 2120 Campus Drive, Evanston, IL 60208, USA
e-mail: dmiller@u.northwestern.edu

For decades, research and public discourse about gender and science have often assumed that women are more likely than men to “leak” from the science pipeline at multiple points after entering college. We used retrospective longitudinal methods to investigate how accurately this “leaky pipeline” metaphor has described the bachelor's to Ph.D. transition in science, technology, engineering, and mathematics (STEM) fields in the U.S. since the 1970s. Among STEM bachelor's degree earners in the 1970s and 1980s, women were less likely than men to later earn a STEM Ph.D. However, this gender difference closed in the 1990s. Qualitatively similar trends were found across STEM disciplines. The leaky pipeline metaphor therefore partially explains historical gender differences in the U.S., but no longer describes current gender differences in the bachelor's to Ph.D. transition in STEM. The results help constrain theories about women's underrepresentation in STEM. Overall, these results point to the need to understand gender differences at the bachelor's level and below to understand women's representation in STEM at the Ph.D. level and above. Consistent with trends at the bachelor's level, women's representation at the Ph.D. level has been recently declining for the first time in over 40 years.

Keywords: doctoral education, STEM education, gender differences, STEM persistence, retrospective methods

INTRODUCTION

For three decades, research and public discourse about gender differences in academic science have often focused on the “leaky pipeline” metaphor (Berryman, 1983; Alper, 1993). According to this metaphor, women are more likely than men to leave science at multiple time points from the beginning of college through academic tenure. Scholars from diverse fields have proposed how specific factors such as cognitive abilities, discrimination, and interests can explain these gender differences in opting out (Ceci et al., 2009). These interlocking factors could collectively cause “leaks” at various segments in the science pipeline and therefore lead to an underrepresentation of women among science Ph.D. holders and faculty. In this way, the leaky pipeline metaphor has explicitly and implicitly served as a core theoretical foundation for several explanations regarding the underrepresentation of women in science, technology, engineering, and mathematics (STEM) fields.

We investigated how accurately the leaky pipeline metaphor has described the bachelor's to Ph.D. STEM pipeline in the U.S. since the 1970s. During this time frame, women's representation in STEM fields has dramatically increased. For instance, women earned 19% of the U.S.'s bachelor's degrees in chemistry in 1966, but earned 48% of them in 2013¹. The increase in women's representation at the Ph.D. and assistant professorship levels has also been dramatic (Ceci et al., 2014). Given this rapid change over time, it is especially worth considering whether the leaky pipeline metaphor (1) was empirically supported in the past, and

(2) continues to be empirically supported today. Current inaccuracies in this metaphor could constrain and potentially prompt revision of diverse theories about current gender differences in STEM fields. Improving such conceptual models could also help policy makers target when and where to allocate limited resources for increasing gender diversity in STEM fields.

Recent research has found some current inadequacies of the leaky pipeline metaphor (Cannady et al., 2014; Miller et al., under review). For instance, plugging leaks in the pipeline from the beginning of college to the bachelor's degree would fail to substantially increase women's representation among U.S. undergraduates in physical science, technology, engineering, and mathematics (pSTEM) fields² (excluding life science and social science). Women currently earn 25% of pSTEM bachelor's degrees in the U.S., and equalizing gender differences in undergraduate pSTEM retention would only increase this percentage to 27% (Miller et al., under review).

Other research has found large gender differences in opting out exist only in some STEM fields, but not others. For instance, the percentage of women among academic biologists substantially declines from receiving a biology Ph.D. to applying for tenure-track positions at Research I institutions; this decline suggests a leaky academic pipeline for female biologists (National Research Council [NRC], 2010). However, such declines are counterintuitively far smaller in the more male-dominated fields of physics and engineering (National Research Council [NRC], 2010). This

¹ <http://ncesdata.nsf.gov/webcaspar/>

² Prior research on gender diversity in STEM has often focused on these pSTEM fields because women are especially underrepresented in them (Ceci et al., 2014).

evidence and related studies have indicated that, when describing academic transitions after the Ph.D., the leaky pipeline metaphor is less accurate for the more male-dominated STEM fields – the fields for which the metaphor was originally intended (see Ceci et al., 2014 for a review).

We contribute to this research on persistence in STEM fields by investigating men's and women's transition from undergraduate to graduate education. During this formative period, students start to develop identities as scientists and engineers capable of independently producing scientific knowledge and technological innovations (Herzig, 2004). Several scholars have suggested that women face more challenges than men in completing this transition. Such challenges could include gender discrimination from academic mentors (Moss-Racusin et al., 2012; Milkman et al., 2014), male advantages on gatekeeper mathematics and science tests (Wai et al., 2010; Lakin and Gambrell, 2014), concerns about raising young children (Williams and Ceci, 2012), and support from peers and family (Herzig, 2004). Collectively, these diverse challenges could present themselves at many points between earning a bachelor's and Ph.D. degree, including choosing and then applying to graduate school, getting accepted, choosing the graduate school and mentor, completing coursework, developing research ideas and professional relationships, completing research projects, writing the Ph.D. thesis, and defending the thesis.

As described above, various factors at multiple time points could compel women to leave STEM fields at higher rates during the transition from the bachelor's to the Ph.D. degree. However, empirically investigating such gender differences is methodologically challenging, especially because (1) few students pursue a Ph.D. after earning the bachelor's and (2) the time in between the bachelor's and Ph.D. can often exceed a decade (National Science Board [NSB], 2014). These challenges make prospective longitudinal studies exceptionally expensive, considering the large sample sizes and long time intervals needed.

Consequently, “[s]tudies of sex differences in Ph.D. completion are hampered by a lack of data” (Ceci et al., 2014, p.99). For instance, few research studies have systematically investigated gender differences in Ph.D. completion using representative samples (though see Xie and Killewald, 2012 for persistence rates ~1–2 years after the bachelor's). Prior studies have instead used students from self-selected fellowship programs (Bowen and Rudenstine, 1992; Myers and Pavel, 2011) or non-representative groups of institutions (Zwick, 1991; Council of Graduate Schools, 2008; Ampaw and Jaeger, 2011).

Other studies (Ceci et al., 2014; Gillen and Tanenbaum, 2014) have used population-level data to compare artificial cohorts of bachelor's and Ph.D. degree earners (e.g., compare the percentage of women among physics bachelor's degree earners in a given year and then among physics Ph.D. earners 8 years after). However, these studies make somewhat restrictive assumptions about the artificial cohorts. For example, these methods assume that students do not switch fields between the bachelor's and Ph.D. degree and that students take similar amounts of time between the bachelor's and Ph.D. degree. Hence, results even from these population-based studies could be strengthened and extended with alternate methods.

To help overcome these prior limitations, we used nationally representative samples and *retrospective* methods to investigate gender differences in the bachelor's to Ph.D. STEM pipeline in the U.S. since the 1970s. As with all retrospective studies, the relevant events (e.g., earning of bachelor's and Ph.D. degrees) had already occurred at the time of the survey; participants simply recalled their prior educational histories. This retrospective design allowed us to investigate changes in STEM persistence over three decades – a unique advantage of a retrospective, compared to prospective, longitudinal design.

Supplementing these retrospective analyses, cross-sectional analyses investigated how gender differences in three other characteristics (career goals, employment status, and family outcomes) varied across cohorts of bachelor's degree holders. These supplemental analyses helped provide clues about why bachelor's to Ph.D. persistence rates might have changed over time.

MATERIALS AND METHODS

OVERVIEW

For this study, the term *STEM persistence rate* refers to the percentage of students who earned a Ph.D. in a particular STEM field (e.g., engineering) among students who had earlier received bachelor's degrees in that same field. We estimated persistence rates separately by field of study (e.g., engineering vs. physical science), bachelor's degree cohort (e.g., 1980s vs. 1990s), and gender. These rates were estimated by two sets of numbers: (1) numbers of students who earned a bachelor's degree in a particular field during a certain time frame and (2) numbers of those students who also later earned a Ph.D. in that same field. We used two national probability samples to estimate these two sets of numbers: National Survey of College Graduates (NSCG) and Survey of Doctoral Recipients (SDR).

SAMPLES

The 2010 NSCG sample ($n = 77,188$) provided estimates for numbers of bachelor's degree earners. The NSCG's target population was college graduates living in the U.S. in 2010 under 76 years old who were not institutionalized (Fecso et al., 2012). The 2010 SDR sample ($n = 31,462$) provided estimates for numbers of Ph.D. earners. The SDR's target population was a subpopulation of NSCG's target population who also earned a Ph.D. from a U.S. institution in a science, engineering, or health field³. Although the NSCG sample could have also provided estimates for numbers of Ph.D. earners, the SDR sample provided more precise estimates given its exclusive focus on Ph.D. earners.

ANALYZED VARIABLES

In both the NSCG and SDR surveys, participants were asked to recall their educational histories (e.g., the field of study and year of their first bachelor's degree). Although retrospective studies such as ours can often have various recall biases (e.g., students misremembering how interested they were in science as children), it is unlikely that participants systematically misremembered concrete details such as what year they earned their first bachelor's degree. These educational histories, participants' demographics, and probability survey weights formed the basis for our analyses.

³<http://www.nsf.gov/statistics/srvydoctoratework/>

Officials at the National Science Foundation created the survey weights to adjust for unequal sampling probabilities and non-response bias (Finamore et al., 2011). All the variables analyzed were available in the public-use versions of the 2010 NSCG and SDR surveys, which can be downloaded from the National Science Foundation's website⁴. Lists of the analyzed variables and R analysis scripts are available in the supplemental materials for this paper.

DEFINITION OF STEM FIELDS

We separated the category of "STEM" into five major subcategories as defined by the National Science Foundation's classification system: (1) computer and mathematical science, (2) engineering, (3) life science, (4) physical science, (5), and social science (National Science Board [NSB], 2014). We also estimated persistence rates for pSTEM fields as a collective whole (categories #1, #2, and #4), given the focus on these fields in prior research on gender diversity in STEM (e.g., Riegle-Crumb et al., 2012).

ESTIMATING PERSISTENCE RATES

We divided participants into cohorts based on field of study and year of the first bachelor's degree (e.g., individuals who earned their first bachelor's degree in engineering during 1976–1980). The NSCG sample provided estimates on the size of these cohorts. The SDR sample provided estimates on the numbers of Ph.D. holders within these cohorts (e.g., individuals who earned their first bachelor's degree in engineering during 1976–1980 and who also earned an engineering Ph.D. before 2010). For any particular cohort, the persistence rate was estimated by dividing the sum of relevant SDR survey weights (i.e., the number of Ph.D. holders) by the sum of corresponding NSCG survey weights (i.e., the number of bachelor's degree holders).

Analyses were restricted to U.S. citizens; the high proportion of international students among U.S. Ph.D. earners, but not bachelor's degree earners, could have artificially inflated estimates of persistence rates (Xie and Killewald, 2012). Analyses included cohorts of students who earned their first bachelor's degree between the years 1971–2000. These cohorts were divided into 5-year intervals (e.g., 1971–1975, 1976–1980, etc.) to increase sample sizes for individual cohorts and thus reduce fluctuations due to noise. See **Table 1** for sample sizes.

ESTIMATING STANDARD ERRORS

The 2010 NSCG survey used a complex two-phase sampling design in which individuals for the NSCG were sampled from respondents to the 2009 American Community Survey. As such, traditional approaches for estimating SEs in survey research (e.g., analytical formulas, jackknife replicates) are no longer appropriate. We therefore contacted the National Center for Science and Engineering Statistics and obtained custom SEs for this study's specific estimates. These SEs were estimated using successive difference replication, which is appropriate for such two-phase sampling designs (White, 2014; Opsomer et al., under review). The sample design for the SDR survey was less complex and we therefore used

Table 1 | Sample sizes by cohort, field of bachelor's degree, and gender.

Cohort	Field of bachelor's degree				
	Engineering	Life science	Math/computer science	Physical science	Social science
National Survey of College Graduates					
'71–5	1073	543	275	347	1034
	40	271	146	109	825
'71–75	1116	584	250	407	839
	127	469	162	117	851
'81–85	1553	426	373	468	653
	284	376	239	186	777
'86–90	1483	323	564	306	695
	292	377	330	167	886
'91–95	1361	411	439	264	782
	329	440	275	135	1285
'95–00	1197	487	487	250	731
	338	659	270	195	1296
Survey of Doctorate Recipients					
'71–75	297	482	137	492	516
	18	212	61	87	346
'71–75	361	538	154	475	395
	34	315	47	135	341
'81–85	584	423	159	559	309
	94	291	63	195	355
'86–90	497	377	174	406	314
	120	338	62	189	407
'91–95	392	396	151	378	331
	132	390	90	182	558
'95–00	259	420	106	303	254
	120	425	66	180	428

Blue entries refer to men, and red entries refer to women.

Source: 2010 National Survey of College Graduates and 2010 Survey of Doctoral Recipients.

standard "equivalent sample size" formulas to derive SEs for the SDR estimates (Potthoff et al., 1992).

ACCOUNTING FOR THE LENGTH OF TIME BETWEEN DEGREES

By restricting analyses to cohorts in the year 2000 and before, we allow for at least 10 years between when students earned their first bachelor's degree and when the surveys were conducted in 2010. Nevertheless, estimates especially for the last cohort (1996–2000) should be interpreted somewhat cautiously because a non-trivial proportion of students may earn Ph.D.'s after 2010.

The time in between the first bachelor's degree and Ph.D. degree can be long. For instance, among U.S. citizens earning pSTEM Ph.D.'s in the U.S. between 2000 and 2010, the time in between degrees exceeded 10 years in 26% of cases, 15 years in 11% of

⁴<http://sestat.nsf.gov/datadownload/>

cases, and 20 years in 7% of cases⁵. Given this long time between degrees, bachelor's to Ph.D. persistence rates were likely somewhat underestimated especially among later cohorts (e.g., those who earned first bachelor's degree in 1996–2000).

For these reasons above, we conducted additional analyses that compared persistence rates across cohorts based on the same length of time after the first bachelor's degree. For instance, we compared persistence rates for the 1986–1990 cohort based on Ph.D.'s earned by 2000 with the persistence rates for the 1996–2000 cohort based on Ph.D.'s earned by 2010. Such additional analyses effectively control for the confound between cohort and length of time after the bachelor's degree.

ACCOUNTING FOR SAMPLE RESTRICTIONS

In both the NSCG and SDR samples, the target populations were restricted to non-institutionalized individuals living in the U.S. in 2010 aged 75 years old or younger. These restrictions on the target populations likely had only modest effects on our estimates. For instance, the restriction to non-institutionalized populations likely had little influence because of low incarceration rates among college-educated populations (Lochner and Moretti, 2004). The age restriction might have modestly influenced estimates especially for the oldest cohort (1971–1975). The age restriction, for instance, would have excluded individuals who earned their first bachelor's degree in 1971 past the age of 36 years. However, few students in the U.S. earn bachelor's degrees past the age of 36 years. For instance, only 3% of pSTEM bachelor's degrees in 1993 were awarded to students older than 36 years⁶. Finally, the restriction to individuals living in the U.S. likely had small effects on our estimates because few U.S. bachelor's degree earners move outside the U.S. after college graduation. For instance, less than 1% of pSTEM bachelor's degree holders in 1993 moved outside the U.S. by the year 2003⁵.

INTERACTIVE WEBSITE

We made an interactive website⁷ of our results to help interested readers inspect the effects of alternate analytic decisions (e.g., effects of using an alternate grouping of STEM fields or including non-U.S. citizens). All code to make this interactive website is also available in the supplemental materials.

RESULTS

RESULTS FROM RETROSPECTIVE METHODS

Among students earning pSTEM bachelor's degrees in the 1970s and 1980s, women were 0.6–0.7 times as likely as men to later earn a pSTEM Ph.D. (Figure 1). However, this gender difference closed in the 1990s. Gender differences in persistence rates were statistically significant for cohorts in the 1970s and 1980s (all p s < 0.0005), but not in the 1990s (both p s > 0.60). See Table 2 for count estimates that were used for calculation of persistence rates and SE for gender differences in persistence rates.

⁵These estimates were calculated from the NSCG sample. Variables to compute the exact number of years between the first bachelor's degree and Ph.D. degree were not available in the public-use SDR dataset.

⁶These estimates were based on our own analysis of the 1993 Baccalaureate and Beyond study, which can be analyzed on the PowerStats website for the National Center for Educational Statistics (<http://nces.ed.gov/datalab/powerstats/>).

⁷<http://d-miller.shinyapps.io/bachelorsPHD>

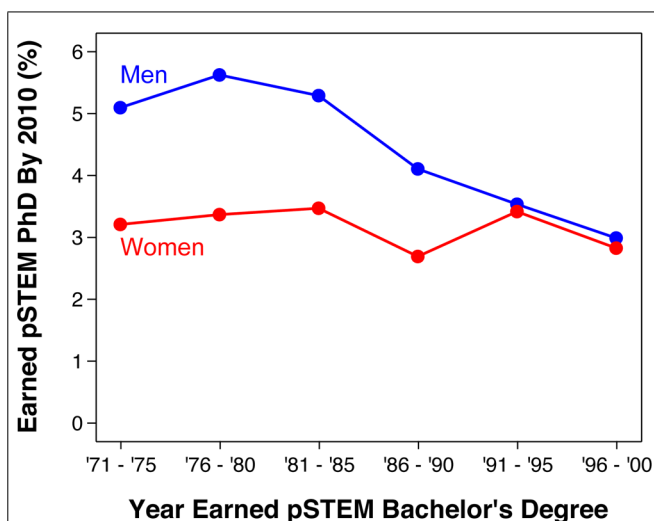


FIGURE 1 | Bachelor's to Ph.D. STEM persistence rates by gender and bachelor's degree cohort (excluding life and social science). Rates especially for the last cohort (1996–2000) should be interpreted cautiously because a non-trivial proportion of students may earn Ph.D.'s in the future. Source: 2010 National Survey of College Graduates and 2010 Survey of Doctoral Recipients.

As shown in Figure 2, similar results were found when disaggregating pSTEM fields (engineering, mathematics/computer science, physical science). Life science also showed a similar recent convergence between men and women. Social science had male advantages in persistence rates among cohorts in the 1970s, small non-significant female advantages in the early 1980s, and little to no gender differences since the late 1980s. Reasons for these convergences among cohorts in the 1990s varied across disciplinary fields (e.g., sometimes the convergence was driven by declines in men's rates or increases in women's rates, or both). No gender difference in persistence rates was significant for these 1990s cohorts (all p s > 0.19, except $p = 0.054$ for the 1991–1995 mathematics/computer science cohort). See Supplementary Table 1 for count estimates and SE across disaggregated fields.

As discussed earlier (see Accounting for the Length of Time between Degrees), one concern about these results was that cohort was confounded with the length of time after the bachelor's degree. The analyses shown in Figure 3 controlled for this confound by comparing pSTEM persistence rates over time using the same length of time after the bachelor's degree. As shown, results were qualitatively similar compared to Figure 1: male advantages in pSTEM persistence rates were found in earlier cohorts in the 1970s, but not in later cohorts in the 1990s. Results were similarly unchanged for the groupings of STEM fields shown in Figure 2; see the interactive website for detailed results⁷.

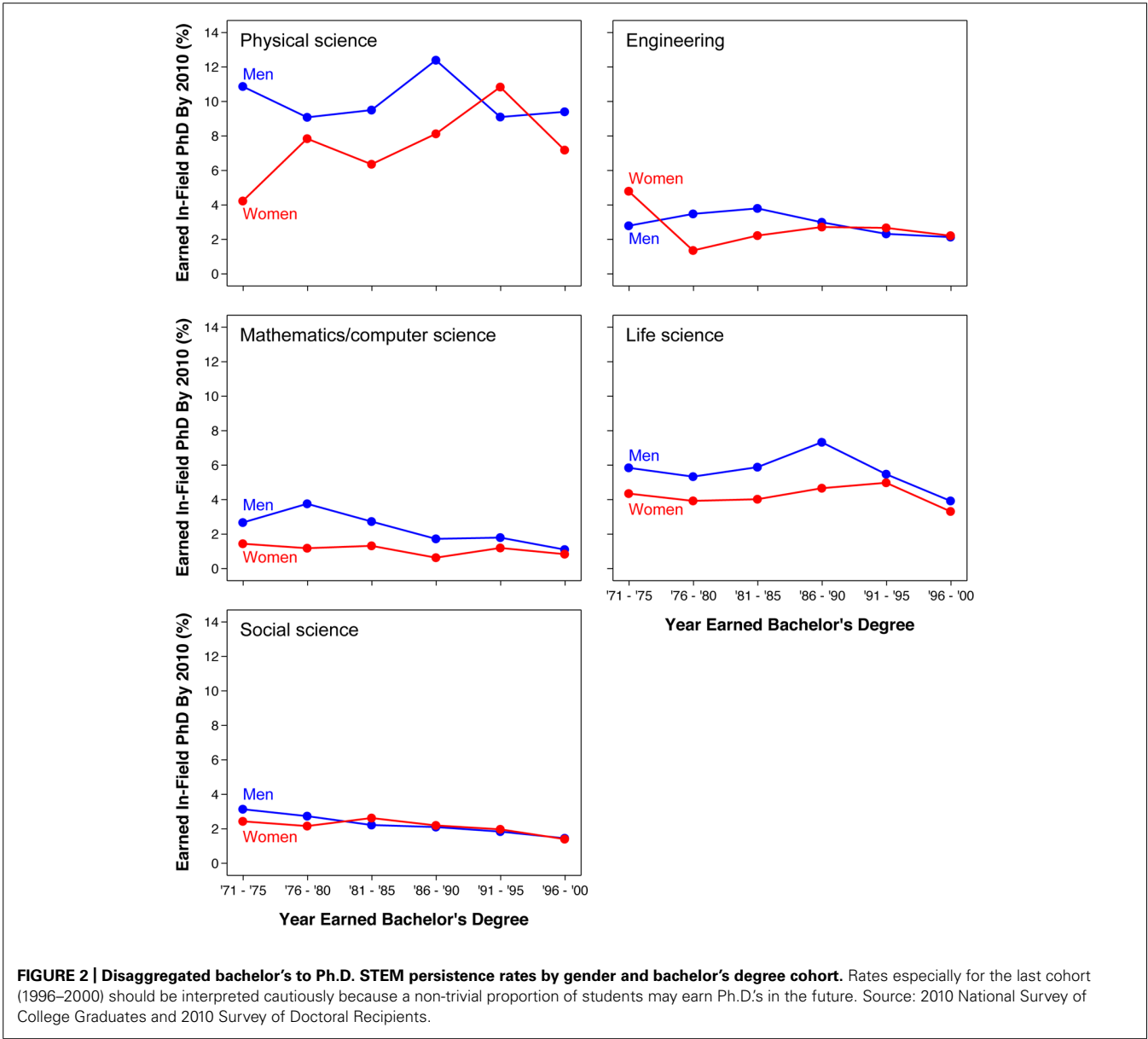
RESULTS FROM CROSS-COHORT COMPARISONS

The results for persistence rates help to explain the continual increases in women's representation among STEM Ph.D. holders. As shown in Figure 4, women earned less than 3% of the U.S.'s pSTEM Ph.D.'s in 1966, but earned 27% of them in 2012. This increase in women's representation at the Ph.D. level has been

Table 2 | Count estimates and gender differences in persistence rates for pSTEM fields.

Cohort	Count estimates				Persistence rates		
	Bachelor's		Ph.D.				
	Female	Male	Female	Male	Difference	SE	<i>p</i>
'71–'75	77450	454251	2477	23098	1.89	0.51	<0.001
'76–'80	114323	448251	3838	25163	2.26	0.46	<0.001
'81–'85	185007	646401	6401	34129	1.82	0.37	<0.001
'86–'90	238176	635401	6385	26020	1.41	0.30	<0.001
'91–'95	174975	541446	5960	19085	0.12	0.34	0.727
'96–'00	196555	472214	5535	14065	0.16	0.31	0.604

"Difference" refers to the percentage point difference in men's minus women's persistence rate. *ps* < 0.05 are bolded.



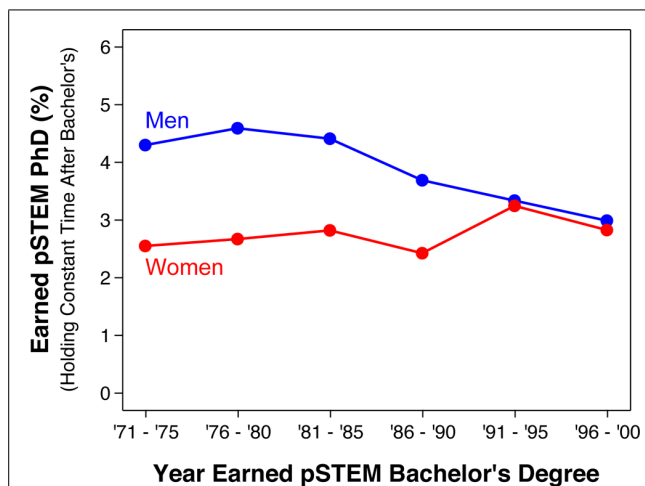


FIGURE 3 | Bachelor's to Ph.D. STEM persistence rates by gender and bachelor's degree cohort awarded (excluding life and social science), holding constant the length of time after the first bachelor's degree.

For instance, rates for the 1996–2000 cohort were based on Ph.D.s earned by 2010, rates for the 1991–1995 cohort were based on Ph.D.s earned by 2005, rates for the 1986–1990 cohort were based on Ph.D.s earned by 2000, and so on. Source: 2010 National Survey of College Graduates and 2010 Survey of Doctoral Recipients.

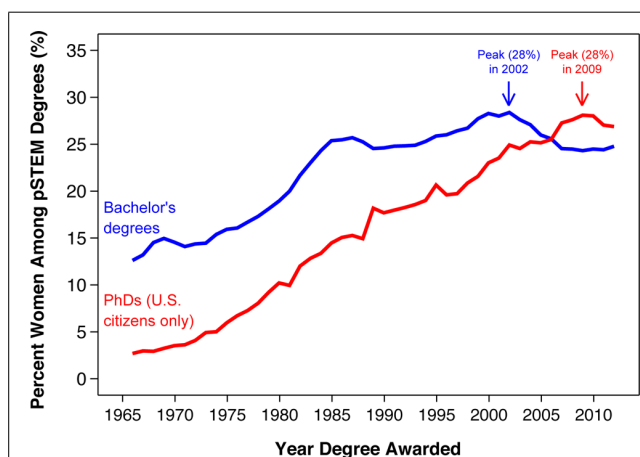


FIGURE 4 | Women's representation among STEM bachelor's and Ph.D. degree earners by year of degree awarded (excluding life and social science). Ph.D. data after 2012 are not available. Source: WebCASPAR Integrated Science and Engineering Resource Data System (2014).

steady over these four decades, and qualitatively similar trends are found across all STEM disciplines (e.g., life science and physics; Ceci et al., 2014).

Our results indicate that changes over time at the Ph.D. level can be attributed to two major factors: (1) the increase of women's representation at the bachelor's level among cohorts in the early 1970s to mid 1980s (Figure 4), and (2) the narrowing of gender differences in persistence rates among bachelor's degree cohorts in the early 1980–1990s (Figures 1 and 2).

Although women's representation among pSTEM Ph.D. holders has been continually increasing since the 1970s, this trend may

not continue in the future for two major reasons: (1) women's representation among STEM bachelor's degree holders has been declining since 2000 (Figure 4), and (2) gender differences in STEM persistence rates have already closed (Figures 1 and 2). Current data indicate that women's representation at the Ph.D. level has started to decline for the first time in over 40 years. The percent women among pSTEM Ph.D.s awarded to U.S. citizens peaked at 28% in 2009 and has been declining ever since (Figure 4). Women would need to overtake men in bachelor's to Ph.D. STEM persistence rates to reverse this trend. Otherwise, this trend will likely continue over the next few years.

CHANGES IN OTHER CHARACTERISTICS AMONG BACHELOR'S DEGREE HOLDERS

To help place the bachelor's to Ph.D. persistence findings in context, we investigated changes in other characteristics (e.g., career goals) for our focal bachelor's degree holder population (i.e., U.S. citizens who earned a pSTEM bachelor's degree during the years 1971–2000). These supplemental analyses used the NSCG data to characterize this focal population. Results revealed some stable gender differences regarding career goals (e.g., men were more likely to rate *salary* as a very important factor when thinking about a job, and women were more likely to rate *contribution to society* as very important) and employment outcomes (e.g., men were more likely than women to be working in 2010, working women were more likely than working men to be precollege teachers) in this focal population. However, these gender differences generally showed no consistent increase or decrease across cohorts; see the interactive website for complete, detailed results. Hence, gender differences in these characteristics likely cannot explain the cross-cohort changes observed for bachelor's to Ph.D. persistence rates. Gender differences in having children and getting married were small across the cohorts and therefore also likely cannot explain the changes in persistence rates.

DISCUSSION

The leaky pipeline metaphor has partially explained historical gender differences in the U.S., but it no longer describes current gender differences in the bachelor's to Ph.D. transition in STEM. Remarkably, these recent convergences in persistence rates were found in all major groups of STEM fields (i.e., engineering, life science, mathematics, and computer science, social science, and physical science). These results align with and extend other recent studies that used alternate methods to investigate the bachelor's to Ph.D. transition (Ceci et al., 2014; Gillen and Tanenbaum, 2014). Our study helps to place these recent convergences in historical context; some of the mixed results in prior literature likely reflect genuine change over time (e.g., Zwick, 1991; Herzig, 2004; Council of Graduate Schools, 2008; Gillen and Tanenbaum, 2014). Male Ph.D. holders still outnumber female Ph.D. holders by approximately three to one in pSTEM fields. However, our results indicate that gender differences in bachelor's to Ph.D. persistence rates no longer help to explain this male overrepresentation. In fact, women's representation in pSTEM is now *higher* at the Ph.D. than bachelor's level.

Reasons for the convergences in persistence rates remain unclear. Sometimes the convergence was driven by declines in

men's rates (e.g., in mathematics/computer science), increases in women's rates (e.g., in physical science), or both (e.g., in engineering). Our results helped eliminate potential hypotheses for these changes over time. For instance, convergences in persistence rates were likely unrelated to changes in some characteristics among bachelor's degree holders. For instance, among pSTEM bachelor's degree holders, gender differences in career goals and employment outcomes generally showed no consistent increase or decrease across the relevant cohorts. To explore other hypotheses, future research should investigate how changes in doctoral education might help account for the changes in persistence rates. For instance, gender diversity initiatives at the graduate level might have helped increase women's rate of persisting in a doctoral program after entering graduate school.

THEORETICAL IMPLICATIONS: GENERAL

Regardless of reasons why persistence rates might have changed over time, the recent convergences between women's and men's rates inform theories about women's current representation in STEM. The convergences in rates may seem surprising given the multitude of factors that could cause women to leave STEM fields at higher rates than men (e.g., gender discrimination, gender-science stereotypes, right tail differences in cognitive abilities, or a combination of multiple factors). As reviewed in the introduction, many theories of women's underrepresentation in STEM have often either explicitly or implicitly assumed that women are less likely than men to persist and pursue doctoral training in STEM. However, our results indicate that this foundational assumption may have been accurate in the past, but is no longer accurate.

One possible interpretation of recent gender similarity is that some factors could create male advantages in persistence rates (e.g., factors such as discrimination, right tail ability differences), but other factors create female advantages. For instance, self-selection among STEM undergraduates might create female advantages at the graduate level. As Hunt (2012, p. 1) hypothesized, various obstacles that female STEM undergraduates may face could "cause women entering science and engineering to be more positively selected for interest and aptitude than their male counterparts." In other words, given the obstacles for female STEM undergraduates, only women with the strongest interest and aptitude for STEM would successfully earn STEM bachelor's degrees.

This self-selection hypothesis, however, does not seem to align with the changes over time that we found. If anything, obstacles facing female STEM undergraduates were likely more extreme earlier in time when fewer women were earning STEM bachelor's degrees (Ceci et al., 2014) and gender-science stereotypes were stronger (Miller et al., under review). According to this hypothesis, self-selection among female STEM undergraduates might have then been stronger in the 1970s and 1980s, meaning that those women might have been especially likely to pursue doctoral education. However, our results contradict this prediction because male advantages in persistence rates were larger earlier in time.

Another possible interpretation of these results is that various factors such as gender discrimination may not contribute

substantially to current gender differences in bachelor's to Ph.D. STEM persistence rates. In the following section, we consider this possibility for two specific factors especially relevant to doctoral education: gender discrimination among academic mentors and right tail differences in cognitive abilities. Of course, these possibilities are not necessarily mutually exclusive with the one discussed earlier (i.e., some factors such as self-selection create female advantages in persistence rates).

THEORETICAL IMPLICATIONS: GENDER DISCRIMINATION

Two recent field experiments found that STEM faculty's biases favor male students on average (Moss-Racusin et al., 2012; Milkman et al., 2014). For instance, in one nationally representative sample, STEM faculty ignored emails more frequently from prospective female graduate students than prospective male graduate students (Milkman et al., 2014). Such biases might therefore create a "leaky pipeline" for female STEM college majors by discouraging them from applying to graduate school or impeding their academic progress once in graduate school. However, our results do not agree with this basic prediction. Men and women now persist at roughly equal rates in STEM fields between the bachelor's and Ph.D. degree, despite evidence of pro-male biases among academic mentors (Moss-Racusin et al., 2012; Milkman et al., 2014).

One possibility is that STEM faculty's biases favor male students on average, but women overcome these biases by persisting at equal rates compared to men. Some empirical evidence supports this hypothesis. For instance, Milkman et al.'s (2014) study found biases favoring White males in nearly all academic fields. However, the size of the gender discriminatory gap in a particular academic field did not predict the representation of women in that field at the Ph.D. or faculty level. For instance, compared to non-STEM faculty, STEM faculty were not particularly biased against women. In fact, gender discrimination against White females was stronger among faculty in health fields than in the male-dominated fields of computer science and engineering (Milkman et al., 2014, **Figure 1B**). These results demonstrate that stronger pro-male biases do not necessarily translate to a lower representation of women at the Ph.D. or faculty level (see Ceci et al., 2014 for discussion of other related studies about gender discrimination in academic science).

These considerations above should not be used to discount the crucial importance of accurately assessing and changing gender biases in science. Gender discrimination can negatively affect many potential outcomes other than the numeric percentage of women in STEM fields. For instance, gender discrimination may cause some women to not feel respected or limit women's equitable access to resources (e.g., equal salaries). Such realizations raise the question of whether diversity initiatives in STEM should focus more on increasing the representation of particular groups (e.g., women, non-Asian racial minorities) or improving the quality of experiences for members of those groups.

THEORETICAL IMPLICATIONS: COGNITIVE ABILITIES

Some scholars have proposed that gender differences in mathematics and science reasoning performance might partially contribute to the underrepresentation of women in academic science

(Benbow, 1988; Wai et al., 2010; Ceci et al., 2014). Although males and females often perform similarly on standardized mathematics tests on average, males are overrepresented in the right tail of mathematics and science reasoning performance (e.g., top 5% of performance or higher; Wai et al., 2010; Miller and Halpern, 2014). These right tail differences could be related to women's representation among STEM Ph.D. holders because such individuals disproportionately come from this right tail of performance (Lubinski and Benbow, 2006) and individual differences in SAT-Mathematics scores at age 12 predict later differences in earning STEM Ph.D.'s even within the top 1% of performance (Wai et al., 2005).

Although these right tail differences could be relevant to women's representation at the Ph.D. level, they are likely less relevant to representation at the bachelor's level. Many students successfully earn STEM bachelor's degrees without being in the right tail of mathematics performance (though see Hsu and Schombert, 2010 for additional discussion). For instance, only one-fifth (18%) of STEM bachelor's degree holders in 2008⁸ had received a SAT-Mathematics score above 700⁹ in high school. Right tail differences in mathematics performance therefore likely do not substantially contribute to women's representation in STEM fields at the bachelor's level; longitudinal studies support this hypothesis (Riegle-Crumb et al., 2012).

If extremely high mathematics performance is required at the Ph.D. but not bachelor's level, right tail differences in performance might be especially important for persisting from the bachelor's to Ph.D. degree. For instance, low scores on challenging gatekeeper tests (e.g., GRE-Mathematics) could directly reduce students' likelihood of being admitted to a STEM Ph.D. program. Hence, if right tail gender differences contribute to women's representation among STEM Ph.D. holders, one might predict these right tail differences do so through their influence on persisting from the bachelor's to Ph.D. degree.

Our results, however, do not agree with this basic prediction because men and women now persist at equal rates from the bachelor's to Ph.D. in various STEM disciplines. Moreover, in pSTEM fields, men and women also persist at equal rates in the academic pipeline past the Ph.D. (see Ceci et al., 2014 for a review). For instance, in physical science and engineering fields, female and male Ph.D. holders are equally likely to earn assistant professorships (National Research Council [NRC], 2010) and academic tenure (Ginther and Kahn, 2009; Kaminski and Geisler, 2012; National Research Council [NRC], 2010). Hence, despite males outnumbering females in the top fraction of math and science reasoning performance (Wai et al., 2010; Miller and Halpern, 2014), males and females now persist at equal rates in the most intellectually challenging segments of the academic pipeline (e.g., earning Ph.D.'s and academic tenure) in some of the most math-intensive STEM fields (e.g., physical science and engineering). These results

suggest women's underrepresentation among high mathematics performers might be a more minor factor contributing to women's underrepresentation among pSTEM Ph.D. holders and faculty.

WHAT IS THE RIGHT METAPHOR?

Our research shows that the leaky pipeline metaphor is a dated description of gender differences in the transition between earning bachelor's and Ph.D. degrees in STEM in the U.S. Related prior research indicates that the pipeline metaphor is also misleading for some other academic pathways in STEM. For instance, the metaphor fails to acknowledge the multiple entry points into STEM prior to the bachelor's degree. Many students successfully earn STEM bachelor's degrees despite not having traveled the traditional STEM "pipeline." In one nationally representative study, 39% of STEM bachelor's degree earners had not intended to enter STEM when asked in either 8th or 12th grade of secondary education (Cannady et al., 2014). Moreover, in another nationally representative sample, female science and engineering majors were likely to have joined STEM for the first time during college than have entered college already intending to major in STEM (Xie and Shauman, 2003). As Cannady et al. (2014, pp. 447–448) argued, such results make "the pipeline an ill-suited frame to understand STEM identity formation, particularly for women and underrepresented minorities."

Nevertheless, the pipeline metaphor may be an apt description of academic transitions after the Ph.D. Academic pathways are considerably more rigid after the Ph.D. degree than before the bachelor's degree. For instance, transitioning from a humanities Ph.D. to physical science tenure-track position would be nearly impossible without a physical science Ph.D.; the analogous transition between high school and college would be relatively open. However, as reviewed earlier, the post-Ph.D. academic pipeline leaks more women than men only in some STEM fields such as life science, but surprisingly not the more male-dominated fields of physical science and engineering (Ceci et al., 2014).

Although the leaky pipeline metaphor may aptly describe the post-Ph.D. pathways in life science, the metaphor as a whole may nevertheless do more harm than good. It is an inappropriate description for nearly all other academic pathways in STEM. Moreover, the metaphor may even burden some women who leave academic science with a sense of guilt about being "leaks" in the pipeline. The Twitter user *biochembelle* wrote that, "Sometimes I think the way we talk about women in science and the 'leaky pipeline' makes more guilt for women to follow paths they want" (post on 26 August 2013). This sentiment resonated with other users who replied with tweets such as, "Every time someone talks about the 'leaky pipeline,' they are calling me a 'drip'" (user *elakdawalla*, tweet also on 26 August 2013). See **Figure 5** for other selected responses or the associated blog post by *biochembelle* for additional discussion¹⁰. These examples are of course anecdotal, but help illustrate how some individuals are personally impacted by the metaphor.

Along with other researchers (e.g., Xie and Shauman, 2003; Cannady et al., 2014), we propose replacing the metaphor of a

⁸These estimates were based on our own analysis of the 2008 Baccalaureate and Beyond study, which can be analyzed on the PowerStats website for the National Center for Educational Statistics (<http://nces.ed.gov/datalab/powerstats/>).

⁹SAT-Mathematics scores of 700+ corresponded to the top ~5–10% of performance among the SAT test-taking population. See http://professionals.collegeboard.com/profdownload/sat_percentile_ranks_2008_males_females_total_group_math.pdf

¹⁰<http://biochembelle.com/2013/08/28/the-pipeline-isnt-leaky/>



singular pipeline with a network of multiple pathways into and out of STEM. This concept of pathways more accurately describes the multiple entry points into STEM prior to the bachelor's degree. The idea also more positively portrays women who leave academic science as women pursuing other potentially fulfilling goals outside of academia (Webb et al., 2002; Xie and Killewald, 2012). And perhaps most importantly, this reconceptualization provides policy makers and educators with a wider range of strategies for increasing diversity in STEM. For instance, compared to “plugging the leaky pipeline” for female STEM undergraduates, equalizing gender differences in rates of joining STEM from non-STEM fields would more potently increase women's representation among STEM bachelor's degrees (Miller et al., under review).

LIMITATIONS

The retrospective methods we used extended and complemented prior methods for studying gender differences in bachelor's to Ph.D. persistence rates (e.g., Ceci et al., 2014;

Gillen and Tanenbaum, 2014). However, the use of retrospective methods also limited our inferences to a subpopulation of degree earners who were included in the surveys' target populations: non-institutionalized adults aged 75 years or younger living in the U.S. in 2010. However, as discussed earlier (see Accounting for Sample Restrictions), this limitation likely did not introduce large biases into our results. Our conclusions were also restricted to the U.S., though the retrospective methods that we used could be applied to any other nation with appropriate data.

Cohort was confounded with the length of time after the bachelor's degree (see Accounting for the Length of Time between Degrees). As such, estimated persistence rates may have been modestly underestimated especially among the later cohorts; a non-trivial proportion of those students may earn Ph.D.'s after when the surveys were conducted in 2010. Importantly, however, our cross-cohort results were qualitatively similar when holding constant the length of time after the bachelor's degree (e.g., see Figure 3). Hence, this limitation cannot account for the changes in gender differences over time. Future changes in persistence rates are unclear. Gender gaps could reemerge in the future, although our data offer no particular indication that they will reemerge.

Our methods revealed changes in persistence rates over time, but not in other outcomes relevant to doctoral education (e.g., performance in graduate school, subjective experiences of students). As we discussed earlier, future research should investigate whether some factors such as gender discrimination affect these other outcomes without substantially affecting persistence rates. Finally, continuing to investigate *why* persistence rates changed over time would also be invaluable.

CONCLUSION

Overall, these results and supporting literature point to the need to understand gender differences at the bachelor's level and below to understand women's representation in STEM at the Ph.D. level and above. Women's representation in computer science, engineering, and physical science (pSTEM) fields has been decreasing at the bachelor's level during the past decade. Our analyses indicate that women's representation at the Ph.D. level is starting to follow suit by declining for the first time in over 40 years (Figure 2). This recent decline may also cause women's gains at the assistant professor level and beyond to also slow down or reverse in the next few years. Fortunately, however, pathways for entering STEM are considerably diverse at the bachelor's level and below. For instance, our prior research indicates that undergraduates who join STEM from a non-STEM field can substantially help the U.S. meet needs for more well-trained STEM graduates (Miller et al., under review). Addressing gender differences at the bachelor's level could have potent effects at the Ph.D. level, especially now that women and men are equally likely to later earn STEM Ph.D.'s after the bachelor's.

ACKNOWLEDGMENTS

This research was supported by a NSF Graduate Research Fellowship awarded to D. I. Miller (DGE-0824162). We thank John Finamore at the National Center for Science and Engineering Statistics for his generous efforts to help derive SEs for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2015.00037/abstract>

REFERENCES

- Alper, J. (1993). The pipeline is leaking women all the way along. *Science* 260, 409–411. doi: 10.1126/science.260.5106.409
- Ampaw, F. D., and Jaeger, A. J. (2011). Understanding the factors affecting degree completion of doctoral women in the science and engineering fields. *New Dir. Inst. Res.* 152, 59–73. doi: 10.1002/ir.409
- Benbow, C. P. (1988). Sex differences in mathematical reasoning ability among the intellectually talented: their characterization, consequences, and possible explanations. *Behav. Brain Sci.* 11, 169–183.
- 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*. New York, NY: Rockefeller Foundation.
- Bowen, W. G., and Rudenstine, N. L. (1992). *In Pursuit of the Ph.D.* Princeton, NJ: Princeton University Press. doi: 10.1515/9781400862474
- Cannady, M. A., Greenwald, E., and Harris, K. N. (2014). Problematising the STEM pipeline metaphor: is the STEM pipeline metaphor serving our students and the STEM workforce? *Sci. Educ.* 98, 443–460. doi: 10.1002/sce.21108
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412s
- Council of Graduate Schools. (2008). *Ph.D. Completion and Attrition: Analysis of Baseline Demographic Data from the Ph.D. Completion Project*. Washington, DC: Council of Graduate Schools.
- Fecso, R. S., Frase, M. J., and Kannankutty, N. (2012). *Using the American community survey as the sampling frame for the national survey of college graduates*. Working Paper NCSES 12-201. Arlington, VA: National Science Foundation.
- Finamore, J., Cohen, S., Hall, D., Walker, J., and Jang, D. (2011). "NSCG estimation issues when using an ACS-based sampling frame," in *JSM Proceedings* (Alexandria, VA: ASA), 4408–4422.
- Gillen, A., and Tanenbaum, C. (2014). *Broadening Participation in Stem Graduate Education*. Washington, DC: American Institutes for Research.
- Ginther, D. K., and Kahn, S. (2009). "Does Science Promote Women? Evidence from Academia 1973–2001," in *Science and Engineering Careers in the United States*, eds R. B. Freeman and D. F. Goroff (Chicago, IL: University of Chicago Press for NBER).
- Herzig, A. H. (2004). Becoming mathematicians: women and students of color choosing and leaving doctoral mathematics. *Rev. Educ. Res.* 74, 171–214. doi: 10.3102/00346543074002171
- Hsu, S. D. H., and Schombert, J. S. (2010). *Nonlinear Psychometric Thresholds for Physics and Mathematics*. Available at: <http://arxiv.org/abs/1011.0663>
- Hunt, J. (2012). *Why do women leave science and engineering?* Available at: <http://www.econstor.eu/bitstream/10419/67325/1/727519247.pdf>
- Kaminski, D., and Geisler, C. (2012). Survival analysis of faculty retention in science and engineering by gender. *Science* 335, 864–866. doi: 10.1126/science.1214844
- Lakin, J. M., and Gambrell J. L. (2014). Sex differences in fluid reasoning: manifest and latent estimates from the cognitive abilities test. *J. Intell.* 2, 36–55. doi: 10.3390/jintelligenc2020036
- Lochner, L., and Moretti, E. (2004). The effect of education on crime: evidence from inmates, arrests, and self-reports. *Am. Econ. Rev.* 94, 155–189. doi: 10.1257/000282804322970751
- Lubinski, D., and Benbow, C. P. (2006). Study of Mathematically Precocious Youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspect. Psychol. Sci.* 1, 316–345. doi: 10.1111/j.1745-6916.2006.00019.x
- Milkman, K. L., Akinola, M., and Chugh, D. (2014). *What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations*. Available at: <http://ssrn.com/abstract=2063742>
- Miller, D. I., and Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends Cogn. Sci.* 18, 37–45. doi: 10.1016/j.tics.2013.10.011
- Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16474–16479. doi: 10.1073/pnas.1211286109
- Myers, C. B., and Pavel, D. M. (2011). Underrepresented students in STEM: the transition from undergraduate to graduate programs. *J. Divers. High. Educ.* 4, 90–105. doi: 10.1037/a0021679
- National Science Board [NSB]. (2014). *Science and Engineering Indicators 2014*. Arlington, VA: National Science Foundation.
- National Research Council [NRC]. (2010). *Gender Differences at Critical Transitions in the Careers of Science, Engineering, and Mathematics Faculty*. Washington, DC: National Academies Press.
- Potthoff, R. F., Woodbury, M. A., and Manton, K. G. (1992). "Equivalent sample size" and "equivalent degrees of freedom" refinements for inference using survey weights under superpopulation models. *J. Am. Stat. Assoc.* 87, 383–396. doi: 10.1080/01621459.1992.10475218
- Riegle-Crumb, C., King, B., Grodsky, E., and Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *Am. Educ. Res. J.* 49, 1048–1073. doi: 10.3102/0002831211435229
- Wai, J., Cacchio, M., Putallaz, M., and Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: a 30-year examination. *Intelligence* 38, 412–423. doi: 10.1016/j.intell.2010.04.006
- Wai, J., Lubinski, D., and Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youths: an age 13 to age 33 longitudinal study. *J. Educ. Psychol.* 97, 484–492. doi: 10.1037/0022-0663.97.3.484
- Webb, R. M., Lubinski, D., and Benbow, C. P. (2002). Mathematically facile adolescents with math/science aspirations: new perspectives on their educational and vocational development. *J. Educ. Psychol.* 94, 785–794. doi: 10.1037/0022-0663.94.4.785
- White, M. (2014). *Generalized Variance Parameters for the 2010 National Survey of College Graduates: New Cohort and Combined Cohort*. Suitland, MD: U.S. Bureau of the Census, Demographic Statistical Methods Division, Health Surveys and Supplements Branch.
- Williams, W. M., and Ceci, S. J. (2012). When women scientists choose motherhood. *Am. Sci.* 100, 138–145. doi: 10.1511/2012.95.138
- Xie, Y., and Killewald, A. A. (2012). *Is American Science in Decline?* Cambridge, MA: Harvard University Press. doi: 10.4159/harvard.9780674065048
- Xie, Y., and Shauman, K. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.
- Zwick, R. (1991). *Differences in Graduate School Attainment Patterns Across Academic Programs and Demographic Groups*. Princeton, NJ: Educational Testing Service.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 October 2014; paper pending published: 13 November 2014; accepted: 08 January 2015; published online: 17 February 2015.

Citation: Miller DI and Wai J (2015) The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: a 30-year analysis. *Front. Psychol.* 6:37. doi: 10.3389/fpsyg.2015.00037

This article was submitted to *Developmental Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Miller and Wai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The role of school performance in narrowing gender gaps in the formation of STEM aspirations: a cross-national study

Allison Mann^{1*}, Joscha Legewie² and Thomas A. DiPrete¹

¹ Department of Sociology, Columbia University, New York, NY, USA

² Department of Humanities and Social Sciences, New York University, New York, NY, USA

Edited by:

Stephen J. Ceci, Cornell University, USA

Reviewed by:

Shulamit Kahn, Boston University School of Management, USA
Andrew Penner, University of California, Irvine, USA

*Correspondence:

Allison Mann, Department of Sociology, Columbia University, Knox Hall – MC9649, 606 W. 122nd St., New York, NY 10027, USA
e-mail: alm2174@columbia.edu

This study uses cross-national evidence to estimate the effect of school peer performance on the size of the gender gap in the formation of STEM career aspirations. We argue that STEM aspirations are influenced not only by gender stereotyping in the national culture but also by the performance of peers in the local school environment. Our analyses are based on the Program for International Student Assessment (PISA). They investigate whether 15-year-old students from 55 different countries expect to have STEM jobs at the age of 30. We find considerable gender differences in the plans to pursue careers in STEM occupations in all countries. Using PISA test scores in math and science aggregated at the school level as a measure of school performance, we find that stronger performance environments have a negative impact on student career aspirations in STEM. Although girls are less likely than boys to aspire to STEM occupations, even when they have comparable abilities, boys respond more than girls to competitive school performance environments. As a consequence, the aspirations gender gap narrows for high-performing students in stronger performance environments. We show that those effects are larger in countries that do not sort students into different educational tracks.

Keywords: education, school context, gender inequality, careers in science, technology, engineering, mathematics, cross-cultural research

1. INTRODUCTION

A growing body of research documents the under-representation of women in science, technology, engineering, and mathematics (STEM) occupations and fields of study (Xie and Shauman, 2003; Eccles, 2007; Ceci and Williams, 2011; Ceci et al., 2014). In order to understand the sources of these differences, we need to study the formation of career aspirations in high school, because high school aspirations are strong predictors of initial college major choice and the attainment of a Bachelor degree in STEM fields (Tai et al., 2006; Morgan et al., 2013; Legewie and DiPrete, 2014a). Recent research from the United States demonstrates that the high school environment—and particularly the strength of the STEM curriculum and the gender segregation of extra-curricula activities—have a substantial impact on gender differences in plans to major in STEM fields in college (Legewie and DiPrete, 2014b). Data from South Korea suggest that single-sex schools for boys increase the level of interest in STEM fields, but single-sex schools for girls do not have a corresponding effect on the STEM aspirations of girls (Park et al., 2012). Related research finds that the school context also plays an important role for gender differences in educational performance (Legewie and DiPrete, 2012).

In this study, we contribute to research on the role of the school context for the gender gap in STEM aspirations, by examining the impact that peer ability has on gender differences in the formation of STEM orientations across 55 countries. Researchers

have found that the school performance environment has a negative impact on student career aspirations in science (Marsh and Hau, 2003; Shen and Tam, 2008; Nagengast and Marsh, 2012). There is strong theoretical justification for expecting a gender difference in the responsiveness to the school performance climate. High performance in the environment arguably raises the level of competition. It has important implications for the self evaluation of performance, which in turn shapes the aspirations for different fields of study. Indeed, the self evaluation of performance plays a central role in previous research. With respect to women and STEM fields, (Correll, 2001) argues that gender status beliefs lead boys to evaluate their math and science abilities more highly than girls do, either because girls believe that the relative competency assessment is valid or because girls expect that others will accept the ranking as valid. Correll (2001) found that the undervaluation by girls of their own competence in math had behavioral consequences in that it discouraged them from pursuing quantitative coursework and fields of study. Other researchers have reached similar conclusions with regard to the influence of self evaluations on course choices in high school (Marsh and Yeung, 1997; Nagy et al., 2008) and career aspirations (Eccles et al., 1999; Nagy et al., 2006; Riegle-Crumb et al., 2010; Eccles, 2011; Sikora and Pokropek, 2012).

Similar to gender status beliefs that influence performance expectations, the ability of peers in the school context provides an important reference for performance evaluations and an

important influence on the formation of STEM field aspirations. This influence is presumably twofold. First, peer ability influences the self-evaluation of math and science ability and aspirations for STEM fields directly. Second, peer ability mediates the role of performance for self-evaluation and aspirations insofar as the influence of performance on aspirations (returns to performance) varies depending on the performance of peers. Previous research on the country level supports this idea. Mann and DiPrete (unpublished manuscript) show that boys and girls have lower STEM aspirations and stronger returns to math-science performance when they live in countries with stronger overall performance levels. This finding is attributed to the higher risk of failure in more competitive environments and the concomitant need for stronger evidence that one is good at math-science before forming a STEM orientation. Mann and DiPrete (unpublished manuscript) also find that the effect on the math-science slope of a stronger math-science country environment is stronger for girls than for boys, which is linked to gender status beliefs in the national culture. If this is true, we would expect to find a similar pattern in school environments, particularly during the high school years.

The influence of peer ability most likely differs across countries. We examine these variations in a sample of 55 countries and point to the importance of tracking systems as a mediating factor for peer influence on STEM aspirations. The track into which a student is placed affects the composition of the student's peer group and provides an independent signal of the student's ability and potential. The organization of national education systems has been shown to influence student's educational aspirations in previous studies. Research shows that in relatively undifferentiated (unstructured) systems—where there are fewer tracks and a later age at first selection into tracks—peer and parent attitudes have significantly greater influences on student aspirations to complete college and to pursue high-status occupations (Buchmann and Dalton, 2002; Buchmann and Park, 2009). Furthermore, students in course tracking appear to experience the opposite patterns: although lower self assessments typically emerge in higher-performance environments, students in higher tracks have higher self assessments (Chmielewski et al., 2013). Environmental and contextual factors also have been shown to influence academic self assessments and career intentions aside from the aggregate impact of school performance or SES (Alwin and Otto, 1977; Legewie and DiPrete, 2014b). Accordingly, structural features of national and school education systems might influence the extent to which peer ability shapes educational aspirations.

2. DATA AND METHODS

Measures and sample data are from the Program for International Student Assessment (PISA). PISA is a triennial international study that tests the reading, mathematical and scientific literacy level of 15-year-old students who are still in school. The database is hierarchically structured such that students are nested within schools, and schools are nested within countries. We use the 2006 data collection, which included 57 countries. In 2006, science was the major content domain.

We restrict our sample in three ways. First, we exclude data from Qatar because the students were not asked about STEM

aspirations. Second, we exclude students in schools that have fewer than 10 students considering that we are interested in understanding school effects (about 6800 observations). Finally, we remove data from Liechtenstein because of the small number of schools (about 12 schools and 300 observations). With these restrictions, there are 55 countries, 12,846 schools, and 331,834 students in the final sample.

2.1. DEPENDENT VARIABLE: STEM ASPIRATIONS

The dependent variable is whether the student expects to have a STEM job at the age of 30. The question taken from the student questionnaire was “What kind of job do you expect to have when you are about 30 years old? Write the job title —.” The responses were coded using the International Standard Classification of Occupations. Our definition excludes some of the occupations that have been treated as STEM occupations in previous research (Kjærnsli and Lie, 2011; Sikora and Pokropek, 2012)—specifically, nursing and associate or technician level occupations—because we are interested in a measure of aspirations for STEM careers among high-performing students. In some models, we use the STEM subfields of physical sciences and life sciences as the dependent variables (always relative to those with non-STEM aspirations). The Appendix includes a detailed list of occupations for STEM fields and the breakdown between the physical and life sciences.

2.2. MATH AND SCIENCE PERFORMANCE

PISA does not contain information about student grades or other performance feedback given directly to students. We use test scores—the best measure of performance—as a proxy for all observed and unobserved performance feedback available to students. The composite math and science test scores for each student were averaged to form an individual math-science test score. We standardized the average of the math-science test scores for the students in each country; within each country the math-science test score measure has a mean of zero and a standard deviation of 1. Then, we aggregated the standardized test score measure to the school level to create a measure of the school performance environment. With these measures, we are able to identify the high- and low-performing students and schools in each country, but we obscure the relative position of students in the global sample.

2.3. DEMOGRAPHIC CHARACTERISTICS

We use demographic information about each respondent—specifically sex, immigrant status, a broad measure of socioeconomic status (ESCS)—and an indicator for whether either parent has a science-related career. PISA respondents are all 15 years old so that age is not a relevant predictor, but we do include the student's grade level relative to the modal grade for the country in which the student lives.

2.4. COUNTRY CHARACTERISTICS

We use measures of the structural features of the nation's education system as they pertain to the tracking of students between schools. We use a binary measure for whether assignment into tracks occurs before the age of 16. Countries with an early age at first selection into tracks are also countries that tend to have more programs in which 15-year old students are enrolled. Thus, as an

alternative measure of national tracking, we use the number of separate programs in which 15-year old students can be enrolled (a binary variable that measures whether this number is greater than one). Because these variables represent the same underlying concept, they are not used in the same models.

2.5. PROCEDURES

To determine whether the school context is related to the gender gap in STEM aspirations, we use regression analyses with country fixed effects and standard errors clustered on schools. We use logistic regression predicting three different dependent variables—STEM aspirations, physical science aspirations, and life science aspirations. The dependent variable was regressed onto standardized test scores, standardized school performance measures and their interaction, and gender. In addition, we include gender interactions with all performance measures and also with the background measures described above. To assess cross-national variation in the magnitude of these effects, we use hierarchical logistic regression models.

3. RESULTS

This section begins with descriptions of the sample countries in terms of our variable of interest – STEM-related aspirations. **Table 1** shows the results of the descriptive statistics—both overall and by gender. Because PISA has a complex, two-stage stratified sample design, all descriptive statistics are weighted using the student-level weights provided in the dataset to compensate for unequal selection probabilities of students.

Across the 55 countries, the average proportion of students with STEM aspirations is 22 percent, ranging from a low of about 9 percent in Montenegro to a high of about 47 percent in Colombia. The average proportion of students with life science aspirations is about 12 percent, as is the average proportion of students with physical science aspirations. These proportions mask significant variability; some countries—Mexico, Chile, Brazil, and Colombia—have 25 percent of students or more with life science aspirations, and other countries—Switzerland, the Netherlands, Austria, and Germany—have only 5–6 percent of students with

life science aspirations. The Latin American countries also have large proportions of students with physical sciences aspirations, while several European and Asian countries have very low proportions of students with physical science aspirations compared with the global average.

In most countries, we observe substantial gender differences in STEM aspirations. There is a male advantage in physical science aspirations in 52 of 55 countries (with no significant gender difference in 3 countries). There is a female advantage in life science aspirations in 48 countries, a male advantage in 1 country, and no significant gender difference in life science aspirations in 6 countries. With all STEM occupations combined (referred to as “combined STEM” below), males have an advantage in STEM aspirations in 34 countries in the study, females have an advantage in 6 countries, and there is no significant gender difference in STEM aspirations in the remaining countries. The magnitude of the gender gap in STEM aspirations varies considerably, with a 10-point difference in proportions favoring girls in Kyrgyzstan and a 16-point difference in proportions favoring boys in Chinese Taipei.

3.1. ANALYTICAL RESULTS

We begin our analysis of gender differences in STEM aspirations by pooling the students across countries and estimating three logistic regression models that predict overall STEM aspirations, physical science aspirations, and life science aspirations. Because we are interested in the average effects of school performance environments, these models use country fixed effects to condition on all observed and unobserved factors on the country level. These models use cluster robust standard errors to account for clustering on schools. **Table 2** displays the results.

Generally speaking, girls respond differently to the school performance environment than do boys. Strong environments decrease only slightly the propensity for boys to develop STEM aspirations at the mean of the individual-level performance distribution. However, the negative interaction between own performance and school performance means that strong performance environments more powerfully suppress STEM aspirations for

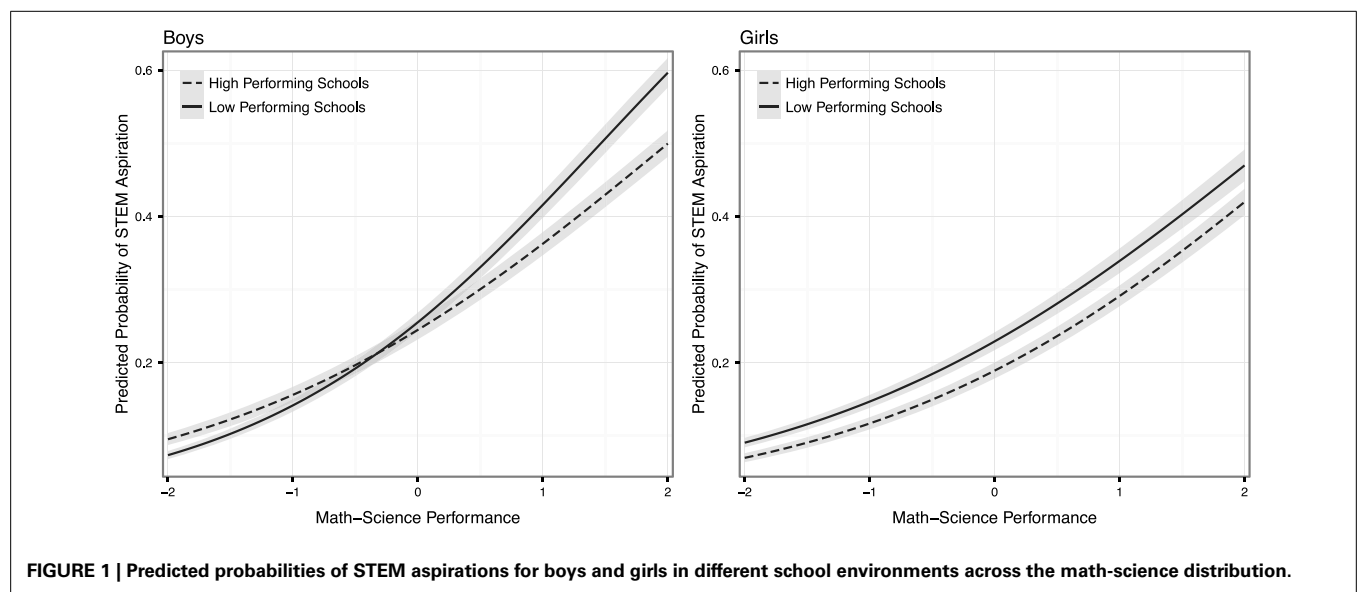
Table 1 | Nation-level descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min.	Max.	Male mean	Female mean
Own math-science (MS)	55	0	1		0.064	−0.060	
School math-science (SchMS)	55	0	0.6	0.3	0.8	−0.014	0.013
ESCS	55	−0.175	0.487	−1.434	0.823	−0.145	−0.202
Immigrant	55	0.092	0.126	0.001	0.734	0.092	0.092
Parent STEM career	55	0.072	0.030	0.007	0.139	0.074	0.070
Relative grade level	55	−0.131	0.302	−0.947	0.553	−0.167	−0.010
STEM aspirations	55	0.219	0.080	0.087	0.469	0.239	0.199
Physical science aspirations	55	0.121	0.050	0.045	0.292	0.175	0.069
Life science aspirations	55	0.127	0.067	0.046	0.344	0.097	0.151
First age of selection into tracks	54	14.176	1.901	10	17		
Number of programs	54	2.315	1.226	1	5		
No ability grouping	54	0.341	0.206	0.003	0.895	0.340	0.342
Ability grouping-some classes	54	0.456	0.243	0.033	0.918	0.455	0.458
Ability grouping-all classes	54	0.203	0.176	0.007	0.770	0.205	0.201

Table 2 | Gender differences in the effects of the local performance environment on STEM aspirations, with country fixed effects.

	Overall STEM		Physical sciences		Life sciences	
	Coef	S. E.	Coef	S. E.	Coef	S. E.
Female	−0.23***	0.01	−1.14***	0.02	0.60***	0.02
Math-science score (MS)	0.65***	0.01	0.66***	0.01	0.64***	0.01
ESCS	0.09***	0.01	0.05***	0.01	0.12***	0.01
Immigrant	0.49***	0.02	0.41***	0.03	0.55***	0.03
Parent in STEM Occup.	0.47***	0.02	0.43***	0.02	0.51***	0.02
Relative grade level	−0.04***	0.01	−0.00	0.01	−0.07***	0.01
School math-science (SchMS)	−0.03*	0.01	−0.04**	0.02	−0.02	0.03
INTERACTIONS						
SchMS × MS	−0.10***	0.02	−0.11***	0.01	−0.10***	0.02
Female × MS	−0.10***	0.01	−0.01	0.02	−0.11***	0.02
Female × SchMS	−0.11***	0.02	−0.10**	0.03	−0.13***	0.03
Female × MS × SchMS	0.12***	0.02	0.10***	0.02	0.10***	0.02
Constant	−1.29***	0.07	−2.38***	0.11	−2.05***	0.08
Number of observations	322947		284663		285972	

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

**FIGURE 1 | Predicted probabilities of STEM aspirations for boys and girls in different school environments across the math-science distribution.**

stronger-performing boys ($p < 0.001$). The negative interaction between school environment and female means that the gender gap in physical science aspirations widens in favor of boys in stronger school environments for students at the mean of the math-science distribution ($p < 0.01$), while the female advantage in life science aspirations shrinks in stronger school environments ($p < 0.001$). At the same time, however, in the aspirations model, the three-way interaction between school environment, own performance, and female is significantly positive (Female \times MS \times SchMS = 0.12, $p < 0.001$). This means that the widening gender gap in high-performance schools applies more to weaker performing students than stronger performing students. Boys have a tendency to “de-differentiate” by own performance in stronger environments. Girls show no such tendency; their tendency to

differentiate by own performance when forming STEM aspirations remains as strong in high performance environments as in low performance environments. This pattern applies both to physical science and to life science STEM aspirations. As a result, the remaining analysis focuses on gender differences in the response to performance environments for combined-STEM aspirations.

To further illustrate the gender differences in the response to school performance environments for STEM aspirations, **Figure 1** plots the predicted probabilities of having a STEM aspiration across the math-science distribution for boys and girls in schools at the 10th percentile (“low performing schools,” SchMS = −0.74) and at the 90th percentile (“high performing schools,” SchMS = 0.89) of the distribution of school

Table 3 | Predicted probabilities of STEM aspirations for boys and girls in different school environments and at different positions on the math-science distribution.

	High-performing schools		Low-performing schools	
	Boys	Girls	Boys	Girls
High-performing student	0.38	0.31	0.44	0.35
Average-performing student	0.23	0.18	0.24	0.22
Low-performing student	0.13	0.10	0.12	0.13

High-performing and low-performing schools are defined as schools in the 90th and 10th percentiles of the school MS distribution, respectively. High-, average-, and low-performing students are defined as students at the 90th, 50th, and 10th percentiles of the MS distribution, respectively.

math-science environments. **Table 3** contains the corresponding predicted probabilities of a high-, average- and low-performing student in high- and low-performing schools. **Figure 1** (and all subsequent figures) assume the “base case” (i.e., setting all independent variables to zero), which corresponds at a substantive level to a native-born student in the modal grade for the country, with average socio-economic status, parents in non-STEM occupations, and average values on test-score measures except as otherwise indicated. As **Figure 1** shows, girls have lower STEM aspirations than boys in most circumstances, but girls have an advantage relative to boys in the difference between the returns to math-science in strong performance environments and in low performance environments. This is because boys receive higher returns to math-science scores in lower performance environments than they do in higher performance environments while girls receive similar returns without regard to the strength of the school performance environment. To put it another way, the gender gap in STEM aspirations among high performing students is smaller when these students are in higher performance environments. High performance school environments provide lower costs to girls than they do to boys.

3.2. COUNTRY DIFFERENCES IN THE IMPORTANCE OF SCHOOL PERFORMANCE FOR GENDER DIFFERENCES IN STEM ASPIRATIONS

It is important to keep in mind that these results are averages across all the PISA countries and themselves mask potentially strong environmental heterogeneity. Having established the average importance of school performance environments for STEM aspirations and gender differences in the response to school performance environments, we therefore next use hierarchical models to examine heterogeneity across countries in the effects of school environments on STEM aspirations. We estimate separate models for boys and girls that use STEM aspirations as the dependent variable. Each model includes own math-science, school math-science, and their interaction, as the predictor variables, as well as controls for socio-economic status, immigrant status, having a parent with a STEM occupation, and relative grade level. Each model includes random intercepts at the country and school level and random country slopes for own math-science, school math-science, and their interaction. In these models, the “fixed”

effects are consistent with the output shown in **Table 2** (see also the first set of models in **Table 5**).

Table 4 contains the total effects for each country (including the random components). **Figure 2** displays these results graphically by presenting the male effect on the y-axis and the female effect on the x-axis, with a 45° reference line representing gender parity, for each of the four estimates of interest. As expected, the regression intercepts are larger for boys (i.e., above the 45° line) in most but not all country environments. The returns to math-science test scores are positive in all countries and are stronger for boys (i.e., above the 45° line) in most country environments. The returns to school performance environments are negative in the majority of countries, but there is a sizable minority of countries where the returns to school performance environments are positive. Many of the countries with large positive coefficients for SchMS (Austria, Croatia, Germany, Italy, Montenegro, Slovenia, and the Slovak Republic) have structural features of their national education system that facilitate tracking into homogeneous environments. Gender differences in the interaction of own math-science times school math-science ($MS \times SchMS$) favor girls in the majority of countries. This corresponds to the finding from the country fixed-effects models (**Table 2**) that the female response to own math-science performance is greater (depending on the country, increases more or decreases less) than is the male response in schools with stronger math-science environments.

As **Figure 2** (bottom right panel) shows, however, this pattern—while widely present—is not universal. While most countries are below the 45° line, a few countries are above it. **Figure 3** displays the predicted probabilities of having a STEM aspiration across the math-science distribution for boys and girls in schools at the 10th and 90th percentile of the school math-science distribution in 8 selected countries that show nation-level variability in the relative effect of school environments on STEM aspirations for boys and for girls. In Italy and Korea, girls have lower STEM aspirations than boys in the base case ($MS = 0$, $SchMS = 0$), but the relative difference in aspirations is smaller in higher-performing schools; conversely, in Japan girls have higher aspirations than boys in the base case, and the aspirations gap widens in higher-performing schools. In all three of these countries, girls have higher STEM aspirations in higher-performing schools than they do in lower-performing schools. In Italy and Japan, boys have higher math-science test score slopes. However, in Italy the male math-science slopes are smaller in higher-performing schools—that is, own performance has a bigger effect on STEM aspirations for boys in low-performing schools—but girls’ slopes do not respond to the school environment. Conversely, in Japan, boys’ math-science slopes respond very little to the school environment, but girls’ slopes increase in stronger performance environments. In Korea, girls have slightly higher returns to math-science, but those slopes do not change in the school environment, while the male slopes increase in high-performing environments¹.

¹Recall that Korea is one of the few countries in the sample in which boys have a larger slope on the interaction of individual math-science and school performance than girls.

Table 4 | Total effects of performance and performance environment on STEM aspirations, by gender.

	Intercept		Math-science slope		School M-S slope		Interaction	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
COUNTRIES WITH EARLY TRACKING (BEFORE AGE 16)								
Argentina	−1.119	−1.102	0.267	0.560	−0.172	−0.161	0.002	−0.065
Austria	−2.725	−2.625	0.477	0.598	0.665	0.828	−0.017	−0.125
Azerbaijan	−1.236	−1.374	0.316	0.352	−0.217	−0.100	−0.014	−0.155
Belgium	−2.280	−1.683	0.926	0.960	−0.087	−0.127	0.047	0.022
Bulgaria	−0.906	−0.995	0.144	0.219	−0.097	−0.195	0.033	−0.013
Chile	−0.667	−0.478	0.592	0.652	−0.074	0.084	−0.095	−0.211
Chinese Taipei	−2.268	−1.087	0.883	0.631	−0.039	−0.132	0.005	−0.081
Colombia	−0.136	0.003	0.210	0.390	−0.274	−0.314	−0.063	−0.041
Croatia	−2.434	−2.670	0.539	0.690	0.122	0.296	0.027	−0.016
Czech Republic	−2.172	−1.832	0.798	0.797	0.109	0.126	−0.098	−0.181
Estonia	−1.645	−1.609	0.365	0.597	−0.217	−0.117	0.046	−0.095
France	−2.270	−1.803	0.956	0.948	−0.082	−0.114	0.194	0.129
Germany	−2.569	−2.233	0.764	0.640	0.104	0.237	0.136	0.030
Greece	−1.595	−1.290	0.925	0.756	−0.128	0.080	−0.015	−0.173
Hong Kong-China	−2.372	−1.609	0.879	0.799	−0.198	−0.194	0.137	−0.035
Hungary	−2.147	−1.765	0.650	0.669	0.227	0.595	−0.004	−0.188
Indonesia	−0.971	−1.079	0.127	0.134	0.148	0.157	0.008	−0.032
Ireland	−2.054	−1.370	0.914	0.658	−0.239	−0.225	0.088	−0.027
Israel	−1.237	−1.363	0.587	0.553	−0.438	−0.378	−0.004	−0.034
Italy	−1.663	−1.370	0.314	0.446	0.506	0.440	−0.060	−0.144
Japan	−2.002	−2.280	0.418	0.860	0.211	0.099	0.045	0.009
Korea	−2.346	−1.532	0.755	0.723	0.081	−0.495	−0.023	0.078
Kyrgyzstan	−0.640	−1.312	−0.004	0.558	−0.391	−0.563	−0.062	−0.062
Lithuania	−1.664	−1.369	0.615	0.666	−0.128	−0.160	−0.020	−0.091
Luxembourg	−2.337	−1.868	0.678	0.752	0.294	−0.183	−0.008	−0.057
Macao-China	−2.402	−2.011	0.679	0.748	−0.018	0.006	0.069	0.028
Mexico	−0.773	−0.198	0.291	0.348	−0.101	−0.102	−0.008	−0.109
Montenegro	−2.185	−2.486	0.118	0.388	0.273	0.142	−0.019	0.003
Netherlands	−2.923	−2.764	0.930	0.905	0.322	0.070	0.052	0.124
Portugal	−1.127	−0.983	0.777	0.848	−0.235	0.009	−0.011	−0.219
Romania	−1.703	−1.575	0.342	0.753	0.072	0.382	0.022	−0.165
Russian Federation	−1.834	−1.428	0.249	0.519	−0.009	0.184	0.082	−0.181
Serbia	−2.091	−1.989	0.578	0.720	0.112	−0.033	0.000	−0.005
Slovak Republic	−2.115	−1.586	0.599	0.674	0.119	0.272	−0.074	−0.139
Slovenia	−1.690	−1.130	0.585	0.383	0.326	0.590	−0.064	−0.222
Switzerland	−2.691	−2.043	0.669	0.766	0.165	−0.058	0.021	0.000
Turkey	−1.453	−0.837	1.127	0.692	−0.356	0.003	0.121	−0.056
Uruguay	−1.003	−1.037	0.343	0.390	0.030	0.234	−0.117	−0.122
COUNTRIES WITHOUT EARLY TRACKING								
Australia	−2.007	−1.603	0.785	0.813	−0.203	−0.223	0.053	−0.054
Brazil	−0.505	−1.016	0.072	0.288	0.001	−0.166	−0.040	−0.116
Canada	−1.291	−1.315	0.605	0.728	−0.409	−0.236	0.030	−0.036
Denmark	−2.199	−2.118	0.852	0.802	−0.210	−0.168	0.078	−0.001
Finland	−2.157	−2.302	0.597	0.730	−0.241	−0.093	0.137	−0.004
Iceland	−1.215	−1.303	0.809	0.769	−0.289	−0.096	0.006	−0.056
Jordan	−0.705	−0.165	0.911	0.853	−0.401	−0.316	0.057	0.028
Latvia	−1.719	−1.489	0.420	0.574	0.002	−0.200	0.007	0.016
New Zealand	−1.819	−1.848	0.756	0.588	−0.212	−0.181	0.038	0.016
Norway	−1.745	−1.542	0.591	0.678	−0.063	−0.315	0.028	−0.027
Poland	−1.457	−1.362	0.643	0.791	−0.044	−0.291	−0.082	−0.046

(Continued)

Table 4 | Continued

	Intercept		Math-science slope		School M-S slope		Interaction	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Spain	-1.411	-1.225	0.944	0.945	-0.378	-0.239	0.098	0.037
Sweden	-2.141	-2.246	0.622	0.690	-0.331	0.068	0.142	-0.070
Thailand	-0.841	-1.037	0.780	0.894	-0.254	-0.395	0.064	0.038
Tunisia	-0.750	-0.695	0.881	0.658	-0.212	0.058	0.155	-0.044
United Kingdom	-2.378	-1.720	1.060	0.847	-0.333	-0.152	0.126	0.064
United States	-1.229	-1.237	0.445	0.699	-0.461	-0.444	0.145	-0.026

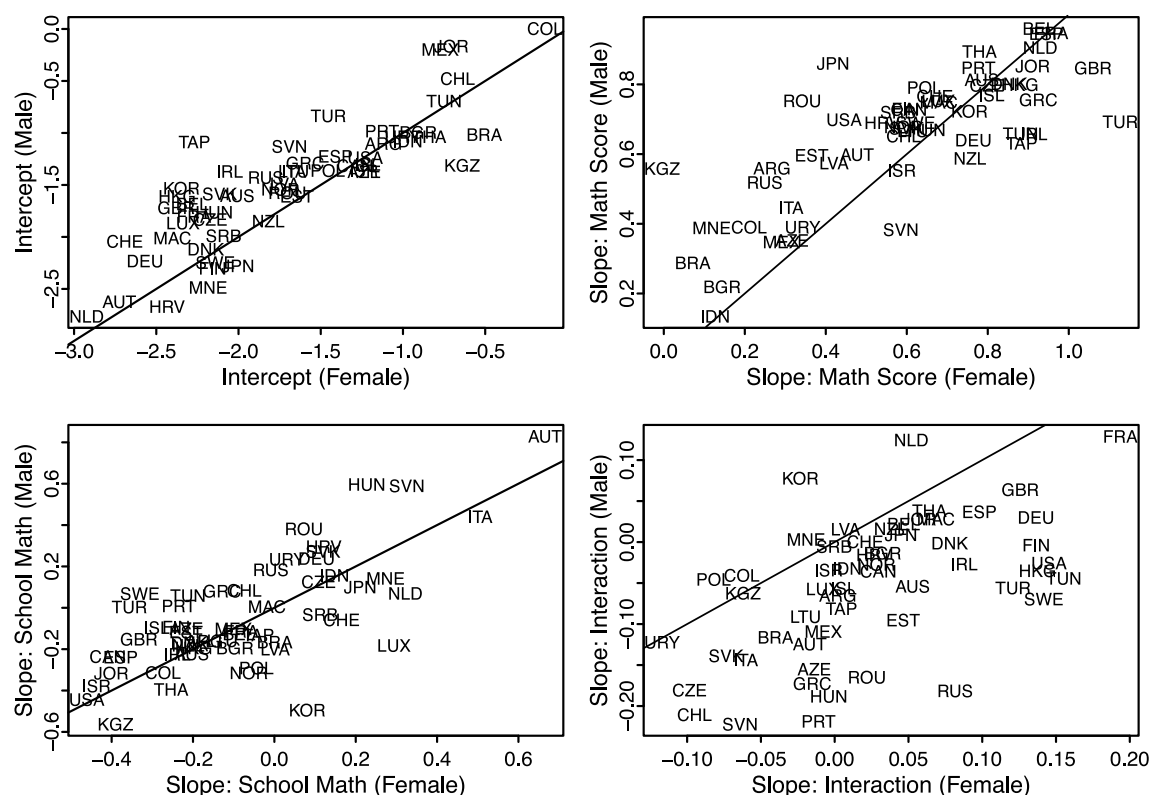


FIGURE 2 | Country random-effects estimates from models predicting STEM aspirations for boys and girls.

School environments affect STEM aspirations differently in the remaining 5 countries, which are examples of the dominant pattern found in the PISA data in the bottom right panel of **Figure 2**. In Finland, girls have higher STEM aspirations at average ability levels ($MS = 0$), especially when they are in low-performing schools; boys' math-science slopes are larger, but those returns diminish in stronger performance environments. In the United States, boys and girls have comparable aspirations at average ability levels ($MS = 0$) without regard to school environments, but (similar to Finland) boys have larger math-science slopes, with diminishing gender differences in effects in stronger performance environments (where girls' slopes converge across school performance levels and boys' slopes diverge). In Poland and Great Britain, girls have lower STEM aspirations in

the base case; however in Great Britain, the effects widen in stronger performance environments, while in Poland, they narrow. Similarly, girls' math-science slopes are larger than boys' slopes in Great Britain, while the reverse is true in Poland. In both cases, those gender differences are heightened in stronger performance environments.

To explore whether structural features of country and school education systems explain the variation in the effects of the performance environment on gender differences in STEM aspirations, we estimated a similar set of models on subsets of the data selected according to the characteristics of the school systems. **Table 5** presents the estimates from the random-effects models predicting STEM aspirations. The first set of models for boys and for girls use the full sample. The second set of models use

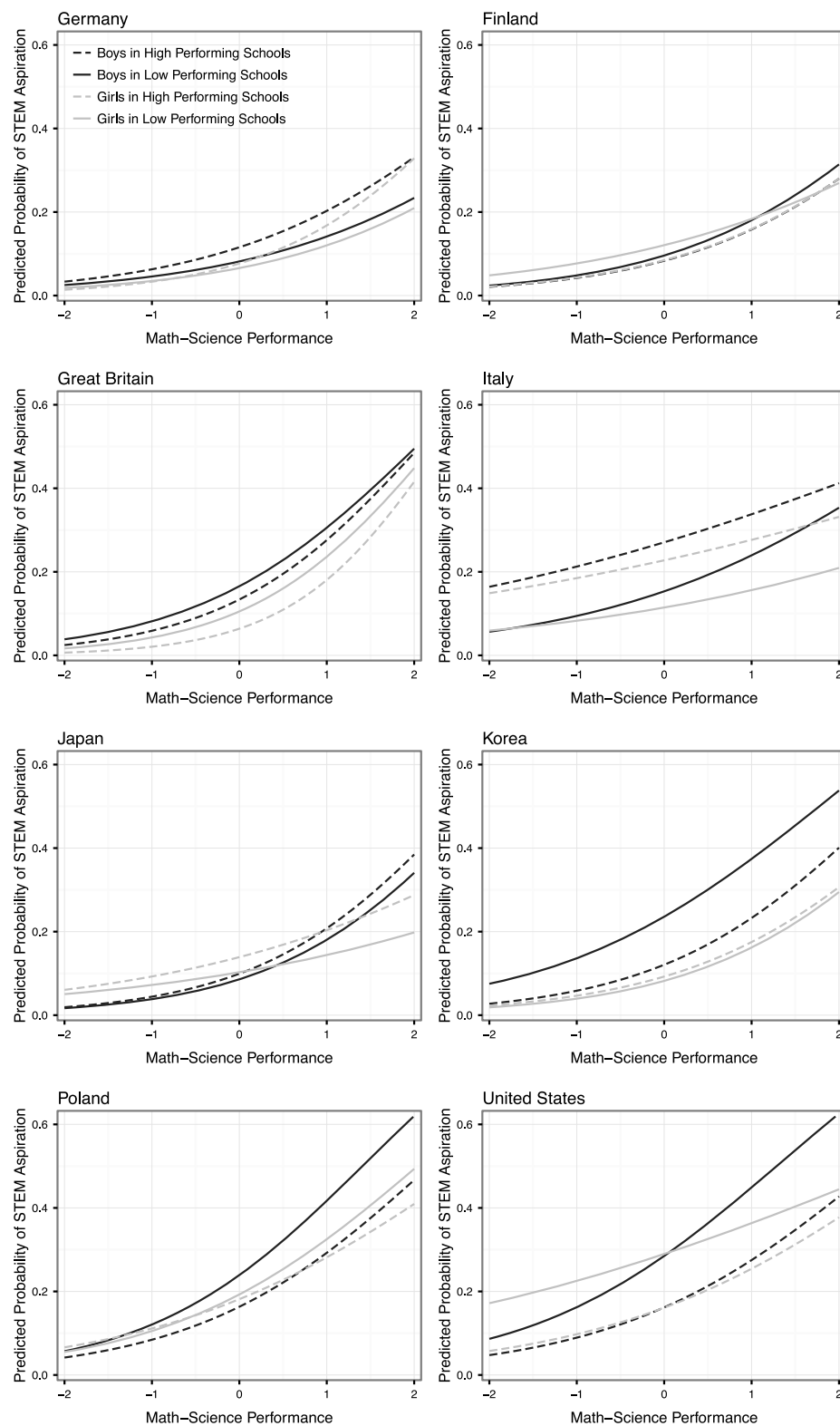


FIGURE 3 | Predicted probabilities of STEM aspirations for boys and girls in different school environments across the math-science distribution (selected countries).

Table 5 | Gender differences in the effects of the local performance environment on STEM aspirations, with country and school random effects.

	Full sample				No tracking before age 16				Tracking before age 16			
	Girls		Boys		Girls		Boys		Girls		Boys	
	Coef	S. E.	Coef	S. E.	Coef	S. E.	Coef	S. E.	Coef	S. E.	Coef	S. E.
Math-science score (MS)	0.62	0.04	0.67	0.03	0.65	0.05	0.69	0.04	0.56	0.06	0.63	0.04
ESCS	0.09	0.01	0.10	0.01	0.11	0.01	0.13	0.01	0.08	0.01	0.06	0.01
Immigrant	0.52	0.03	0.57	0.03	0.57	0.03	0.65	0.03	0.39	0.05	0.39	0.05
Parent in STEM Occup.	0.36	0.02	0.54	0.02	0.36	0.03	0.47	0.03	0.36	0.04	0.63	0.04
Relative grade level	−0.10	0.01	0.01	0.01	−0.13	0.02	0.02	0.02	−0.08	0.02	−0.01	0.02
School math-science (SchMS)	−0.06	0.04	−0.03	0.05	−0.26	0.04	−0.21	0.04	0.19	0.06	0.20	0.08
SchM × MS	0.03	0.02	−0.05	0.02	0.05	0.02	−0.06	0.02	−0.01	0.03	−0.06	0.03
Intercept	−1.75	0.09	−1.54	0.08	−1.52	0.11	−1.39	0.10	−2.00	0.13	−1.70	0.14
Error terms	Std. Dev.		Std. Dev.		Std. Dev.		Std. Dev.		Std. Dev.		Std. Dev.	
School intercept	0.51		0.50		0.44		0.43		0.61		0.58	
Country intercept	0.67		0.61		0.61		0.53		0.65		0.68	
Country MS slope	0.29		0.20		0.29		0.20		0.27		0.20	
Country SchMS slope	0.28		0.31		0.17		0.15		0.24		0.34	
Country MS × SchMS slope	0.10		0.11		0.06		0.07		0.08		0.12	
Number of observations	169457		154754		92678		82168		74867		69710	

the subset of countries that have no between-school tracking before age 16. The final set of models use the subset of countries that have between-school tracking before age 16. Looking across the columns, the most noticeable difference is the effects on students of average ability levels ($MS = 0$) of the school performance environment. In countries without tracking, STEM aspirations significantly decrease in stronger performance environments, which is consistent with a social comparison effect, while in countries with tracking, STEM aspirations significantly increase in stronger performance environments, which is consistent with a signal associated with placement in a higher track. Strong environments in the absence of a signal about a student's track placement appears to weaken student intentions to pursue a STEM career. But in the presence of a signal about track placement, strong environments (which invariably means placement in an academic track) enhance student intentions to pursue a STEM career. In models predicting science self assessments (available upon request), the relative pattern of the school performance slopes is similar; the school performance effects on self assessments are more strongly negative in less structured school environments than in environments with national tracking systems. This suggests that part of the mechanism for the effect of the performance environment on STEM aspirations runs through science self assessments².

The relative difference across genders in the interaction effect of own math-science and school math-science also is greater in countries that do not track students between schools. **Figure 4**

displays the results of models estimated separately for students in different types of national education systems; the top two panels show the results for girls and boys in countries that begin tracking students into schools before the age of 16 compared to those that do not begin tracking students into schools before age 16, and the bottom two panels show the results for girls and boys in countries that have multiple tracks into which students are assigned compared to those students in countries where only one track is possible. In general, students in low-performing schools (the solid lines in **Figure 4**) who live in countries with institutional tracking receive the lowest returns to math-science performance. Our interpretation is that the signal given to these students by their track placement crowds out the signal they are receiving from their own math-science performance in these countries. In the countries without tracking, on the other hand, the effect of the own-performance signal is relatively strong. The own-performance signal is especially strong in low-performing schools. In high performing untracked schools, the social comparison effect of high performing peers reduces the probability of STEM aspirations for high performing students. **Figure 4** also shows in the top two panels that the gender gap in STEM aspirations in favor of boys is diminished in untracked schools when these schools are high performance schools³.

²We use the science self-concept scale as a measure of self assessment. Because our models reveal no significant gender differences in the effects of peer performance on self assessments or in the returns to own math-science for self assessments, we do not report them here.

³To further explore the sources of cross-national variation in effects of performance and performance environments, we included the country Gender Gap Index (2006) in the separate models for boys and girls in tracking and non-tracking subsets of the sample. The inclusion of the GGI had no significant effect on the estimates of interest. The GGI lowers STEM aspirations to a comparatively greater extent for girls than for boys, and it reduces the variability in the country regression intercept.

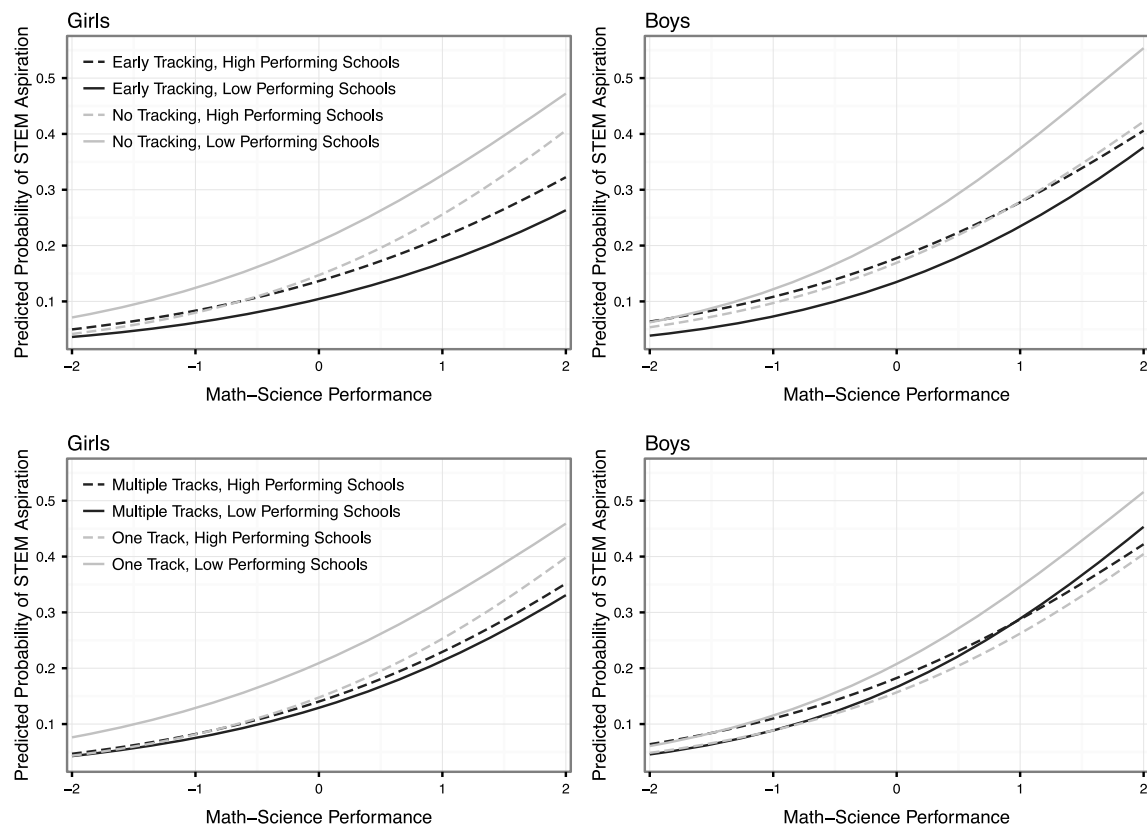


FIGURE 4 | Predicted probabilities of STEM aspirations for boys and girls in different school environments across the math-science distribution and by characteristics of national tracking system.

4. DISCUSSION

This paper examines the impact that peer ability has on gender differences in the formation of STEM orientations. Peer ability is measured by a school's math and science performance level. High performance in the environment arguably raises the level of competition. Across the 55 countries in our sample, we show that girls and boys are more likely to develop STEM orientations if they have stronger performance in math and science; yet in high performance school environments, boys and girls require stronger evidence that they are good in math and science before deciding to pursue a STEM orientation. This is consistent with the pattern for nation-level performance and STEM aspirations (Mann and DiPrete, unpublished manuscript). In general, however, strong environments have different effects for girls and for boys. Strong environments generally widen the gender gap in physical science aspirations in favor of boys and shrink the female advantage in life science aspirations, but—as **Table 2** makes clear—this impact primarily falls on low performing students. Among high performing students, stronger math-science environments shrinks the overall STEM gender gap. These patterns are not universal, however. Countries display heterogeneity in the effects of the school performance environment on STEM aspirations and in particular the impact of the performance environment on student decision-making in response to their own level of math-science performance.

Some of this country variation can be attributed to country differences in the structure of tracking. Our analysis made clear that the strength of the own-performance signal on STEM aspirations is stronger in countries that do not use early tracking in their school systems than in countries with early tracking. In early tracking school systems, STEM aspirations are generally higher in the high performing schools (the “academic” track). In untracked school systems, STEM aspirations are generally higher at any given level of own performance in low-performing schools, and this gap in favor of low-performing schools grows as own performance increases. We see this as clear evidence of a social comparison effect in strong performance environments. Moreover, there is a clear gender difference in the workings of this social comparison effect. Boys respond more strongly to their own performance than do girls in environments that provide weak signals from tracking and in environments where peer performance is weak, which seems to induce strongly performing boys more than girls to draw the conclusion that they belong in STEM occupations. In environments with strong environmental performance, the gender gap in STEM aspirations shrinks. In other words, girls who perform well in environments filled with other strong performing students behave more similarly to boys in the formation of their STEM aspirations. Again, however, there is country-heterogeneity in the responses to own performance and environmental signals that our models cannot fully account for.

ACKNOWLEDGMENTS

This project was supported by Award Number R01EB010584 from the National Institute Of Biomedical Imaging And Bioengineering. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of Biomedical Imaging And Bioengineering or the National Institutes of Health.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2015.00171/abstract>

REFERENCES

- Alwin, D. F., and Otto, L. B. (1977). High school context effects on aspirations. *Sociol. Educ.* 50, 259–273. doi: 10.2307/2112499
- Buchmann, C., and Dalton, B. (2002). Interpersonal influences and educational aspirations in 12 countries: the importance of institutional context. *Sociol. Educ.* 75, 99–122. doi: 10.2307/3090287
- Buchmann, C., and Park, H. (2009). Stratification and the formation of expectations in highly differentiated educational systems. *Res. Soc. Stratif. Mob.* 27, 245–267. doi: 10.1016/j.rssm.2009.10.003
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: explaining the gap. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3157–3162. doi: 10.1073/pnas.1014871108
- Chmielewski, A. K., Dumont, H., and Trautwein, U. (2013). Tracking effects depend on tracking type an international comparison of students' mathematics self-concept. *Am. Educ. Res. J.* 50, 925–957. doi: 10.3102/0002831213489843
- Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. Sociol.* 106, 1691–1730. doi: 10.1086/321299
- Eccles, J. S. (2007). "Where are all the women? Gender differences in participation in physical science and engineering," in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, eds S. J. Ceci and W. M. Williams (Washington, DC: APA), 199–210.
- Eccles, J. S. (2011). Gendered educational and occupational choices: applying the eccles model of achievement-related choices. *Int. J. Behav. Devel.* 35, 195–201. doi: 10.1177/0165025411398185
- Eccles, J. S., Barber, B., and Jozefowicz, D. (1999). "Linking gender to educational, occupational, and recreational choices: applying the eccles et al. model of achievement-related choices," in *Sexism and Stereotypes in Modern Society: The Gender Science of Janet Taylor Spence*, eds W. B. Swann, J. H. Langlois, and L. A. Gilbert (Washington, DC: American Psychological Association), 153–192.
- Kjærnsli, M., and Lie, S. (2011). Students' preference for science careers: international comparisons based on PISA 2006. *Int. J. Sci. Educ.* 33, 121–144. doi: 10.1080/09500693.2010.518642
- Legewie, J., and DiPrete, T. A. (2012). School context and the gender gap in educational achievement. *Am. Sociol. Rev.* 77, 463–485. doi: 10.1177/0003122412440802
- Legewie, J., and DiPrete, T. A. (2014a). Pathways to science and engineering bachelor's degrees for men and women. *Sociol. Sci.* 1, 41–48. doi: 10.15195/v1.a4
- Legewie, J., and DiPrete, T. A. (2014b). The high school environment and the gender gap in science and engineering. *Sociol. Educ.* 87, 259–280. doi: 10.1177/0038040714547770
- Marsh, H. W., and Hau, K.-T. (2003). Big-fish–little-pond effect on academic self-concept: a cross-cultural (26-country) test of the negative effects of academically selective schools. *Am. Psychol.* 58, 364–376. doi: 10.1037/0003-066X.58.5.364
- Marsh, H. W., and Yeung, A. S. (1997). Coursework selection: relations to academic self-concept and achievement. *Am. Educ. Res. J.* 34, 691–720. doi: 10.3102/00028312034004691
- Morgan, S. L., Gelbgiser, D., and Weeden, K. A. (2013). Feeding the pipeline: gender, occupational plans, and college major selection. *Soc. Sci. Res.* 42, 989–1005. doi: 10.1016/j.ssresearch.2013.03.008
- Nagengast, B., and Marsh, H. W. (2012). Big fish in little ponds aspire more: mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *J. Educ. Psychol.* 104, 1033–1053. doi: 10.1037/a0027697
- Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., and Eccles, J. S. (2008). "Gendered high school course selection as a precursor of gendered careers: the mediating role of self-concept and intrinsic value," in *Gender and Occupational Outcomes: Longitudinal Assessments of Individual, Social, and Cultural Influences*, eds J. S. Eccles and H. M. G. Watt (Washington, DC: American Psychological Association), 115–143.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O., and Garrett, J. (2006). Gender and course selection in upper secondary education: effects of academic self-concept and intrinsic value. *Educ. Res. Eval.* 12, 323–345. doi: 10.1080/13803610600765687
- Park, H., Behrman, J., and Choi, J. (2012). "Do single-sex schools enhance students' stem (science, technology, engineering, and mathematics) outcomes?," in *SSRN Scholarly Paper ID 2153812* (Rochester, NY: Social Science Research Network).
- Riegle-Crumb, C., Moore, C., and Ramos-Wada, A. (2010). Who wants to have a career in science or math? exploring adolescents' future aspirations by gender and race/ethnicity. *Sci. Educ.* 95, 458–476. doi: 10.1002/sce.20431
- Shen, C., and Tam, H. (2008). The paradoxical relationship between student achievement and self-perception: a cross-national analysis based on three waves of TIMSS data. *Educ. Res. Eval.* 14, 87–100. doi: 10.1080/13803610801896653
- Sikora, J., and Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Sci. Educ.* 96, 234–264. doi: 10.1002/sce.20479
- Tai, R. H., Liu, C. Q., Maltese, A. V., and Fan, X. (2006). Planning early for careers in. *Science* 312, 1143–1144. doi: 10.1126/science.1128690
- Xie, Y., and Shauman, K. (2003). *Women in Science: Career Processes and Outcomes*, Vol. 26. Cambridge, MA: Harvard University Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 December 2014; accepted: 03 February 2015; published online: 25 February 2015.

Citation: Mann A, Legewie J and DiPrete TA (2015) The role of school performance in narrowing gender gaps in the formation of STEM aspirations: a cross-national study. *Front. Psychol.* 6:171. doi: 10.3389/fpsyg.2015.00171

This article was submitted to *Developmental Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2015 Mann, Legewie and DiPrete. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Perceived mathematical ability under challenge: a longitudinal perspective on sex segregation among STEM degree fields

Samantha Nix^{1*}, Lara Perez-Felkner¹ and Kirby Thomas²

¹ Department of Educational Leadership & Policy Studies, Florida State University, Tallahassee, FL, USA, ² Department of Sociology, Florida State University, Tallahassee, FL, USA

OPEN ACCESS

Edited by:

Stephen J. Ceci,
Cornell University, USA

Reviewed by:

Shulamit Kahn,
Boston University School of
Management, USA
Yi-Miau Tsai,
University of Michigan, USA

*Correspondence:

Samantha Nix,
Department of Educational Leadership
& Policy Studies, Florida State
University, 1114 W. Call St.
Tallahassee, FL, 32306-4452 USA
snix@fsu.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 24 December 2014

Accepted: 14 April 2015

Published: 09 June 2015

Citation:

Nix S, Perez-Felkner L and Thomas K
(2015) Perceived mathematical ability
under challenge: a longitudinal
perspective on sex segregation
among STEM degree fields.
Front. Psychol. 6:530.
doi: 10.3389/fpsyg.2015.00530

Students' perceptions of their mathematics ability vary by gender and seem to influence science, technology, engineering, and mathematics (STEM) degree choice. Related, students' perceptions during academic difficulty are increasingly studied in educational psychology, suggesting a link between such perceptions and task persistence. Despite interest in examining the gender disparities in STEM, these concepts have not been considered in tandem. In this manuscript, we investigate how *perceived ability under challenge*—in particular in mathematics domains—influences entry into the most sex-segregated and mathematics-intensive undergraduate degrees: physics, engineering, mathematics, and computer science (PEMC). Using nationally representative Education Longitudinal Study of 2002 (ELS) data, we estimate the influence of perceived ability under challenging conditions on advanced high school science course taking, selection of an intended STEM major, and specific major type 2 years after high school. Demonstrating the importance of specificity when discussing how gender influences STEM career pathways, the intersecting effects of gender and perceived ability under mathematics challenge were distinct for each scientific major category. Perceived ability under challenge in secondary school varied by gender, and was highly predictive of selecting PEMC and health sciences majors. Notably, women's 12th grade perceptions of their ability under mathematics challenge increased their probability of selecting PEMC majors over and above biology. In addition, gender moderated the effect of growth mindset on students' selection of health science majors. Perceptions of ability under challenge in general and verbal domains also influenced retention in and declaration of certain STEM majors. The implications of these results are discussed, with particular attention to access to advanced scientific coursework in high school and interventions aimed at enhancing young women's perceptions of their ability, in particular in response to the potentially inhibiting influence of stereotype threat on their pathways to scientific degrees.

Keywords: higher education, gender, STEM, pipeline, perceived ability, ability-related beliefs, college major

Introduction

Socially influenced beliefs about mathematics ability have been studied as possible explanations for the gender gap in science, technology, engineering, and mathematics (STEM) higher education. Nevertheless, there remains insufficient conceptual and empirical clarity about how beliefs influence gendered differences over time, specifically during upper secondary and post-secondary school—the primary years for attrition from pathways to science careers (Berryman, 1983; Morgan et al., 2013). Notably, theories have emerged suggesting that persistence when encountering potentially negative or challenging situations is influenced by students' perceived ability to complete specific tasks (*self-efficacy*) (Bandura, 1977; Pajares, 1996), beliefs about the malleability of their abilities (*mindset*) (Dweck, 2007, 2008), the alignment of their skills to the challenge presented by the material (*flow*) (Csikszentmihályi and Csikszentmihályi, 1988; Shernoff et al., 2003), and fear of confirming negative stereotypes related to their identities (*stereotype threat*) (Steele, 1997; Beilock, 2008). Related, students' self-assessments of their mathematics ability appear to vary by gender and influence STEM degree choice (Correll, 2001; Parker et al., 2012; Perez-Felkner et al., 2012). These studies indicate a growing interest in examining the puzzling persistence of gender disparities in STEM. These concepts have not been considered in tandem however, to investigate how domain-specific and domain-general *perceived ability under challenging conditions* influence the gender gap in the most sex-segregated undergraduate degrees: physics, engineering, mathematics, and computer science (PEMC).

This study takes on this gap in the literature. Using the nationally representative Education Longitudinal Study of 2002 (ELS) data, we estimate the influence of mindset and self-perceptions of mathematics ability in challenging contexts on each subsequent step in the STEM pipeline: completing advanced high school science courses, persistence in a STEM major, and specific STEM major selection. Importantly, we compare and control for variation in students' response to challenge in verbal and mathematics tasks, while also controlling for more objective measures of verbal and mathematics ability. Moreover, this study uses the most recent and complete U.S. panel data available to examine how perceptions of mathematics ability on difficult tasks changes over time¹, during the years that appear to be when most girls who exit the STEM pipeline conclude that they are more capable in other domains.

Previous Research

Empirical studies demonstrate a persistent gender gap in post-secondary degree attainment in certain mathematically-intensive STEM disciplines, both internationally (OECD, 2013) and domestically (NSF, 2013). Students' perceptions in response to

challenges and negative feedback may be particularly informative to enhancing our understanding of how to encourage women's persistence in these fields. Performance feedback is formally given to students through grades, which some have suggested can imply subject-field difficulty to students (Drew, 2011; Putman et al., 2014). Research on the influence of STEM grades is mixed, however. For instance, in his longitudinal study of a single, elite research institution, Ost (2010) found that female physical science majors were more likely than their male physical sciences or female life sciences counterparts to change majors in response to lower grades in their STEM courses. In contrast, Griffith's (2010) findings from an analysis of multi-institutional datasets suggests that the positive effects of higher STEM GPAs on STEM persistence is likely more important for men. Such findings have led some scholars to conclude that grades cannot adequately predict students' responses to challenge, and instead suggest investigations of social psychological factors that may play an even larger role in student choice-making processes (Rask, 2010; Stearns et al., 2013).

This study builds upon these efforts by looking specifically at the role of beliefs about difficult mathematics material, a vital competency area for success in postsecondary STEM fields. To frame our study, we discuss factors that have been shown to impact persistence in scientific fields. In particular, we focus on self-perceptions of ability in mathematics with difficult material, from tenth grade through university major selection.

How Demographic, Academic, and Schooling Contexts Influence Scientific Ambitions

Previous scholars have demonstrated links between family background, high school preparation, and environmental factors that have played a role in students' decisions to pursue degrees in scientific fields. Overall, female gender has been widely shown to differentially affect youths' preparedness for and persistence in certain STEM fields, both across and within racial-ethnic groups. In a qualitative study of prospective STEM majors at seven campuses in the early 1990s, Seymour (1999) found that women who entered college as potential STEM majors were less rigid in their choice of major than were men, with the exception of those most socioeconomically disadvantaged. Interestingly, Hanson (2008) finds that contemporary labor norms in the black community contribute to black girls' resilience in pursuing scientific careers. Moreover, black and Latina girls seem to take more advanced high school mathematics course sequences than their male peers (Riegle-Crumb, 2006). In a study of Latino STEM majors, Cole and Espinoza (2008) found that participants' gender had the third largest positive impact on GPA, providing evidence that Latinas outperform Latinos in STEM postsecondary classrooms. A national longitudinal study using ELS data similarly found a nuanced relationship between gender and race/ethnicity in who chooses STEM majors in college, with Latino males being the group least likely to pursue STEM and black males being the most likely among those who had completed pre-collegiate STEM coursework (Perez-Felkner et al., 2014).

In the later years of high school, students may elect to take advanced mathematics and science courses. Gendered patterns in completion of these courses have been found (Riegle-Crumb

¹The National Center for Educational Statistics (NCES) released ELS cohort data through 2012 regarding educational attainment. NCES released postsecondary transcript data on this cohort in mid-April, 2015. At the time of this writing, the most up-to-date accurate information regarding majors and degree fields was from the third wave of data, the 2nd follow-up in 2006 (NCES personal communication with authors, 2014).

et al., 2006), as girls may be less inclined to pursue areas that have not been associated with female success. Notably, some gaps have closed in recent years. For example, the National Center for Education Statistics reported gender parity in high school calculus completion in 2009 (Kena et al., 2014). While research on mathematics course taking is more extensive than that on science course taking (e.g., Davenport et al., 1998), the latter may be more important given the persistence of gendered patterns in science. For example, this report also found that girls were less likely to complete high school physics (33% of girls as compared to 39% of boys).

These high school course decisions can influence postsecondary STEM major selection and degree completion, in particular in PEMC fields. Across three nationally representative cohorts attending high school in the 1980s, 1990s, and 2000s, completion of physics and calculus before H.S. graduation each increased students' chances of enrolling in physical science or engineering majors in college (Riegle-Crumb et al., 2012). While completing advanced coursework increases girls' chances of going on to declare postsecondary majors in physical sciences, engineering, mathematics, and computer science, those girls who enrolled in more advanced mathematics and science coursework seemed to have more negative self-assessments of their ability and mindsets regarding mathematics ability (Perez-Felkner et al., 2012). Holding these negative beliefs may contribute to struggles women might encounter as some of the comparatively few women majoring in these fields in college. Nevertheless, this body of research suggests that advanced coursework positions students—including young women—to choose PEMC majors.

Decades of research have indicated that high school contexts contribute to variation in students' postsecondary outcomes (Coleman et al., 1982; Perez-Felkner, in press), which may influence their preparedness for and persistence in scientific majors. Geographic proximity to college may influence where and in what type of college students enroll (e.g., Rouse, 1995); Latinos are especially likely to attend college closer to home (López Turley, 2009). Proximity to college seems to influence enrollment among both advantaged and low-income students, but is less of an issue among students in the northeast, which has both a greater density of post-secondary offerings, selective colleges, and urban areas (Griffith and Rothstein, 2009). Some studies have suggested that students are less likely to select STEM majors if they attend selective postsecondary institutions (Griffith, 2010; Engberg and Wolniak, 2013), while others suggest that institutional selectivity has no effect on women and underrepresented students' pursuit of degrees in scientific fields major (Smyth and McArdle, 2004; Perez-Felkner and Schneider, 2012). At the secondary level, students attending urban schools have tended to have lower postsecondary outcomes (Niu and Tienda, 2013). Moreover, girls in rural schools were found to be more likely than those in suburban or urban schools to choose STEM majors in college, irrespective of their high school preparation for these fields (Perez-Felkner et al., 2014). Gendered differences in scientific degrees may then be partially explained by regional variation in both the density of 4-year colleges and the proportion of students living in cities vs. suburban and rural communities.

Beliefs about Mathematics

While some continue to argue that cognitive ability in mathematics varies by gender and drives the gap in the STEM labor force (Hedges and Nowell, 1995; Summers, 2005), empirical evidence largely refutes this claim (Hyde and Linn, 2006). Notably, a meta-analysis of U.S. state assessments of mathematics performance found that 2nd through 11th grade students did not significantly differ by gender; however limitations in these data did not allow for analyses of complex problem solving and advanced mathematics, areas in which extant research finds that gender differences may be more likely to emerge (Hyde et al., 2008). Research spanning two decades' of nationally representative cohorts reveals gender gaps in some STEM majors are not fully explained by achievement in mathematics (Riegle-Crumb et al., 2012). Many have theorized that individuals' understanding of themselves and mathematics can influence students' major choices. For instance, Perez-Felkner et al. (2012) examined ELS data to show that subjective orientations to mathematics (operationalized as perceived mathematics ability as well as engagement in, valuing, and mindset toward mathematics ability) was positively and significantly correlated with selection of PEMC and Biological sciences majors. Similarly, but with a focus on self-concept, Parker et al. (2012) analyzed large-scale datasets from both Germany and England. Their findings revealed that mathematics self-concept predicted students' entry into physical sciences, engineering, and mathematics. In addition, self-concept was found to be a more powerful predictor of major choice than standardized tests of ability, consistent with studies showing that ability does not explain the gender gap (Perez-Felkner and Schneider, 2012; Riegle-Crumb et al., 2012).

Correll (2001) used National Education Longitudinal Study (NELS) 1988 data to show that girls underrate their abilities in mathematics, even after controlling for performance feedback and objective measures of their abilities. Also using NELS, but extending her research into postsecondary outcomes, Ma (2011) found that perceptions of mathematics ability predicted entry into a STEM field, and that those perceptions were least predictive of entry into life science majors. Consistent with this research, Sax (1994) analyzed Cooperative Institutional Research Program (CIRP) 1985/1989 data and found that mathematics self-concept at the beginning and end of college was significantly lower for women in her sample than men. Importantly, her research also showed that for women in particular, mathematics self-rating at the end of college was significantly and strongly predicted by confidence in mathematics ability before entering postsecondary environments.

Research also reveals that perceptions do not exist in a vacuum. For instance, Correll (2001) found that students compare their progress in mathematics and verbal domains, with higher scores and perceptions of ability in English predicting lower perceptions of mathematics ability and selection out of advanced mathematics courses. Wang et al. (2013) found evidence that ability in both mathematics and verbal domains might lead women to believe that they have a wider range of career choices. In particular, those with high ability in both mathematics and verbal domains were predicted to select out of STEM fields

compared to women with high mathematics ability and moderate verbal ability.

Given the findings summarized above, this study considers beliefs about abilities in general, verbal, and mathematics domains. Further, we focus particularly on students' perceived ability to overcome challenging or difficult material. We hypothesize that variations in those perceptions predict selection of advanced science courses in high school, persistence in STEM fields, and selection of mathematics-intensive majors.

Conceptual Framework

Our research questions and design respond primarily to prominent social psychological theories, which also inform the interpretation of our results.

Self-Efficacy

Bandura's (1977) self-efficacy is perhaps the most widely applied educational motivation theory, especially in investigations of the gender and race/ethnicity variation in STEM fields (Pajares, 1996; Rittmayer and Beier, 2009). Describing students' perceptions of their ability to complete specific tasks in particular domains (such as long division in mathematics), self-efficacy links beliefs, behaviors, and environments to explain students' choice making processes (Pajares, 1996; Zimmerman, 2000). The theory's value arises in part from its wide application—it can be applied across disciplines, given the application is task-specific. In focusing on one's beliefs in their ability to *do* a specific task, self-efficacy measures may miss students' immediate and overall assessment of a domain: whether or not it presents an overwhelming challenge to the student to begin with, before they start contemplating their ability to complete specific tasks within that field of study. Therefore, our analysis focuses on domain-specific (rather than task-specific) perceptions of ability under challenge.

Flow

In contrast to self-efficacy, Csikszentmihalyi's flow theory integrates people's perceptions of challenge and their corresponding perceptions of ability. Flow theory, at its heart, is about "optimal" experience (Nakamura and Csikszentmihalyi, 2002, p. 89)—the moment when people become so involved in their tasks, that they lose their sense of self-consciousness and the passage of time. According to the theory, people arrive in this state of being when a task just meets the threshold of their abilities, and thus are perceived as challenging, but not overwhelming (Csikszentmihályi and Schneider, 2000; Nakamura and Csikszentmihályi, 2002). Additionally, people gain such satisfaction from moments when they are in flow, that they seek out tasks that will continue to provide them with such experiences (Csikszentmihályi and Csikszentmihályi, 1988; Csikszentmihályi and Csikszentmihályi, 1991). We propose that students who believe that they can overcome challenge in mathematics domains will continue to seek those experiences out, via selecting mathematics-related majors while in college.

Mindset Theory

Dweck's (2000, 2006) mindset theory proposes that students do not have a universal response to challenge. Instead, their response

to challenge is mediated by their mindset, or their belief that abilities can be developed or are innate. Those who believe that intelligence is innate—people with a fixed mindset—tend to be much less likely to select challenging tasks, because they do not want to disconfirm their intelligence in front of others. In contrast, those who believe that intelligence is malleable or can be developed—growth mindset individuals—tend to take on challenging material because they do not believe the task at hand implies anything specific about their overall intelligence. Thus, fixed mindset individuals are thought to have helpless responses to challenging material, while growth mindset individuals are thought to have mastery responses to challenging material (Dweck, 2000, 2006).

Importantly, girls have been shown to be more likely to hold a fixed mindset (Dweck, 2007), suggesting that they may implement helpless behaviors when confronting a difficult task. Further, much of the research on this topic argues that adjustments to women's and underrepresented minorities' mindsets could help with gaps in STEM participation (Dweck, 2008; Good et al., 2012; Mangels et al., 2012). If girls are more inclined to view their abilities as fixed rather than malleable, they may also be more likely to believe that they are not capable when they encounter setbacks on challenging mathematics tasks. How gender moderates perceived ability becomes particularly important, given the prevailing stereotypes that girls encounter regarding their mathematics ability.

Stereotype Threat

According to Steele (1997) stereotype threat occurs when an individual internalizes the stereotypes of a group with which they identify, such as women's perceived weakness in mathematics. Biellock and colleagues have proposed a link between stereotype threat and task success via working memory (Beilock, 2008; Rydell et al., 2009; DeCaro et al., 2010). Importantly, these studies use experimental research design to establish that women's working memory is inhibited when they are reminded about the gender stereotype that women are less successful at mathematics, and propose interventions to help mitigate that effect (Good et al., 2003). Therefore, we recognize that females' perceptions of ability to overcome challenge might be particularly important as they move into increasingly more gender-segregated academic environments while advancing toward STEM degrees, in which stereotypic beliefs may be more salient.

Research Questions and Hypotheses

We build upon the previous research presented above to examine the complex interplay between gender, perceptions, and participation in STEM, particularly under difficult or challenging conditions in mathematics and other domains. Specifically, four research questions guided our research:

1. To what degree do domain-specific and domain-general perceptions of ability under challenge differ by gender?
2. What is the relationship between perceived ability under challenge in mathematics and advanced high school science course enrollment?
3. To what extent does perceived ability under challenge in mathematics predict staying in a STEM field as intended before

entering postsecondary education? How is this relationship moderated by gender?

4. What is the relationship between perceived ability under challenge in mathematics and selection of mathematics-intensive science majors (PEMC), and how is that relationship moderated by gender?

As emphasized in the research questions above, we hypothesize that gender moderates the relationships between perceived ability under mathematics challenge and outcomes for subsequent steps in the STEM pipeline. Therefore, our research questions build upon one another, leading to our primary focus: an examination of the relationships between perceived ability under challenge, gender, and selection of mathematics-intensive majors (see **Figure 1**).

Methods

Data Source and Participants

We used nationally representative Education Longitudinal Study (ELS) panel data to address our research questions. Collected by the National Center for Education Statistics, probability sampling was implemented for the base year data collection effort in 2002, yielding a sample of 17,591 eligible 10th graders from 752 high schools across the United States. Parents, administrators, staff, and teachers were also surveyed. Follow-ups were then conducted in 2004 (during most students' 12th grade year), 2006, and 2012 (Ingels et al., 2007). For clarity, we discuss the data primarily in reference to participants' stage in education. For instance, "10th grade" refers to 2002 or base year data, "12th grade" refers to 2004 or first follow-up data, and "2 years after high school" refers to 2006 or second follow-up data. This study uses the high school (10th and 12th grades) and 2 years after high school student surveys, including some control variables (such as family income, education, and high school environment measures) gleaned from the accompanying parent and administrator surveys (see the Appendix in Supplementary Material for more details). Survey administrators reported an 88% weighted response rate for students participating in these first three waves: 10th grade (2002), 12th grade (2004), and 2 years after high school (2006) (Ingels et al., 2007).

Our analytic sample represents the college-going population of U.S. students who were tenth graders in the spring of 2002 and enrolled in college between 2004 and 2006. We include only students who attended either 2- and 4-year institutions by 2 years

after high school, as our second and third research questions are related to college major choice. Any students who remained undecided or undeclared were coded as such but retained in our analyses. Therefore, of the 16,197 observations in the ELS dataset, we found that 10,534 had enrolled in a postsecondary institution by 2 years after high school. Because of our interest in race as well as gender, we additionally excluded respondents from groups with overly low representation in the sample². We then used listwise deletion for any remaining missing observations on the independent and dependent variables, yielding a final analytic sample of 4450 cases. Lastly, we used response adjusted, calibrated bootstrap replicate weights (ELS variables f2byp1-f2byp200), and panel survey weighting (with f2bywt) to adjust for stratification in the sample design. Sample descriptive statistics are discussed throughout the measures section.

Measures

Outcome Variables

Science pipeline

First, we examined the most advanced science course students took in high school. We collapsed the original categories from eight to three to enhance the interpretability of our analyses, as those on the lowest end of the science pipeline tended not to attend nor complete college. Thus, the science pipeline variable focuses on the upper end of the scale and represents students' completion of three levels of science coursework: (1) chemistry I or physics I or less, (2) both chemistry I and physics I, and (3) chemistry II and physics II. Biology and other sciences were included in the science pipeline variable, but the ranking privileges chemistry and physics as indicators of having completed the "science pipeline" in high school. We report on the relationship between gender and completion of these science pipeline courses in **Table 1**. Fewer women participated in the highest and middle level of science coursework compared to men (**Table 1**). Correspondingly, there is a higher percentage of women (53.1%) who only completed the lowest level of science coursework (Chemistry I or Physics I and below) compared to men (45.4%).

Major retention

Next, we were interested in what encouraged retention in STEM fields. To understand this, we compared participants' intended

²While we hoped to include Native Americans/Alaskan Natives in our study, they would have comprised only 1.49% of our sample; meaningful results would not have been attainable. We therefore excluded this group from our analysis.

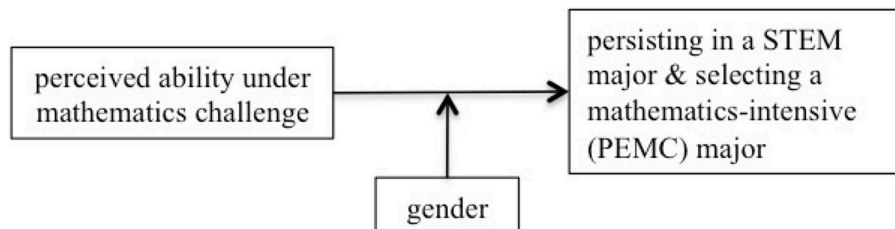


FIGURE 1 | Conceptual model of how gender moderates perceived ability and major choice.

TABLE 1 | Sample descriptive statistics on dependent variables.

	Men		Women		Min	Max
	%	SE	%	SE		
SCIENCE PIPELINE						
Chemistry I or Physics I and below	45.4%	1.6	53.1%	1.6	0.0	100.0
Chemistry I and Physics I	26.0%	1.6	22.6%	1.3	0.0	100.0
Chemistry II and Physics II	28.6%	1.6	24.3%	1.3	0.0	100.0
MAJOR RETENTION						
Abstainers	68.1%	1.4	86.9%	0.8	0.0	100.0
Stayers	16.6%	1.2	3.7%	0.4	0.0	100.0
Leavers	10.1%	0.9	4.0%	0.5	0.0	100.0
Newcomers	5.3%	0.7	5.4%	0.5	0.0	100.0
MAJOR TYPE						
Undeclared/undecided	24.0%	1.4	20.6%	1.0	0.0	100.0
Non-STEM	39.7%	1.5	42.5%	1.3	0.0	100.0
PEMC	17.1%	1.2	4.0%	0.5	0.0	100.0
Biology	4.7%	0.6	5.0%	0.5	0.0	100.0
Health	4.0%	0.6	15.9%	1.0	0.0	100.0
Social/Behavioral and other sciences	10.4%	0.9	11.9%	0.7	0.0	100.0

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002/2006 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Restricted-use NCES data requires rounding these descriptive results to the nearest tenth.

major to their declared major 2 years after high school. The intended major variable was retrospective, as students were asked 2 years after high school which field they intended on entering before starting their postsecondary educations. Further, due to the original coding of the intended major data, intended PEMC and biology majors could not be disaggregated (see the Appendix in Supplementary Material for more detail). Therefore, the major retention variable includes four categories: (1) abstainers (never intended or majored in PEMC and/or biology), (2) stayers (intended and majored in PEMC and/or biology), (3) leavers (intended but did not major in PEMC and/or biology), and (4) newcomers (did not intend but majored in PEMC and/or biology).

Overall, a larger proportion of men (32.0%) in our sample participated in a PEMC and/or biology major in some way, compared to women (13.1%, **Table 1**). A full 86.9% of women in our sample abstained from PEMC and/or biology, neither intending nor enrolling in those fields by 2 years after high school. This high lack of engagement drives the lower percentages in the other major retention categories for women in our sample. For instance, only 3.7% of women persisted in a PEMC and/or biology field as intended compared to 16.6% of men. In addition, 4.0% of women left a PEMC and/or biology field, compared to 10.1% of men. Finally, a comparable proportion of men and women (5.3 and 5.4%, respectively) were considered newcomers, entering a PEMC and/or biology field by 2 years after high school, even though it was not their intended major.

Major type

The last outcome variable provides information on the specific type of major students selected 2 years after leaving high school. Because of the importance of disaggregating STEM majors by fields of study (Perez-Felkner et al., 2012; Ceci et al., 2014), we looked specifically at students majoring in the physical sciences, engineering, mathematics, and computer sciences (PEMC) against other STEM majors; specifically, we compare PEMC to biology, health, social/behavioral and other sciences, and non-STEM majors. We additionally compare with undecided/undeclared students, to achieve a more representative set of analyses from high school through college. Full details (including the list of majors included in each category) are provided in the Appendix in Supplementary Material.

Looking specifically at this outcome for our sample on **Table 1**, we see that 2 years after high school, 24.0% of men and 20.6% of women had an undeclared or undecided major. A larger proportion of women (42.5%) had a non-STEM major, compared to men (39.7%). Consistent with the previous literature, a far smaller percentage of women majored in PEMC fields (4.0%) compared to men (17.1%), though these results are roughly mirrored when looking at health fields (15.9% of women vs. 4.0% of men). Men and women participated at comparable levels in biology fields (4.7% of men and 5.0% of women), with a slightly higher proportion of women (11.9%) declaring a social/behavioral and other science major compared to men (10.4%).

Perceived Ability under Challenge in Domain-General and Verbal and Mathematics Domains

As noted above, this study is primarily concerned with students' perceived ability under challenge. We operationalized this concept by selecting ELS items that represented students' perceptions of their ability to use mastery-oriented behavior and comfort with complex or difficult material. A brief discussion of each measure of perception of ability to overcome challenge is discussed below, with full details in the Appendix in Supplementary Material. All perceived ability under challenge variables were mean-centered for interpretability. We report mean scores for men and women in the Results Section.

General index

Five items from the 10th grade survey were used to assess students' perceived ability under challenge in general, as opposed to within a particular subject domain. Original scores on each ranged from 1 to 4, with higher values representing higher agreement with each statement. From these five statements, we developed a mean item index to represent domain-general perceived ability under challenge ($\alpha = 0.865$).

Verbal index

Tenth graders were also asked to report their agreement with three statements related to their comfort with difficult verbal tasks and use of mastery behavior in that field. Similar to the items on the general index, each of these responses were originally coded 1–4; a score of 4 indicates the highest agreement with each statement. These variables were averaged into a mean item index ($\alpha = 0.881$).

Mathematics index (10th and 12th grades)

Three questions were repeated on the 10th and 12th grade surveys related to students' perceptions of ability to overcome challenge in mathematics domains. As with the questions on the other indices, responses to these questions were originally coded 1–4; a score of 4 indicates agreement with each statement and higher perception of ability to overcome challenge in mathematics. Scores on each set of questions were averaged into two separate mean item indices, one each for the 10th grade ($\alpha = 0.892$) and the other for the 12th grade ($\alpha = 0.871$).

Growth mindset

Finally, one question from the base year survey asked students about their level of agreement with a statement related to Dweck's (2000, 2006) concept of growth mindset (whether or not people could learn to be good at mathematics). Because this is a question specific to one theory and not necessarily related to the other mathematics measures identified in the questionnaires, we let it stand alone. As with the other measures of perceived ability under challenge, this variable was coded such that 1 indicated less agreement and 4 indicated more agreement.

Control Variables

Demographic characteristics

Demographic variables included dichotomous variables for gender, race/ethnicity (white, Asian/Pacific Islander, black, Latino, multi-race/ethnic), parents' education (high school degree or less, less than a 4-year degree, 4-year degree, more than a 4-year degree), and family income by quartiles (\$0–\$35,000 per year, \$35,001–\$50,000 per year, \$50,001–\$100,000 per year, \$100,001 or more per year). Detailed information on the coding of these variables is available in the Appendix in Supplementary Material.

Table 2 presents descriptive statistics for the sample and indicates that there are more women in the sample (57.4%) compared to men (42.6%)³. Additionally, white students constitute the majority of the sample (73.7%), while the rest of the sample consists of 5.0% Asian American/Pacific Islander students, 9.1% each for black and Latino students, and 3.2% multi-race/ethnicity students. About a quarter of the participants in our sample had parents that earned more than a bachelor's degree, and almost 60.0% of the sample had parents that attended some college or earned a bachelor's degree. 16.1% of the sample had parents with a high school diploma or less. Turning to family income, the largest percentage of our sample (40.7%) came from families that earned between \$50,001 and \$100,000 per year. 18.0% of the sample had families that earned \$35,001–\$50,000 per year. Finally, about 20.0% of the sample had families that earned either the lowest level of family income (up to \$35,000 per year) or the highest level of family income (more than \$100,000 per year).

³For comparison, according to the U.S. Census Bureau Current Population Surveys, in 2004, the percentage of women enrolled in postsecondary institutions was at 41.2%, exceeding men at 34.7%. As noted in the literature review, women have been exceeding men in postsecondary enrollments in both the U.S. and other industrialized nations, since the 1990s or in some cases earlier. See: http://nces.ed.gov/programs/digest/d13/tables/dt13_302.60.asp

TABLE 2 | Sample descriptive statistics.

	Mean	SE	Min	Max
DEMOGRAPHIC CHARACTERISTICS				
Gender				
Men	42.6%	1.1	0.0	100.0
Women	57.4%	1.1	0.0	100.0
Race/Ethnicity				
White	73.7%	1.1	0.0	100.0
Asian/Pacific Islander	5.0%	0.4	0.0	100.0
Black	9.1%	0.7	0.0	100.0
Latino	9.1%	0.7	0.0	100.0
Multi-race/Ethnicity	3.2%	0.4	0.0	100.0
Parent Education				
High school or less	16.1%	0.9	0.0	100.0
Some college	29.9%	1.0	0.0	100.0
Bachelor's degree	29.0%	0.9	0.0	100.0
More than a bachelor's degree	25.0%	1.1	0.0	100.0
Family Income				
First quartile (\$0–\$35,000 per year)	20.2%	0.9	0.0	100.0
Second quartile (\$35,001–\$50,000 per year)	18.0%	0.9	0.0	100.0
Third quartile (\$50,001–\$100,000 per year)	40.7%	1.0	0.0	100.0
Fourth quartile (more than \$100,000 per year)	21.1%	1.0	0.0	100.0
STUDENT ABILITY				
Ability With Complex Material				
Mathematics (10th grade)	2.1%	0.2	0.0	100.0
Reading (10th grade)	16.0%	0.6	0.0	100.0
Grade point average (10th grade)	3.0	0.0	0.4	4.0
INSTITUTIONAL EFFECTS				
HS Region				
Northeast	20.4%	1.3	0.0	100.0
Midwest	28.4%	1.3	0.0	100.0
South	30.9%	1.3	0.0	100.0
West	20.3%	1.2	0.0	100.0
HS Urbanicity				
Urban	27.3%	1.3	0.0	100.0
Suburban	54.4%	1.4	0.0	100.0
Rural	18.4%	1.1	0.0	100.0
Institutional Selectivity				
2-year or less institution	27.3%	1.2	0.0	100.0
4-year institution, inclusive	13.4%	0.8	0.0	100.0
4-year institution, moderately selective	30.9%	1.0	0.0	100.0
4-year institution, highly selective	28.4%	1.2	0.0	100.0

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002/2006 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Restricted-use NCES data requires rounding these descriptive results to the nearest tenth. Ability with complex material (mathematics) and ability with complex material (reading) variables are reported in mean percentage form to more meaningfully explain their characteristics descriptively. Our multivariate analyses use the original form of the variables (on a 0.0–1.0 point scale, not percentages).

Student ability

Students' ability was measured through scores on the most complex standardized mathematics and reading questions and grade point average, both in the 10th grade. Scores on the most complex

standardized mathematics and reading questions were measured using a continuous variable ranging from 0.0 to 1.0, representing the probability that students would respond correctly to three of the four questions in each category. We used this original form of the variable in our multivariate analyses for the sake of comparability to other studies on this data, but report a percentage form in **Table 2** to meaningfully interpret the descriptive statistics. Tenth grade GPA was also a continuous variable ranging from 0.0 to 4.0.

Because all of our student ability measures are continuous in nature, these scores are reported in means. Using a 0.0–100.0 point scale to increase the interpretability of our descriptive statistics, we can see that the mean probability that our sample could complete the most difficult standardized mathematics questions was 2.1%. This indicates that much of our sample had almost no probability of answering three of the four most complex standardized mathematics questions⁴. On the other hand, our sample fared better on the mean probability score of completing three of the four most difficult standardized reading questions, at an average of 16.0% on a 0.0–100.0 point scale. Finally, our sample had a mean 10th grade GPA of 3.0/4.0.

High school context.

To control for students' high school contexts, we included measures of their region and urbanicity. Region is based on high school location and corresponds to Census categories: Northeast, Midwest, South, and West. Urbanicity corresponds to NCES classifications: urban, suburban, and rural. Participants in our sample attending high schools across the U.S., with 30.9% concentrated in the South, 28.4% in the Midwest, and just over 20.0% each in the West and Northeast. 54.4% of our sample attended high schools in suburban areas, while 27.3% and 18.4% attended schools in urban and rural areas, respectively.

Institutional selectivity

We also controlled for the institutional selectivity of students' first attended postsecondary institutions as of 2 years after high school. Selectivity is split into four dichotomous categories:

⁴Restricted-use NCES data requires rounding these descriptive results to the nearest tenth.

2-year college or less, 4-year institution (inclusive or not classified), 4-year institution (moderately selective), and 4-year institution (highly selective). 27.3% of the sample started at a 2-year institution, while 72.7% started at a 4-year institution. Of that 72.7% who started at a 4-year institution, 13.4% first attended an inclusive, 30.9% a moderately selective, and 28.4% a highly selective college or university.

Analytic Plan

Our first research question is primarily concerned with understanding if there are gendered differences in perceived ability under challenge. Therefore, we calculated the sample means for men's and women's scores on each of the perceptions of ability to overcome challenge variables and used Adjusted Wald Tests to provide us with information about significant differences between the two groups. To address the second research question related to the highest science course taken in high school, we used ordered logistic regressions. Finally, we used multiple logistic regressions to examine the third and fourth research questions related to STEM retention and specific major choice. More details regarding our analyses are presented with our results.

Results

Gender and Perceived Ability under Challenge in Mathematics

In light of previously cited research indicating differences in boys' and girls' assessments of their abilities, we used sample mean Wald tests to determine whether there were significant gender differences on our measures of perceived ability under challenge. **Table 3** reveals that in fact young men and women rate themselves as similarly confident in their abilities under challenge in general (Wald = 1.5; $p = 0.222$) as well as in the verbal domain (Wald = 0.4; $p = 0.555$). In contrast, mean differences *between* women and men were highly significant for each measure of perceived ability under challenge in mathematics. Young men were between 0.1 and 0.4 points above the mean on each measure, while young women fell either at or just below the mean in their perceived ability under mathematics challenge in 10th and 12th grades. Notably, the gap between women's and men's

TABLE 3 | Perceived ability under challenge (mean centered) by gender, weighted.

	Mean		SE		Range	
	Men	Women	Men	Women	Min	Max
NON-MATHEMATICS MEASURES						
General index (10th grade)	0.1	0.2	0.0	0.0	−1.8	1.2
Verbal index (10th grade)	0.2	0.2	0.0	0.0	−1.7	1.3
MATHEMATICS MEASURES						
Growth mindset (10th grade)	0.1	−0.1***	0.0	0.0	−2.0	1.0
Mathematics index (10th grade)	0.4	0.0***	0.0	0.0	−1.5	1.5
Mathematics index (12th grade)	0.2	0.0***	0.0	0.0	−1.5	1.5

$n = 4450$ respondents from the National Center for Education Statistics' Education Longitudinal Study 2002/2006 restricted data. Wald tests were used to determine the significance of difference between the means for men and women. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Restricted-use NCES data requires rounding these descriptive results to the nearest tenth. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

ratings of their perceived ability under challenge is largest on the 10th grade mathematics index (diff = 0.4; Wald = 102.9; $p = 0.000$) and tapers slightly 2 years later (diff = 0.2; Wald = 58.7; $p = 0.000$). This change results primarily from a loss of confidence for young men, who see a 0.2 mean centered point decrease between 10th and 12th grades. Among young women, perceptions of their ability on difficult mathematics do not appear to fluctuate over time. Young women are less inclined to report a growth mindset than are young men (diff = 0.2; Wald = 30.5; $p = 0.000$). Together, these findings suggest that young men are better positioned psychologically to be resilient in the face of mathematics-related setbacks, as compared to their female peers.

Impacts on Science Course Taking

Next, we turned to the question of how advanced science course taking in high school might be distinctly influenced by perceptions of ability under challenge in general, verbal, and mathematics domains. Given that the lower science pipeline courses are pre-requisites of higher science pipeline courses, and thus are ordinal in nature, we used ordered logistic regressions. The first model included our outcome variable and student demographic characteristics. The second model added student ability, high school context, and institutional selectivity to the variables in the first model. Lastly, the third and final model included all of our predictor variables, including the perceived ability under challenge variables. For simplicity, we present the full model only in **Table 4**. Proportional odds ratios (OR) represent the ratio of odds for completing the highest level of science coursework in high school as compared to the odds of the other combined outcomes (less rigorous courses). When interpreting OR, values under 1 represent negative relationships, values over 1 represent positive relationships, and values approaching 1 represent relationships with less meaningful significance.

Taking all other factors into account, women have about 24.0% lower odds of completing both Chemistry II and Physics II (OR = 0.76; $p = 0.001$) as compared to men, all else being equal. Race/Ethnicity matters as well, as Asian/Pacific Islander students are considerably more likely to complete these more advanced science courses than are white students (OR = 2.44; $p = 0.000$). Conversely, black students are less likely to complete these courses than are their white peers, although this effect is less significant (OR = 0.70; $p = 0.047$).

Objective measures of ability were also meaningfully significant. Students' academic ability with complex material—as measured by test scores—is highly related to completing the most rigorous science courses. Interestingly, this pattern holds for both mathematics and verbal domains. Recall that these scores refer to students' performance on the most challenging sections of the NCES-administered ability tests. A one percentage point increase in one's complex mathematics ability score—an area where most sample respondents struggled, as noted in **Table 2**—corresponds to having 14 times higher odds of completing both physics II and chemistry II (OR = 14.32; $p = 0.001$). Moreover, the same magnitude of increase in complex reading ability (OR = 2.03; $p = 0.000$) notably enhances the likelihood of completing these courses, all else being equal, as does earning higher grades in school (OR = 2.18; $p = 0.000$).

TABLE 4 | Likelihood of advanced science course completion by the end of 12th grade.

	Chemistry II and Physics II	
	OR	SE
DEMOGRAPHIC CHARACTERISTICS		
Female gender (Reference = male)	0.764**	0.062
Race/Ethnicity (Reference = white)		
Asian/Pacific Islander	2.438***	0.338
Black	0.696*	0.126
Latino	0.979	0.135
Multi-race/Ethnicity	1.231	0.255
STUDENT ABILITY		
Ability with Complex Material		
Mathematics (10th grade)	14.324**	11.692
Reading (10th grade)	2.028***	0.342
Grade point average (10th grade)	2.184***	0.156
INSTITUTIONAL EFFECTS		
HS Region (Reference = Northeast)		
Midwest	1.026	0.174
South	0.940	0.147
West	0.694*	0.114
HS Urbanicity (Reference = urban)		
Suburban	0.843	0.102
Rural	0.595**	0.111
PERCEIVED ABILITY UNDER CHALLENGE		
Non-mathematics measures		
General index (10th grade)	1.057	0.064
Verbal index (10th grade)	0.911	0.047
Mathematics Measures		
Growth mindset (10th grade)	1.004	0.061
Mathematics index (10th grade)	1.296***	0.063
Cut1	9.646***	2.599
Cut2	34.210***	9.541
f-statistic	7.050***	
Observations	4450	

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Reference category includes all levels of science course taking less than Physics II and either advanced biology, chemistry, or physics, to conform to the proportional odds assumption. Parent education and family income were included in the model, but are withheld from this table for space. Full tables available from the authors by request. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Although not a focal dimension of this study, the effects of students' high school institutional contexts are worth noting. Access to advanced chemistry and physics coursework is not uniformly available, as noted earlier in this paper. As such, it may not be surprising that students attending high school in Western states (OR = 0.69; $p = 0.028$) are less likely to complete these courses than are students in the Northeast. Correspondingly, students enrolled in rural high schools are less likely to complete these courses than are students in urban high schools (OR = 0.60; $p = 0.006$).

Turning to our primary independent variables of interest, we were surprised that only one perceived ability under challenge measure significantly predicted completion of the most advanced science courses in high school. The 10th grade mathematics index predicts about a 30.0% increase in the odds of taking the highest science courses in high school, holding all other variables constant ($OR = 1.30$; $p = 0.000$). In contrast, the OR on the growth mindset variable did not reach significance, implying that students' mindset does not affect high school science pipeline completion. Neither verbal nor domain-general perceived ability under challenge significantly predicted completion of these courses.

We examined a model (not shown) including product-term interactions between each of the perceived ability under challenge variables and gender. None of the resulting interaction terms were significant, thus we chose not to display results in this paper due to space constraints. However, the lack of significance on

these interaction terms is notable, and combined with the significant effect of the gender variable, indicates that women are less likely to take these courses, but are not affected differentially by perceptions of ability.

PEMC and/or Biology Retention and Perceived Ability under Challenge

Table 5 reports on the results of a multiple logistic regression analysis estimating the likelihood of retention in students' intended major. The reference group is comprised of those who neither intended nor declared PEMC and/or biology majors, compared to those who stayed, left, or were newcomers to PEMC and/or biology fields. As with the science pipeline analysis, we estimated four separate models to understand the movement between intended and declared major. Our first model included demographic characteristics only; the second included student ability, high school context, and institutional selectivity; and the

TABLE 5 | Retention in self-reported intended major, 2 years after high school.

	Never entered PEMC and/or biological sciences majors (reference)					
	PEMC and/or biological sciences stayers		PEMC and/or biological sciences leavers		PEMC and/or biological sciences newcomers	
	RRR	SE	RRR	SE	RRR	SE
DEMOGRAPHIC CHARACTERISTICS						
Female gender (Reference = Male)	0.194***	0.034	0.354***	0.068	0.749	0.142
Race/Ethnicity (Reference = White)						
Asian/Pacific Islander	0.999	0.263	1.059	0.286	1.425	0.515
Black	3.204***	0.940	1.102	0.383	1.777	0.596
Latino	0.920	0.300	1.067	0.319	1.520	0.655
Multi-race/Ethnicity	0.855	0.351	1.423	0.652	1.550	0.672
STUDENT ABILITY						
Ability with Complex Material						
Mathematics (10th grade)	2.358	1.616	2.311	1.826	2.170	1.828
Reading (10th grade)	0.961	0.288	0.830	0.265	0.770	0.282
Grade point average (10th grade)	1.583**	0.218	0.913	0.129	1.427*	0.241
Science Pipeline Completion (Reference = Chemistry I or Physics I or Less)						
Chemistry I and Physics I	1.779*	0.395	1.374	0.316	1.215	0.302
Chemistry II and Physics II	2.722***	0.540	2.578***	0.558	2.512***	0.586
PERCEIVED ABILITY UNDER CHALLENGE						
Non-Mathematics Measures						
General index (10th grade)	0.979	0.133	1.051	0.144	1.480*	0.241
Verbal index (10th grade)	0.582***	0.053	0.701**	0.076	0.786	0.105
Mathematics Measures						
Growth mindset (10th grade)	1.354*	0.177	1.244	0.165	1.062	0.176
Mathematics index (10th grade)	1.617***	0.175	1.276*	0.154	1.084	0.151
Mathematics index (12th grade)	1.561***	0.178	1.239	0.140	1.326*	0.148
Constant	0.020***	0.010	0.070***	0.043	0.012***	0.008
<i>f</i> -statistic	7.050***		7.050***		7.050***	
Observations	4450		4450		4450	

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Family income, parent education, ability with complex material, high school region and urbanicity, and institutional selectivity were included in the model, but are withheld from this table for space. * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$.

third included the perceived ability under challenge indices. For simplicity, we only report the final model using relative risk ratios (RRR). RRR are interpreted as the ratio of the probability that one outcome category will occur compared to that of the reference category (Borooah, 2002; Vogt, 2005). The basic interpretation of ORs and RRRs are congruent: values under 1 represent negative relationships, values over 1 represent positive relationships, and values approaching 1 represent relationships with less meaningful significance.

Comparing PEMC and/or biology stayers, leavers, and newcomers with having never expressing interest in those fields (abstainers), there are two significant findings for women. All else being equal, women have an 80.6% lower risk than men of staying in PEMC and/or biology fields as intended before starting college ($RRR = 0.19$; $p = 0.000$) vs. not entering these fields at all. While the effect is smaller, gender also predicts attrition from PEMC and/or biology fields. Women have a 64.6% lower risk than men of leaving these fields vs. not entering those fields at all ($RRR = 0.35$; $p = 0.000$), holding all other factors constant. There was also a significant relationship for one race/ethnicity category. All else being equal, our results show that black participants' risk of staying in a PEMC and/or biology field was 3.2 times higher as compared to white participants ($RRR = 3.20$; $p = 0.000$), among those who intended to major in that field before enrolling in college.

Lastly, there were notable significant results with respect to our student ability measures and science pipeline completion. While complex mathematics and verbal scores were highly predictive of completing advanced high school science coursework, they are no longer significant with respect to major retention, a more advanced step along the scientific pipeline. A 0.01 point increase in 10th grade GPA increased the risk of staying and entering a PEMC and/or biology field by 58.3% ($RRR = 1.58$; $p = 0.001$) and 42.7% ($RRR = 1.43$; $p = 0.036$), respectively, as compared to never entering these fields. Science course completion generated the second highest effect sizes in this model. Completing chemistry II and physics II in high school increased the likelihood of staying, leaving, and entering PEMC and/or biology fields by over 2.5 times each (all $p < 0.001$), as compared to never intending nor entering those fields. While the similarity of this effect on multiple outcomes may seem puzzling, it perhaps indicates the centrality of high school science course completion to students' entry to the natural sciences at some point early in college, even if it does not singularly predict persistence.

Looking specifically at measures related to our third research question, we see that all perceived ability under mathematics challenge measures—growth mindset, 10th grade mathematics index, and 12th grade mathematics index—positively and significantly predict staying in PEMC and/or biology fields as intended before entering postsecondary education, net of all other factors. The 10th grade mathematics index has the largest effect size here, predicting a 61.7% increased risk of staying in PEMC and/or biology fields as intended ($RRR = 1.62$; $p = 0.000$) vs. never having entered those fields, compared to 56.1% for the 12th grade mathematics index ($RRR = 1.56$; $p = 0.000$), and 35.4% for growth mindset ($RRR = 1.35$; $p = 0.021$). Also consistent with the literature, there is a stronger negative effect on staying in PEMC

and/or biology for the verbal index ($RRR = 0.58$; $p = 0.000$) vs. to leaving ($RRR = 0.70$; $p = 0.001$), compared to abstaining. However, surprisingly, the 10th grade mathematics index also predicts leaving these fields ($RRR = 1.28$; $p = 0.045$). Moreover, both the general index and the 12th grade mathematics index predict new entry to PEMC and/or biology fields 2 years after high school ($RRR_{\text{general index}} = 1.48$; $p = 0.017$ and $RRR_{\text{12th grade mathematics index}} = 1.33$; $p = 0.012$). This finding suggests that either domain-general or mathematics-domain perceived ability to overcome challenge might actually encourage students to cross into mathematics-intensive fields of study from non-STEM fields.

Finally, as with the analysis on science pipeline, we examined a model (not shown) including interactions between gender and each of the perceived ability under challenge variables. The resulting coefficients were not significant, so due to space constraints we decided not to show this particular model. Possible explanations for the lack of significance on gender and perceived ability under challenge interaction terms for the major retention variable will be unpacked in the Discussion Section of this paper.

Specific Scientific Major and Perceptions of Ability to Overcome Challenge

Finally, we examined the relationship between perceived ability under challenge and choice of major 2 years after high school. Since we are primarily interested in how ability-related beliefs might encourage or deter students to major in more mathematics-intensive fields, we disaggregated STEM majors and used non-STEM majors as our reference category. As with the analysis on the major retention variable, we report findings using RRR in **Table 6**.

First, we turn to the main effect of gender, which is strongest as a predictor of PEMC and health sciences majors, albeit in opposite directions. While only the third model is shown in **Table 6**, tables reporting the earlier models are available by request. In the first model, including only demographic characteristics, women have a 0.78 times lower risk than men of majoring in PEMC ($RRR = 0.22$; $p = 0.000$)⁵ and a 3.59 times higher risk than men of majoring in health ($p = 0.000$), as compared to non-STEM fields. When student ability and institutional effects are added in the second model, women's risk of majoring in PEMC declines slightly as the risk ratio becomes more negative ($RRR = 0.20$; $p = 0.000$), but their risk of majoring in health does not meaningfully change ($RRR = 3.59$; $p = 0.000$). The third model adds perceptions under challenge to the model and has a resulting decrease in the negative effect of female gender on the risk of majoring in PEMC ($RRR = 0.26$; $p = 0.000$) and an increase in the positive effect of female gender on the risk of majoring in health sciences ($RRR = 3.69$; $p = 0.000$). Adding perceived ability under challenge variables to our models therefore enhances women's chances (relative to men) of majoring in both PEMC and health fields.

⁵Note that the effect of gender for men can be found by taking the inverse of these relative ratios. Here, the effect of male gender on the risk of majoring in PEMC is $1/0.22$, or 4.55.

TABLE 6 | Specific STEM major category declared 2 years after high school, not including interaction effects.

Variables	Non-STEM (reference)									
	Undeclared/ Undecided		PEMC		Biological sciences		Health sciences		Soc/Behavioral and other sciences	
	RRR	SE	RRR	SE	RRR	SE	RRR	SE	RRR	SE
DEMOGRAPHIC CHARACTERISTICS										
Female gender (reference = male)	0.891	0.099	0.264***	0.048	0.953	0.206	3.691***	0.696	1.009	0.139
Race/Ethnicity (reference = White)										
Asian/Pacific Islander	1.531*	0.276	1.113	0.321	2.497*	0.904	2.498**	0.698	1.423	0.377
Black	1.246	0.332	3.187***	1.006	3.394***	1.196	1.892*	0.504	1.842*	0.522
Latino	1.255	0.298	1.340	0.437	1.397	0.703	1.116	0.323	2.088**	0.557
Multi-race/Ethnicity	1.041	0.323	1.201	0.486	0.967	0.727	1.680	0.659	0.794	0.367
STUDENT ABILITY										
Science Pipeline Completion (reference = Chemistry I or Physics I or less)										
Chem. I and Physics I	1.206	0.165	1.850**	0.396	1.072	0.331	0.960	0.207	1.030	0.169
Chemistry II and Physics II	1.675***	0.258	2.503***	0.512	3.875***	1.025	1.640*	0.314	1.385	0.250
INSTITUTIONAL EFFECTS										
College Selectivity (reference = 4-year institution, highly selective)										
2-year or less institution	1.131	0.210	0.716	0.181	0.595	0.245	2.409***	0.593	0.356***	0.087
4-year institution, inclusive	0.399***	0.085	0.878	0.223	0.440*	0.168	1.807*	0.486	0.447**	0.110
4-year institution, moderately selective	0.587***	0.093	0.708	0.132	1.062	0.293	1.231	0.287	0.728	0.123
PERCEIVED ABILITY UNDER CHALLENGE										
Non-Mathematics Measures										
General index (10th grade)	1.035	0.090	1.023	0.153	1.583**	0.269	1.315*	0.173	1.046	0.126
Verbal index (10th grade)	0.990	0.078	0.629***	0.064	0.838	0.122	0.808*	0.082	1.280*	0.131
Mathematics Measures										
Growth mindset (10th grade)	0.993	0.091	1.340*	0.173	0.953	0.159	1.052	0.104	1.002	0.109
Mathematics index (10th grade)	1.054	0.087	1.599***	0.187	1.030	0.125	1.014	0.097	1.027	0.089
Mathematics index (12th grade)	1.014	0.069	1.542***	0.174	1.259	0.164	1.054	0.096	0.948	0.091
Constant	0.821	0.349	0.047***	0.026	0.014***	0.009	0.031***	0.017	0.233**	0.115
f-statistic	9.930***		9.930***		9.930***		9.930***		9.930***	
Observations	4450		4450		4450		4450		4450	

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002/2006 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Parent education, family income, ability with complex material, 10th grade GPA, high school region and high school urbanicity were included in the model, but are withheld from this table for space. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Race/ethnicity again plays a role, here influencing declared major 2 years after high school. Holding everything else constant, black students had a 3.19 times higher risk than their white peers of majoring in PEMC as compared to non-STEM fields ($p = 0.000$); they had a 3.39 times higher risk than their white peers of majoring in biology fields ($p = 0.001$). Latinos had a 2.09 times higher risk than their white peers of majoring social/behavioral or other science fields ($p = 0.006$), as compared to non-STEM majors. Asian/Pacific Islander students were at a 2.50 times higher risk than their white peers of majoring in biology ($p = 0.012$) and health ($p = 0.001$), respectively, as compared to non-STEM majors.

Results on student ability and course taking were congruent with the previous analysis of major retention. Tenth grade GPA, net of all other effects, significantly and positively predicted the selection of PEMC majors ($RRR = 1.45$; $p = 0.009$)

and biology majors ($RRR = 1.41$; $p = 0.044$) vs. non-STEM majors. In contrast, GPA negatively predicted the selection of undeclared/undecided majors ($RRR = 0.75$; $p = 0.000$), showing that high achieving high school students in our sample tended to select a major by 2 years after high school. Next, the single highest predictor of any major type, holding all other factors constant, was completion of chemistry II and physics II for selection of a biology major ($RRR = 3.88$; $p = 0.000$). Completion of chemistry II and physics II also increased the risk of enrolling in PEMC fields vs. non-STEM fields ($RRR = 2.50$; $p = 0.000$), compared to students who only completed chemistry I or physics I or less. Completing even the middle category of the science pipeline variable also benefitted students, predicting an 85.0% increase in the risk of selecting a PEMC major vs. a non-STEM major ($p = 0.005$), as compared to students who only completed chemistry I or physics I and below in high school.

With respect to institutional effects, high school region and college selectivity were the only notable factors influencing choice of major. Students attending high schools in the Midwest and the South were more likely than their peers in the Northeast to select health sciences majors, as compared to non-STEM majors (full table available by request). Attending a less selective institution decreases students' risk of declaring social/behavioral and other science majors, as compared to non-STEM majors. By contrast, their risk of majoring in health sciences increases, in comparison to non-STEM majors. Together, these results suggest that institutional contexts can influence choice of major, in particular health science fields.

Using the product-term regression method (Jaccard and Turrisi, 2003), we can interpret the interactions between gender and perceived ability under challenge measures as slope differences between men and women. In contrast, the main effects for perceived ability under challenge represent the effects of these perceptions for the reference category on gender. Because this manuscript is primarily concerned with how these perceptions influence *women's* entry into scientific majors, we report the results for the case when the reference category for the

gender variable is female, so that the main effects of perceived ability under challenge represent the effect for women in particular.

We now turn to the version of the full model shown in **Table 7**, with women as the reference category and interactions between gender and the perceived ability under challenge variables. Because our perceived ability under challenge variables are mean-centered, a value of 0 refers to the mean value for each of these terms, for the reference category (in this case, women). In this multinomial logistic regression model then, men have a 3.60 times higher risk of majoring in PEMC than women with average perceived ability under challenge ($p = 0.000$) and a 0.74 times lower risk of majoring in health than women with average perceived ability under challenge ($p = 0.000$), again as compared to non-STEM fields. In sum then, holding all other predictors constant, gender strongly influences students' choice of PEMC and health sciences majors. Gender does not however notably influence choice of biological nor social/behavioral and other sciences majors, as compared to non-STEM majors.

Recall how in **Figure 1**, we show our intent to examine how gender moderates the relationship between perceptions of ability

TABLE 7 | Specific STEM major category declared 2 years after high school, interaction model.

	Non-STEM (reference)									
	Undeclared/ Undecided		PEMC		Biological sciences		Health sciences		Soc/Behavioral and other sciences	
	RRR	SE	RRR	SE	RRR	SE	RRR	SE	RRR	SE
DEMOGRAPHIC CHARACTERISTICS										
Male gender (Reference = female)	1.085	0.122	3.604***	0.806	1.116	0.286	0.257***	0.050	1.032	0.162
PERCEIVED ABILITY UNDER CHALLENGE										
Non-Mathematics Measures										
General index (10th grade)	1.093	0.137	1.118	0.255	1.746*	0.417	1.351*	0.193	1.233	0.189
Verbal index (10th grade)	0.894	0.086	0.647*	0.119	0.766	0.134	0.762*	0.089	1.266	0.166
Mathematics Measures										
Growth mindset (10th grade)	1.070	0.134	1.140	0.269	0.891	0.247	1.212	0.136	1.072	0.166
Mathematics index (10th grade)	1.056	0.113	1.360	0.250	1.095	0.176	0.991	0.106	0.950	0.106
Mathematics index (12th grade)	0.931	0.090	1.650**	0.285	1.193	0.195	1.040	0.110	0.924	0.109
INTERACTIONS BETWEEN GENDER AND PERCEIVED ABILITY UNDER CHALLENGE										
Non-Mathematics Measures										
Male*General index (10th grade)	0.884	0.182	0.832	0.223	0.780	0.271	1.029	0.349	0.663	0.142
Male*Verbal index (10th grade)	1.275	0.178	1.003	0.228	1.273	0.354	1.303	0.308	1.037	0.212
Mathematics Measures										
Male*Growth mindset (10th grade)	0.844	0.162	1.197	0.332	1.148	0.442	0.505**	0.113	0.870	0.187
Male*Mathematics index (10th grade)	0.991	0.155	1.284	0.297	0.850	0.238	1.051	0.257	1.222	0.223
Male*Mathematics index (12th grade)	1.207	0.176	0.934	0.200	1.147	0.304	0.977	0.226	1.056	0.176
Constant	0.730	0.330	0.012***	0.007	0.013***	0.009	0.114***	0.063	0.224**	0.116
<i>f</i> -statistic	7.790***		7.790***		7.790***		7.790***		7.790***	
Observations	4450		4450		4450		4450		4450	

n = 4450 respondents from the National Center for Education Statistics' Education Longitudinal Study 2002/2006 restricted data. Student-level replicate weights particular to the base year through 2nd follow-up (2002/2006) waves were used to enhance the correspondence between sample and population results. Family income, parent education, ability with complex material, 10th grade GPA, science pipeline completion, high school region and urbanicity, and college selectivity were included in the model, but are withheld from this table for space. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

under challenge and major choice. Notably, the main effect of the 12th grade mathematics index is the most notable significant perceived ability under mathematics challenge predictor for women, increasing their risk of majoring in PEMC (RRR = 1.65; $p = 0.004$) compared to a non-STEM field⁶. The magnitude and significance of these effects may be somewhat muted, given that there are two indicators in the model for mathematics index (in 10th and 12th grades). This significant result is therefore likely a conservative estimate. To more meaningfully interpret this finding, we used the `prgen` command from `SPost9` (Long and Freese, 2005) to estimate the predicted probabilities for women's selection of each of the major types, given their score on the 12th grade mathematics index. **Figure 2** shows the predicted outcomes on a line graph, for each STEM major category. We see that an increase in perceived ability under challenge in mathematics domains meaningfully changes women's probability of declaring PEMC, biology, and social/behavioral and other sciences. Notably, as women's perceived ability increases, their chances of majoring in social/behavioral and other sciences *decreases*. The opposite is true for PEMC and biology. In particular, women's probability of majoring in PEMC increases in association with an increase in their 12th grade perceptions that they could understand and master difficult and complex mathematics material. Specifically, their probability of majoring in PEMC *rises over and above* that of majoring in biology by the point that their perceptions are one unit above the mean for women in our sample.

⁶The RRR for women, shown in **Table 7**, is smaller than the result for men (RRR = 1.542; $p = 0.004$), when men are in the reference category (table available by request).

Rounding out our discussion of how perceived ability under challenge affects women's choice of major, there are two additional findings of note. Domain-general perceptions positively influence the selection of a STEM field in two other instances: biology (RRR = 1.75; $p = 0.021$) and health science fields (RRR = 1.35; $p = 0.037$). Perceived ability in verbal domains also negatively predicts women's entry into PEMC (RRR = 0.65; $p = 0.019$) and health sciences (RRR = 0.76; $p = 0.020$).

The interaction terms at the bottom of **Table 7** examine the differential impact of gender on perceived ability under challenge. Only one of these interactions is significant in its effect. The male*growth mindset interaction term (RRR = 0.51, $p = 0.003$) indicates that gender moderates the effect of growth mindset on students' choice of health science majors as compared to non-STEM majors. This finding indicates that the belief that anyone can improve their mathematics ability through mastery-oriented behavior (growth mindset) differentially effects men and women in a way that promotes women's selection into health science fields. We again use `prgen` to estimate the predicted probabilities for women's selection of each of the major types, shown in **Figure 3**, given their score on the growth mindset variable. Consistent with the discussion above, women have both a higher and increasing probability of selecting a health science field as their growth mindset score increases, as compared to the other STEM majors. While the effects are not significant, a sizeable enough increase in growth mindset (a half-point above the mean) appears to positively increase the probability such that—all else held constant—women would have a higher likelihood of majoring in PEMC than they would of majoring in biology. This finding further suggests that there are meaningful, tangible implications for enhancing women's perceptions of their ability under challenge.

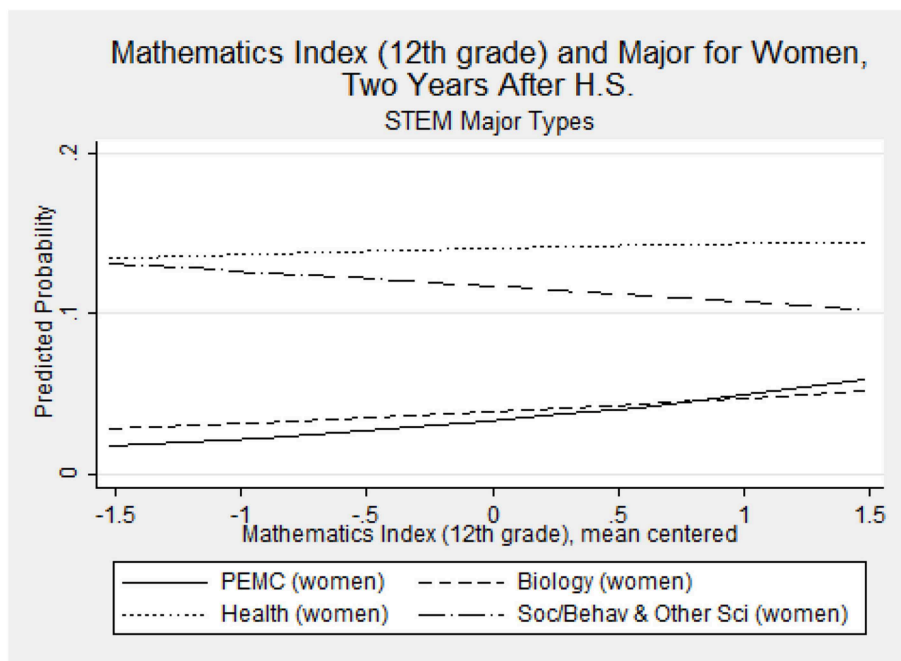


FIGURE 2 | Mathematics index (12th grade) and probability of majoring in specific STEM majors.

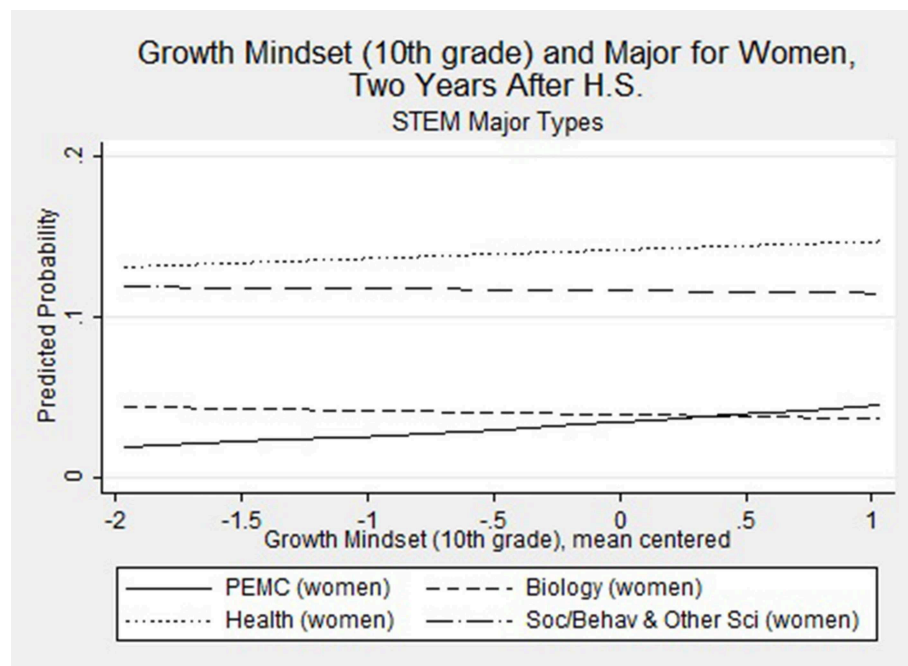


FIGURE 3 | Growth mindset and probability of majoring in specific STEM majors.

Discussion

Limitations

Similar to all studies using secondary data sources, our interpretations are limited by the self-reported nature of the data. For instance, our analysis on major retention was limited because students were retrospectively asked the intended major question 2 years after leaving high school. This measure may be biased by their subsequent choice of major. Additionally, this question focused on students' *intent*, not their actual declared major upon entrance into the institutions. While this gives us some insight, declared major symbolizes commitment and would allow us to be reasonably sure that students participated in gateway coursework in the declared major. Further, the coding of the intended major variable did not permit us to disaggregate PEMC fields from biology in the measurement of students' intended majors. As noted by previous researchers, women tend to be overrepresented in biology fields (NSF, 2013), yet we cannot adequately separate out the effects of staying in biology from staying in PEMC fields. Finally, because we do not currently have information on degree completion, our analyses are limited to students' experiences up through 2 years after high school.

Conclusions

In response to our research questions, we found mixed support for our hypotheses that perceived ability under challenge in mathematics is related to our outcomes of interest: completing advanced science coursework, remaining in intended STEM major fields, and selecting mathematics-intensive science majors

(PEMC). Importantly, both gender and perceived ability under challenge in mathematics influence our prediction of all three outcomes. In addition, 10th grade perceptions of ability under challenge in mathematics positively predict completion of the highest levels of high school science coursework. Moreover, all mathematics perceived ability under challenge measures predict both retention in PEMC and/or biology fields, holding all other factors constant. Finally, in some cases, perceptions of ability under challenge affect women's selection of PEMC and other STEM majors.

Turning first to descriptive differences in high school, women and men's perceived ability under challenge differed, with young men in our sample outscoring young women in all perceived ability under mathematics challenge measures. Intriguingly, while the gender gap in perceived mathematics ability seems to taper during high school, this change seems driven by changes among boys rather than girls. Specifically, boys' perceived ability in mathematics decreased between 10th and 12th grade, while girls' perceived ability stayed constant. This finding suggests the need for further empirical and conceptual studies of boys' experiences in mathematics courses in high school, as their relative strengths in this area have been presumed undeserving of examination.

Next, we turn to the predictions of high school course taking. Our results indicate that perceived ability under mathematics challenge in 10th grade matters, and in fact was the only predictive subjective measure (i.e., beyond demographics and ability test scores) of taking advanced science coursework. Female gender negatively predicts advanced science course taking. While recent research suggests that girls are increasingly successful

in secondary and postsecondary education, including science course completion (Hill et al., 2010; DiPrete and Buchmann, 2013), our results indicate that gender gaps in course taking remain. However, there were no significant findings for the interaction terms in this analysis, suggesting that something other than perceived ability is at work. Indeed, performance indicators of ability—not perceptions—appeared to particularly influence students' course taking.

Future research may be needed to investigate the mechanisms by which students—girls in particular—are advised into and choose to enroll in a second year of both chemistry and physics, which over 25% of our sample elects to do. These decisions have clear ramifications for entering and choosing PEMC and biology majors, as indicated in our findings reported above. Our negative findings for both western and rural measures of high school location suggest that access to higher-level science coursework is differentially distributed around the U.S. and likely varies by the profiles of students' high schools, not limited to region and urbanicity. For instance, Riegle-Crumb and Moore (2014) show how the density of female STEM professionals in the neighborhoods surrounding schools can mitigate the traditional negative relationship between gender and high school physics course taking. Moreover, recent work by Legewie and DiPrete (2014) on U.S. high school students in the early 1990s indicates that school-level curricular and extra-curricular offerings considerably explain the gender gap in intention to major in STEM at the end of high school. Extensive research and policy initiatives have examined increasing access to advanced mathematics courses. This study suggests that similar attention should be paid to increasing access to advanced science coursework in secondary school, physics in particular.

Despite the number of adequately prepared women entering postsecondary education, we know that fewer of them persist in STEM fields (NSF, 2013). Therefore, we turn next to the matter of how perceived ability under challenge might be related to majoring in PEMC and/or biology as intended at enrollment. As mentioned before, perceived ability under challenge in mathematics (growth mindset, 10th grade mathematics index, and 12th grade mathematics index) is positively related to staying in PEMC and/or biology fields, net of all other effects. This suggests that increasing students' confidence in their ability to deal with difficult mathematics material may lead to retention in those fields. However, there were no significant findings for the interactions between gender and these measures on any level of the retention variable, suggesting that gender does not influence the impact of perceived ability under challenge on retention in students' intended major. These modest results may be the consequence of our limited ability to parse out the PEMC and biology categories⁷, as these STEM fields currently have highly distinct patterns of sex segregation at the undergraduate level, as demonstrated in **Table 1**.

⁷Notably, George-Jackson's (2011) study found that while women persist in PEMC-related fields at lower rates than men, only 24.5% of those who initially chose these majors switched to non-STEM fields; 11.4% switched into biology, health, behavioral, and related science fields. There may be considerable movement among the PEMC and/or biology category that we are not able to observe because of limitations in the coding of the intended major variable.

Mathematics is not the only domain in which perceived ability influences choice of major. As reported in **Table 5**, we see a significant and negative relationship between perceived ability under challenge in the verbal domain and persistence in PEMC and/or biology. Similarly in **Table 7** (and corresponding results in **Table 6**), perceived ability in verbal domains negatively predicts women's entry into PEMC and health sciences. These findings are consistent with previous literature suggesting that perceived high verbal ability may act as a stronger influence on major choice than actual mathematics ability (Correll, 2001; Wang et al., 2013). Related, we also found that domain-general perceived ability under challenge has a more positive relationship with entering PEMC and/or biology fields than the 12th grade mathematics index, holding all other factors constant. Again, later results on declared majors show that domain-general perceptions positively influence women's selection of biology and health science fields. These results lead us to wonder how domain-general perceived ability may increase interest in certain STEM fields. Future studies, perhaps qualitative in nature, may unpack the mechanisms behind this perhaps puzzling finding.

We were able to disaggregate specific major types in our analysis of the relationships between perceived ability under challenge and declared major 2 years after high school. Results compared across models revealed the specificity of the relationship between gender and each STEM major category. The effects of perceived ability under challenge reported in **Table 6** (without interactions) are robust and in the expected direction with respect to the effects on science majors. Also of note are our findings regarding high school region and college selectivity with respect to health fields. With respect to the latter, it is unclear whether it is the institutional context itself or selection into certain institutions that drives the negative relationship between selectivity and health science majors (and correspondingly, the positive relationship between selectivity and social/behavioral and other science majors). As previous research on this topic is inconclusive, this is again an area for potential further investigation.

With respect to the hypothesized moderating effect of gender, the gender-specific results reported in **Table 7** did not neatly correspond to our hypotheses. Importantly, we did find an effect for women's selection of a PEMC field when looking at the main effect of the 12th grade mathematics index. Notably, an increase in women's perceived ability with difficult and complex mathematics material increases their probability of majoring in PEMC, such that they become more likely to major in PEMC than in biology. This is notable, as PEMC fields are those that have thus far been the most persistently sex segregated STEM disciplines. As biology and health fields have become more gender egalitarian, and even female-dominant in recent years, this result suggests that interventions aimed at enhancing secondary school girls' perceptions of their mathematics ability can have real effects on their participation in mathematics-intensive fields in postsecondary school, and preventing the loss of scientific talent among young women.

Examining our results on gender moderation further, we found positive gender moderation on the effect of growth mindset for selection into only one STEM field: health sciences. It may

be that the intensive and cumulative investment of girls and boys on the scientific pipeline may track those girls with more negative ability-related beliefs out of PEMC fields before they select college majors (Perez-Felkner et al., 2012). Notwithstanding, the effects of perceived ability under challenge for women among mathematics, verbal, and general domains, as well as this finding regarding gender moderation on growth mindset, indicate that there are indeed notable effects to consider and continue to investigate.

Intriguingly, the predicted probabilities shown in **Figure 3** indicate that a positive enough mindset among women will increase their probability of majoring in PEMC, even over and above their probability of majoring in biology. Because we did not find significance on the interaction effects between gender and growth mindset on PEMC, we cannot be sure that women who believe that anyone can develop their mathematics ability will enter PEMC majors, a finding seemingly inconsistent with mindset theory (Dweck, 2008; Good et al., 2012; Mangels et al., 2012). It could be that mathematics-intensive fields, such as PEMC, are losing growth mindset women as a result of environmental factors, such as messages that they would fit better or be happier elsewhere (such as the health science fields). These messages may foster stereotype threat.

Stereotype threat occurs when individuals with stereotyped identities fear that they will confirm negative stereotypes (Steele, 1997), and has been widely discussed related to women's choices to leave STEM fields. It is possible that the null findings on most of our interaction terms are masked by the effects of stereotype threat—something we could not directly measure. Although gender did not consistently moderate the relationship between

perceived ability under mathematics challenge and our dependent variables, there were strong gender differences in perceptions under challenge across our results, from secondary school through the early postsecondary years. Moreover, while gender did not show a consistent moderating effect, this may be the case for race/ethnicity—a topic beyond the scope of this paper, though no less important to the issue of increasing participation in STEM. Future studies using similar constructs would benefit from additional analyses on the interactive effects of race/ethnicity and perceptions of ability to overcome challenge in pathways to mathematics and science careers.

Acknowledgments

This material is based upon work supported by Florida State University's Center for Higher Education Research, Teaching, and Innovation as well as the National Science Foundation, under Grant No. 1232139. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We appreciate the clear and sophisticated feedback from our reviewers as well as the support of our undergraduate research assistants Mitchell D'Sa, Melissa Magalhaes, and Richea Osei in preparing this manuscript.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.00530/abstract>

References

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84:191. doi: 10.1037/0033-295X.84.2.191
- Beilock, S. L. (2008). Math performance in stressful situations. *Curr. Dir. Psychol. Sci.* 17, 339–343. doi: 10.1111/j.1467-8721.2008.00602.x
- 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.* New York, NY: Rockefeller Foundation.
- Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinomial Models.* Thousand Oaks, CA: SAGE Publications, Inc.
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Cole, D., and Espinoza, A. (2008). Examining the academic success of latino students in science technology engineering and mathematics (STEM) majors. *J. Coll. Stud. Dev.* 49, 285–300. doi: 10.1353/csd.0.0018
- Coleman, J. S., Hoffer, T., and Kilgore, S. (1982). *High School Achievement: Public, Catholic, and Private Schools Compared.* New York, NY: Basic Books.
- Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. Sociol.* 106, 1691–1730. doi: 10.1086/321299
- Csikszentmihályi, M., and Csikszentmihályi, I. S. (eds.). (1988). *Optimal Experience: Psychological Studies of Flow in Consciousness.* New York, NY: Cambridge University Press.
- Csikszentmihályi, M., and Csikszentmihályi, M. (1991). *Flow: The Psychology of Optimal Experience*, Vol. 41. New York, NY: HarperPerennial.
- Csikszentmihályi, M., and Schneider, B. L. (2000). *Becoming Adult: How Teenagers Prepare for the World of Work.* New York, NY: Basic Books.
- Davenport, E. C., Davison, M. L., Kuang, H., Ding, S., Kim, S.-K., and Kwak, N. (1998). High school mathematics course-taking by gender and ethnicity. *Am. Educ. Res. J.* 35, 497–514. doi: 10.3102/00028312035003497
- DeCaro, M. S., Rotar, K. E., Kendra, M. S., and Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on mathematical problem solving. *Q. J. Exp. Psychol.* 63, 1619–1630. doi: 10.1080/17470210903474286
- DiPrete, T. A., and Buchmann, C. (2013). *The Rise of Women. The Growing Gender Gap in Education and what it Means for American Schools.* New York, NY: Russell Sage Foundation.
- Drew, C. (2011, November 4). Why science majors change their minds (it's just so darn hard). *The New York Times*. Available online at: <http://www.nytimes.com/2011/11/06/education/edlife/why-science-majors-change-their-mind-its-just-so-darn-hard.html?pagewanted=all>
- Dweck, C. (2000). *Self-Theories: Their Role in Motivation, Personality, and Development.* New York, NY: Psychology Press.
- Dweck, C. (2006). *Mindset: The New Psychology of Success.* New York, NY: Random House.
- Dweck, C. (2007). "Is math a gift? Beliefs that put females at risk," in *Why Aren't More Women in Science? Top Researchers Debate the Evidence*, eds S. C. Ceci and W. M. Williams (Washington, DC: American Psychological Association), 47–55.
- Dweck, C. (2008). *Mindsets and Math/Science Achievement.* New York, NY: Carnegie Corp. of New York-Institute for Advanced Study Commission on Mathematics and Science Education.
- Engberg, M. E., and Wolniak, G. C. (2013). College student pathways to the STEM disciplines. *Teach. Coll. Rec.* 115, 1–27.

- George-Jackson, C. E. (2011). STEM switching: examining departures of undergraduate women in STEM fields. 17, 149–171. doi: 10.1615/JWomenMinorScienEng.2011002912
- Good, C., Aronson, J., and Inzlicht, M. (2003). Improving adolescents' standardized test performance: an intervention to reduce the effects of stereotype threat. *J. Appl. Dev. Psychol.* 24, 645–662. doi: 10.1016/j.appdev.2003.09.002
- Good, C., Rattan, A., and Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *J. Pers. Soc. Psychol.* 102, 700–717. doi: 10.1037/a0026659
- Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: is it the school that matters? *Econ. Educ. Rev.* 29, 911–922. doi: 10.1016/j.econedurev.2010.06.010
- Griffith, A. L., and Rothstein, D. S. (2009). Can't get there from here: the decision to apply to a selective college. *Econ. Educ. Rev.* 28, 620–628. doi: 10.1016/j.econedurev.2009.01.004
- Hanson, S. (2008). *Swimming Against the Tide: African American Girls and Science Education*. Philadelphia, PA: Temple Univ Press.
- Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. doi: 10.1126/science.7604277
- Hill, C., Corbett, C., and St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: American Association of University Women. Available online at: <http://eric.ed.gov/PDFS/ED509653.pdf>
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., and Williams, C. C. (2008). Gender similarities characterize math performance. *Science* 321, 494–495. doi: 10.1126/science.1160364
- Hyde, J. S., and Linn, M. C. (2006). Gender similarities in mathematics and science. *Sci. Mag.* 314, 599. doi: 10.1126/science.1132154
- Ingels, S. J., Pratt, D. J., Wilson, D., Burns, L. J., Currivan, D., Rogers, J. E., et al. (2007). *Educational Longitudinal Study of 2002 (ELS: 2002): Base-year to second follow-up data file documentation (Report No. NCES 2008-347)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Jaccard, J., and Turrisi, R. (2003). *Two-Way Interactions Interaction Effects in Multiple Regression, 2nd Edn*. Thousand Oaks, CA: Sage Research Methods.
- Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., et al. (2014). *The Condition of Education 2014*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Legewie, J., and DiPrete, T. A. (2014). The high school environment and the gender gap in science and engineering. *Sociol. Educ.* 87, 259–280. doi: 10.1177/0038040714547770
- Long, J. S., and Freese, J. (2005). *Regression Models for Categorical Outcomes Using Stata, 2nd Edn*. College Station, TX: Stata Press.
- López Turley, R. N. (2009). College proximity: mapping access to opportunity. *Sociol. Educ.* 82, 126–146. doi: 10.1177/003804070908200202
- Ma, Y. (2011). Gender differences in the paths leading to a STEM baccalaureate. *Soc. Sci. Q.* 92, 1169–1190. doi: 10.1111/j.1540-6237.2011.00813.x
- Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., and Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Soc. Cogn. Affect. Neurosci.* 7, 230–241. doi: 10.1093/scan/nsq100
- Morgan, S. L., Gelbgiser, D., and Weeden, K. A. (2013). Feeding the pipeline: gender, occupational plans, and college major selection. *Soc. Sci. Res.* 42, 989–1005. doi: 10.1016/j.ssresearch.2013.03.008
- Nakamura, J., and Csikszentmihalyi, M. (2002). "The concept of flow," in *Handbook of Positive Psychology* eds C. R. Snyder and S. J. Lopez (New York, NY: Oxford University Press), 89–105.
- NSF. (2013). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2013*. Arlington, TX: National Science Foundation (NSF) National Center for Science and Engineering Statistics.
- Niu, S. X., and Tienda, M. (2013). High school economic composition and college persistence. *Res. High. Educ.* 54, 30–62. doi: 10.1007/s11162-012-9265-4
- OECD. (2013). *Education at a Glance 2013: OECD Indicators*. Paris: Organisation of Economic Cooperation and Development.
- Ost, B. (2010). The role of peers and grades in determining major persistence in the sciences. *Econ. Educ. Rev.* 29, 923–934. doi: 10.1016/j.econedurev.2010.06.011
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578. doi: 10.3102/0034654306004543
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., and Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Dev. Psychol.* 48, 1629. doi: 10.1037/a0029167
- Perez-Felkner, L. (in press). Perceptions and resilience in underrepresented students' pathways to college. [Feature Article]. *Teachers College Record* 117.
- Perez-Felkner, L., McDonald, S.-K., Schneider, B., and Grogan, E. (2012). Female and male adolescents' subjective orientations to mathematics and the influence of those orientations on postsecondary majors. *Dev. Psychol.* 48, 1658. doi: 10.1037/a0027020
- Perez-Felkner, L., and Schneider, B. (2012). *It's not 'Leaking Out' but not Entering: Mathematics Ability, Career Attainment Values, and the Gender Gap in Scientific Fields*. Las Vegas: Paper presented at the Association for the Study of Higher Education.
- Perez-Felkner, L., McDonald, S.-K., and Schneider, B. L. (2014). "What happens to high-achieving females after high school? Gender and persistence on the post-secondary STEM pipeline," in *Gender Differences in Aspirations and Attainment: A Life Course Perspective*, eds I. Schoon and J. S. Eccles (Cambridge: Cambridge University Press), 285–320.
- Putman, H., Greenberg, J., and Walsh, K. (2014). *Training Our Future Teachers: Easy As and What's Behind Them*. Teacher Prep Review Washington, DC: National Council on Teacher Quality.
- Rask, K. (2010). Attrition in STEM fields at a liberal arts college: the importance of grades and pre-collegiate preferences. *Econ. Educ. Rev.* 29, 892–900. doi: 10.1016/j.econedurev.2010.06.013
- Riegle-Crumb, C. (2006). The path through math: course sequences and academic performance at the intersection of race-ethnicity and gender. *Am. J. Educ.* 113, 101–122. doi: 10.1086/506495
- Riegle-Crumb, C., Farkas, G., and Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociol. Educ.* 79, 206–228. doi: 10.1177/003804070607900302
- Riegle-Crumb, C., King, B., Grodsky, E., and Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM College majors over time. *Am. Educ. Res. J.* 49, 1048–1073. doi: 10.3102/0002831211435229
- Riegle-Crumb, C., and Moore, C. (2014). The gender gap in high school physics: considering the context of local communities. *Soc. Sci. Q.* 95, 253–268. doi: 10.1111/ssqu.12022
- Rittmayer, M. A., and Beier, M. E. (2009). "Self-efficacy in STEM," in *Applying Research to Practice (ARP) Resources*. Available online at: <http://www.engr.psu.edu/AWE/ARPresources.aspx>
- Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *J. Bus. Econ. Stat.* 13, 217–224. doi: 10.1080/07350015.1995.10524596
- Rydell, R. J., McConnell, A. R., and Beilock, S. L. (2009). Multiple social identities and stereotype threat: imbalance, accessibility, and working memory. *J. Pers. Soc. Psychol.* 96, 949. doi: 10.1037/a0014846
- Sax, L. (1994). Mathematical self-concept: how college reinforces the gender gap. *Res. High. Educ.* 35, 141–166. doi: 10.1007/BF02496699
- Seymour, E. (1999). The role of socialization in shaping the career-related choices of undergraduate women in science, mathematics, and engineering majors. *Ann. N. Y. Acad. Sci.* 869, 118–126. doi: 10.1111/j.1749-6632.1999.tb08366.x
- Sherhoff, D. J., Csikszentmihályi, M., Schneider, B. L., and Sherhoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *Sch. Psychol. Q.* 18, 158–176. doi: 10.1521/scpq.18.2.158.21860
- Smyth, F. L., and McArdle, J. J. (2004). Ethnic and gender differences in science graduation at selective colleges with implications for admission policy and college choice. *Res. High. Educ.* 45, 353–381. doi: 10.1023/B:RIHE.0000027391.05986.79
- Stearns, E., Jha, N., Giersch, J., and Mickelson, R. A. (2013). "Do gender differences in first-year grades contribute to the gender gap in STEM?," in *Paper Presented at the American Sociological Association, Annual Meeting* (New York, NY).
- Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037/0003-066X.52.6.613

- Summers, L. H. (2005). "Remarks at NBER conference on diversifying the science & engineering workforce," in *Paper Presented at the NBER Conference on Diversifying the Science & Engineering Workforce* (Cambridge, MA).
- Vogt, W. P. (2005). "Risk ratio," in *Dictionary of Statistics & Methodology*, 3rd Edn., ed W. P. Vogt (Thousand Oaks, CA: SAGE Publications, Inc.), 280–281.
- Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937
- Zimmerman, B. J. (2000). Self-efficacy: an essential motive to learn. *Contemp. Educ. Psychol.* 25, 82–91. doi: 10.1006/ceps.1999.1016

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Nix, Perez-Felkner and Thomas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Does Gender of Administrator Matter? National Study Explores U.S. University Administrators' Attitudes About Retaining Women Professors in STEM

Wendy M. Williams*, Agrima Mahajan, Felix Thoemmes, Susan M. Barnett, Francoise Vermeylen, Brian M. Cash and Stephen J. Ceci

Department of Human Development, Cornell University, Ithaca, NY, United States

OPEN ACCESS

Edited by:

Jessica S. Horst,
University of Sussex, United Kingdom

Reviewed by:

Jonathan Wai,
Duke University, United States
Silke Schickanz,
University of Göttingen, Germany

*Correspondence:

Wendy M. Williams
wendywilliams@cornell.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 April 2016

Accepted: 21 April 2017

Published: 22 May 2017

Citation:

Williams WM, Mahajan A,
Thoemmes F, Barnett SM,
Vermeylen F, Cash BM and Ceci SJ
(2017) Does Gender of Administrator
Matter? National Study Explores U.S.
University Administrators' Attitudes
About Retaining Women Professors in
STEM. *Front. Psychol.* 8:700.
doi: 10.3389/fpsyg.2017.00700

Omnipresent calls for more women in university administration presume women will prioritize using resources and power to increase female representation, especially in STEM fields where women are most underrepresented. However, empirical evidence is lacking for systematic differences in female vs. male administrators' attitudes. Do female administrators agree on which strategies are best, and do men see things differently? We explored United States college and university administrators' opinions regarding strategies, policies, and structural changes in their organizations designed to increase women professors' representation and retention in STEM fields. A comprehensive review of past research yielded a database of potentially-effective, recommended policies. A survey based on these policies was sent to provosts, deans, associate deans, and department chairs of STEM fields at 96 public and private research universities across the U.S. These administrators were asked to rate the quality and feasibility of each strategy; 474 provided data, of which 334 contained complete numerical data used in the analyses. Our data revealed that female (vs. male) administrators believed the 44 strategies were higher in *quality* overall—but not higher in *feasibility*—with 9 strategies perceived differently by women and men, after imposing conservative statistical controls. There was broad general agreement on the *relative-quality* rankings of the 44 strategies. Women (vs. men) gave higher quality ratings to increasing the value of teaching, service, and administrative experience in tenure/promotion decisions, increasing flexibility of federal-grant funding to accommodate mothers, conducting gender-equity research, and supporting shared tenure lines enabling work-life balance. Women (vs. men) believed it was more feasible for men to stop the tenure clock for 1 year for childrearing and for universities to support requests for shared tenure lines, but less feasible for women to chair search committees. Our national survey thus supported the belief that placing women into administration creates greater endorsement of strategies to attract and retain women in STEM, although the effectiveness of these strategies was outside the

scope of this research. Topics of disagreement between women and men are potentially important focuses of future policy, because female administrators may have insights into how to retain women that male administrators do not share.

Keywords: underrepresentation of women, women in science, administrator gender, retention strategies, work-life balance, gender bias

INTRODUCTION

Much has been written about the status of women in academic science (e.g., Ginther and Kahn, 2009, 2015; Ceci and Williams, 2010a,b; Williams and Ceci, 2012, 2015; Ceci et al., 2017). To be sure, women have made substantial progress in several STEM fields over the past two decades (e.g., Xie and Shauman, 2003; Hill et al., 2010). For example, female assistant professors are now at or above parity in psychological science and in most social sciences, and they are approaching parity in biological sciences (Ceci et al., 2014). However, women remain less numerous at senior ranks in all fields, and in the mathematically-intensive fields—physics, chemistry, computer science, engineering, economics, and geosciences—women occupy fewer than 20% of combined tenured and tenure-track professorships, as can be seen in **Figure 1**. Women's underrepresentation in academic science has led to the publication of articles, chapters, and books focusing on women's specific, practical, day-to-day needs in their colleges and universities, in the hope of addressing these needs through specific policies and strategies designed to better accommodate women and families (e.g., Williams and Ceci, 2012; Williams et al., 2013, 2015; Ceci and Williams, 2015; Jones et al., 2016).

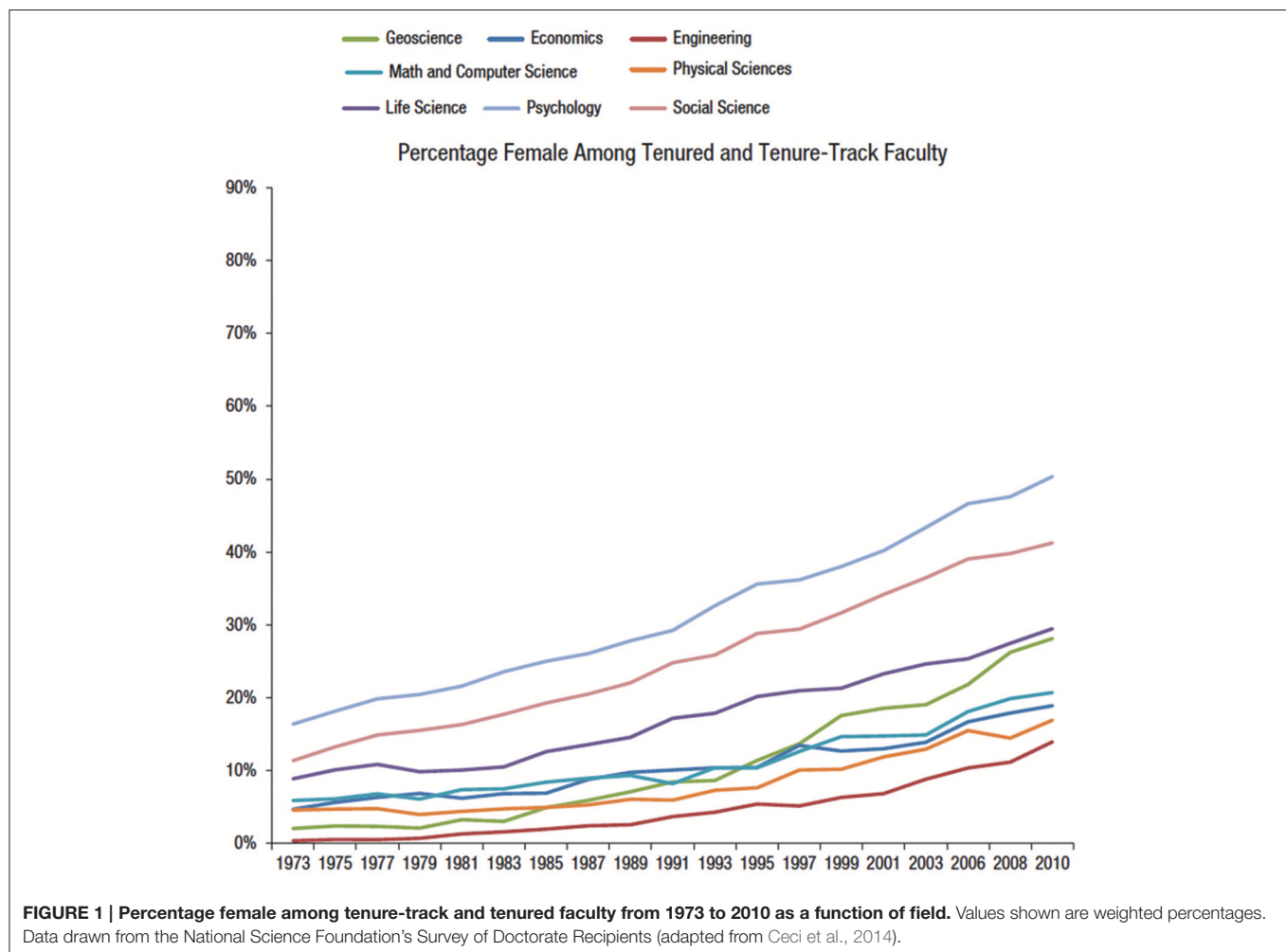
Once hired, women face formidable challenges in academic science, which underscore the need for ongoing strategies and policies to address women's daily needs as professors. One key issue concerns research productivity and how the academic work environment may hinder women's success (Raj et al., 2016). Women professors publish fewer articles, chapters, and books than their male counterparts, a situation that may have implications for sex differences in hiring, salary, and promotion. Numerous researchers have documented productivity differences, using a variety of measures. Women publish less than men, starting in graduate school, and extending through the postdoctoral and pre-tenure years (Ceci et al., 2009).

Assistant professors represent the future of the academy; thus, it is interesting to examine trends in male and female assistant professors' productivity over the past 20 years. Elsewhere we have shown that in many fields, assistant professors of both sexes are publishing more articles in 2008 than in 1995, with some notable exceptions (Ceci et al., 2014). The average difference in publications by gender for assistant professors is 2.1 articles more for men than for women, which is equivalent to 27% of the total male assistant professor publications over the 5-year period. **Figure 2** shows these differences by field. As can be seen, there is no clear-cut temporal trend; in some disciplines women's productivity increased between 1995 and 2008 while in others it declined, vis-a-vis men's productivity. On net, however, women published less in both periods. In each field in both

1995 and 2008, point estimates indicate that the average man published more than the average woman. The largest, statistically significant productivity gaps for assistant professors in 1995 were in engineering, life sciences, math/computer science, and physical science. By 2008, however, the fields of engineering and math/computer science saw these gaps close to the point at which they were no longer statistically significant. In life science by 2008, the gap narrowed but remained statistically significant, whereas in physical science the gap actually grew larger (for details please see pp. 103–107 of Ceci et al., 2014).

Data such as these have motivated administrators and gender-equity advocates to lobby for policies to aid women in the aftermath of childbirth or adoption, such as paid leaves, supplemental funding on grants to hire postdocs to run labs, and paid conference travel for childcare workers. However, it is not clear that the publishing gaps are causally related to family demands, because they exist among single, childless men and women as well (Williams and Ceci, 2012). Although the gap also appears among assistant professors at R1 institutions, with similar teaching responsibilities, it seems largely the result of sex differences in institutional resources, with women disproportionately more likely to work at small teaching-intensive institutions and men at research-intensive ones with greater resources for research (Ceci and Williams, 2011).

Unsurprisingly, women scientists in the academy are more likely to express dissatisfaction with aspects of their work that may be indirectly related to their underrepresentation and lower productivity. There are reports of an unwelcoming, “chilly” climate, indifferent attitudes toward family-work balance, and harassment, all of which may undermine women's success and persistence in the professoriate. Specifically, surveys of faculty indicate that the vast majority of women in science continue to describe an unwelcoming climate, including outright harassment (e.g., Ecklund and Lincoln, 2011). Coinciding with growth in their numbers, women scientists have reported being subjected to various barriers and challenges. Williams et al. (2014) reported the results of a survey in which they recruited 557 women scientists through the Association for Women in Science. Virtually all of the women claimed to have been victims of at least one of five biases they were asked about (e.g., sexual harassment; backlash for exhibiting stereotypically masculine behaviors such as assertiveness or expressing anger). Sixty-four percent of respondents with children reported a stigma when women took parental leave or stopped the tenure clock, leading the authors to conclude: “Motherhood appears to be a no-win proposition for many women in STEM” (Williams et al., 2014, p. 5). (Interestingly, motherhood worked both ways, with women without children also reporting dissatisfaction over being



expected to work longer hours to make up for the schedules of colleagues who do have children.).

Against this broad backdrop of increasing numbers of women in the STEM professoriate, but persistent problems with productivity and allegations of workplace issues that undermine success, we wondered whether the gender of administrators makes a difference in the climate women face in STEM academic science. Are units, departments, colleges, and even universities headed by women likely to endorse more or different interventions and policies to combat the leakage of women STEM faculty? Note that this framing of the question differs from the more common framing, which asks whether women in units led by women are more satisfied. This is because satisfaction can be due to the mere presence of a same-sexed administrator and have nothing to do with any specific policies or procedures she or he advocated. We were interested in knowing whether female administrators endorsed a different constellation of strategies to attract and retain women faculty than were endorsed by their male counterparts. In a search of the literature we found nothing to directly answer this question, so we did the study ourselves.

Could it be that the lower number of women in some fields is associated with less aggressive leadership related to the

recruitment and retention of women? Based on NSF's most recent Survey of Doctorate Recipients (SDR), there are small but statistically-significant sex differences when all types of institutions are combined: Women are less likely to be deans, directors, or department chairs (12.1 vs. 15.1%; $p < 0.01$); however, they are equally likely to be presidents, provosts, and chancellors (1.2 vs. 1.2%). Thus, the question suggests itself: Do departments, colleges, and universities that are headed by women endorse female-friendly practices that male administrators are less likely to endorse?

Some qualitative data suggest that female administrators provide a sense of social capital in the workplace for women that male administrators may not (Smith, 2014). For example, based on interviews, Dunn et al. (2014) reported widely-varying administrative styles of men and women administrators. Intensive interviews they conducted with 19 women in STEM fields at five universities revealed a sense of isolation related to a relative lack of social capital (e.g., connections, tacit knowledge, membership in networks, and possession of material resources). Successful women administrators' style of leadership (building social capital and combining both agentic assertiveness and communal warmth) may be better at communicating and

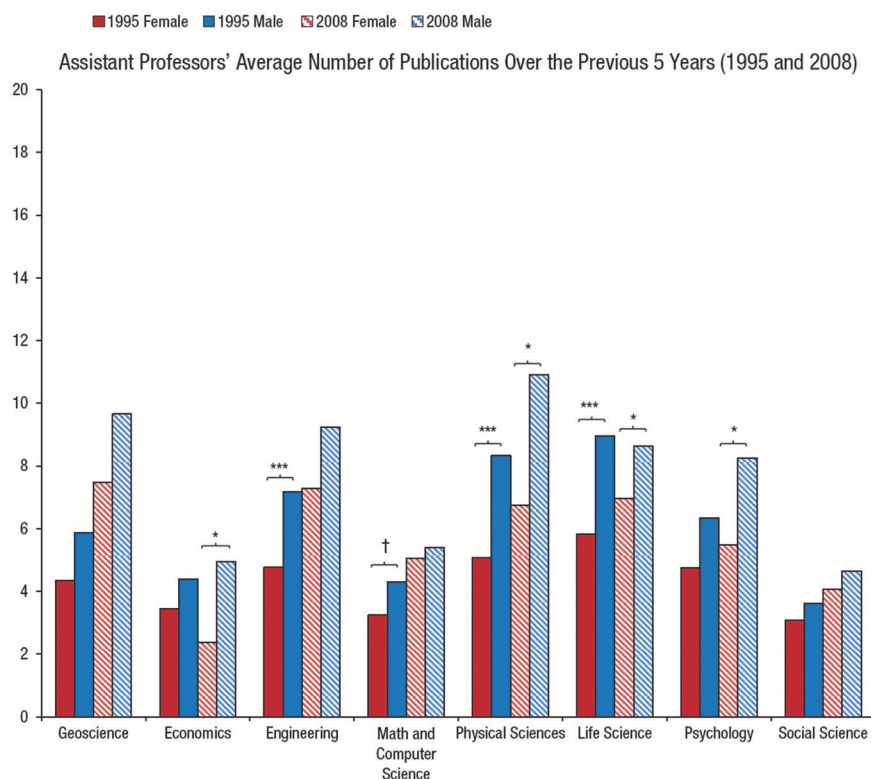


FIGURE 2 | Women vs. Men Assistant Professors: average number of publications over prior 5 years (adapted from Ceci et al., 2014). * $p < 0.05$, $^{\dagger}p < 0.01$, *** $p < 0.001$.

breaking down such feelings of isolation (Eagly and Carli, 2007). Other evidence suggests that women and minorities respond best in more collaborative learning experiences, which is a distinctly female leadership style (see Gorman et al., 2010). Finally, Hough (2010) profiled the leadership styles of 183 female administrators at senior level positions, such as president, chancellor, vice president, and dean, at accredited institutions of higher education. She reported that effective administrators strive to increase a sense of community and collegiality.

In sum, social capital theory explains why having female administrators might work positively to attract and retain women in STEM; however, no direct data exist regarding whether female and male administrators actually endorse different strategies to attract and retain women in STEM. In the survey that follows we asked a national sample of administrators to rate various female-friendly strategies that have been proposed in the literature. Do women and men differ in their support of these interventions? And, what can we learn regarding strategies that were supported by both genders, as opposed to strategies endorsed more by one gender than the other?

METHOD

We began by compiling a list of potential strategies for attracting and retaining women in academic science. We gathered these strategies from articles in the PsycINFO

database and from Google/Google Scholar searches, found via search terms comprised of various combinations of the words “women,” “science,” “STEM,” “underrepresentation,” “gender,” “professor,” “academic,” “hiring,” “tenure,” “retention,” “strategies,” “policies,” “procedures,” “family friendly,” and “university administrator,” (for example, “women in science,” “women in STEM,” “STEM retention,” “academic retention strategies,” and so on). We sifted through 206 articles (by which point we were encountering substantial redundancies in strategies mentioned and/or advocated). We also followed additional leads found in these articles’ reference sections to point us to mentions of further potential strategies. Our overarching goal was to compile a lengthy, representative list of recommendations for increasing the presence and persistence of women in academic science.

Based on this corpus of research, we next whittled the list of policies and recommendations to remove redundancies. For example, numerous researchers recommended establishing committees to monitor women’s progress (i.e., conducting institutional research on gender-equity issues; see, e.g., Committee on Maximizing the Potential of Women in Academic Science and Engineering, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, 2007; National Research Council, 2009), and stopping the tenure clock for family formation (e.g., Goulden et al., 2009; Williams and Ceci, 2012). The resulting list of policies and recommendations was then sent to 24 natural and social science faculty across

ranks, who were asked to comment on remaining redundancies and add any potential missing strategies. Our goal was to include the most important and often-mentioned strategies in a comprehensive list, which we then used as the basis for national empirical data collection. Based on the feedback from professors across ranks in six science and social science disciplines, the list of strategies was iteratively revised until we developed a final survey containing 44 strategies.

The final survey (see **Table 1**) was emailed to 1,529 administrators at 96 public and private research universities across the United States (see **Table 2**). The target population consisted of provosts, deans, associate deans, and department chairs of STEM fields at American Carnegie I research-oriented universities, formerly called R1s. These United States university administrators were asked for two responses to each policy—a rating of its *quality* and a rating of its *feasibility*. Ratings were based on a 9-point Likert scale, with 1 being the lowest score and 9 being the highest. Two-hundred-thirty of the individuals in our database had either left administration, retired permanently, gone on leave, changed universities or had otherwise been separated from their former positions. Our survey received 474 responses (36.5% response rate), of which 334 contained complete numerical data used in the analysis. The other 130 replies contained incomplete data, or responses that consisted of comments about the importance of retaining women, personal anecdotes, and so on, as opposed to complete sets of ratings (Note that we are not asserting that the sample was perfectly representative of the population of U.S. college and university administrators, only that the 334 administrators represented all 96 R1s). For each respondent, publicly-available data was gathered on her or his gender, title, and university type. Data were then de-identified to ensure anonymity of responses. Of the 334 respondents, 246 were men and 88 were women; there were 157 men and 34 women STEM department chairs, 38 men and 22 women associate deans, 42 men and 24 women deans, and nine men and eight women provosts, all from Carnegie I research universities.

RESULTS AND DISCUSSION

The analyses ranked strategies for their quality and feasibility, and examined whether administrator gender affected ratings of the policies. We also evaluated the impact of university type (public or private) and geographical location of institution on ratings of policies; location was not systematically related to strategy ratings, and results for university type, public vs. private, appear in Appendix (**Figures A1, A2**). For each set of comparisons (e.g., comparing all strategies across gender) we adjusted *p*-values using the conservative Benjamini and Hochberg (1995) false-discovery rate at a 5% level.

Overall Effect of Strategy Quality

A repeated-measures ANOVA was conducted on the quality ratings of the 44 strategies by the 334 respondents, to evaluate whether the mean ratings of items differed significantly across items. The result was highly significant— $F_{(16,02,4236,45)} = 81.65$, $p < 0.001$, with Greenhouse-Geiser correction for violations of sphericity. This finding showed that respondents of both

genders perceived strategies as varying systematically in quality, with broad general agreement concerning the strategies' relative rankings. Strategies discussed in the next section ("General Agreement about Relative Quality of Strategies") are ones about which administrators of both genders agreed. Strategies characterized by sex differences in opinions are discussed in the subsequent section ("Gender Differences in Ratings of Strategy Quality or Feasibility"). Ratings of strategy quality correlated 0.98 with strategy feasibility, so we focus on quality ratings in this discussion of results, except for those few occasions when ratings of quality and feasibility differed, such as in the situation discussed below under "Gender Differences in Ratings of Strategy Quality or Feasibility."

General Agreement about Relative Quality of Strategies

The 44 strategies ranked by quality ratings are shown in **Table 3**, with the mean rating for each policy on a 1-to-9 scale (1 = extremely low, 3 = somewhat low, 5 = neutral, 7 = somewhat high, and 9 = extremely high). Twelve strategies had high mean quality ratings of 7.0 or more; we describe these strategies here, proceeding with a brief description of the balance of the strategies in order from highest to lowest quality.

The highest-quality strategy was for universities to provide on-campus childcare centers ($M = 8.36$), which unsurprisingly is a priority for universities across the U.S., reflecting the challenges faced by women (and men) faculty with preschool-aged children. Offering equal opportunities for women and men to lead committees and research groups ($M = 8.26$) was also seen as an extremely high-quality strategy, as was developing mentoring programs to reduce isolation of female faculty ($M = 7.92$).

A policy that has become widely used over the past decade, stopping the tenure clock for raising children for up to 1 year per child ($M = 7.59$), was the next-highest-rated strategy. Providing fully-paid leave for giving birth for tenure-track women only, for a total of one semester, was seen as valuable ($M = 7.52$), as was allowing unpaid sabbaticals and leaves of absence for both genders without penalty, for family-related reasons such as elder caretaking and issues with children ($M = 7.50$).

In recognition of the role played by departmental-level administrators, respondents endorsed training for department chairs on helping faculty manage work-life issues ($M = 7.40$). Respondents also supported the deferred start of fellowships to allow for caregiving ($M = 7.20$), and providing of teaching relief for new tenure-track parents for one semester ($M = 7.17$). Another strategy related to caregiving was also endorsed: Supporting no-cost extensions for caregiving on grants and fellowships ($M = 7.12$). Institutions' need to explore and endorse couples-hiring to help resolve the two-body problem was also rated highly ($M = 7.05$). And finally, there was broad support for providing fully-paid leave for adoption and new parenthood, for tenure-track women and men, for one semester ($M = 7.03$).

Concomitant with the endorsement of providing on-campus childcare centers (discussed above), support was found for the importance of providing subsidies for childcare ($M = 6.84$), and family housing subsidies ($M = 6.77$). Both genders also believed

TABLE 1 | STEM-administrator survey.

Please rate each of the following policy ideas on a 1-to-9 scale for QUALITY and FEASIBILITY, in which 1 = extremely low, 3 = somewhat low, 5 = neutral, 7 = somewhat high, and 9 = extremely high. By QUALITY ("Q") we mean: How good is this strategy, if the goal is to increase the number of women in traditionally-underrepresented STEM fields in the professoriate? By FEASIBILITY ("F") we mean: How workable, cost-effective, and reasonable would this strategy be to implement?

Addressing Gender Biases During Hiring

Have a woman chair search committees whenever possible. Q__F__
 Reward departments that hire women. Q__F__
 Set gender goals for candidate pools. Q__F__
 Set quotas for new lines: women-only lines until critical mass reached. Q__F__
 Explore/endorse couples-hiring. Q__F__
 Guarantee academic employment for professional spouses/partners. Q__F__
 Instruct search committees to ignore family-related gaps in CVs. Q__F__

Addressing Gender Biases After Hiring

Set gender quotas (minimum thresholds) for promotion to higher levels of rank (e.g., full professor). Q__F__
 Set gender quotas for important committees and administrative posts. Q__F__
 For promotion, increase value of teaching and service plus administration. Q__F__
 Conduct (and disseminate) institutional research on gender equity. Q__F__

Attaining Tenure and Maintaining Productivity

Provide fully-paid leave for giving birth (tenure-track women only): For 6 weeks? Q__F__ For 1 semester? Q__F__ For 1 year? Q__F__
 Provide fully-paid leave for adoption/new parenthood (tenure-track women and men): For 6 weeks? Q__F__ For 1 semester? Q__F__ For 1 year? Q__F__
 Provide teaching relief for new tenure-track parents: 1 semester? Q__F__ 1 year? Q__F__
 Stop the tenure clock for raising children for up to 1 year per child: For mothers? Q__F__ For fathers? Q__F__
 Change timing of tenure assessment to not coincide with peak fertility and childrearing demands. Q__F__
 Allow option of changing from full-time to part-time tenure-track: Short Term (up to 1 year) Q__F__ Medium Term (2–5 years) Q__F__ Permanent Q__F__
 Support requests for shared tenure lines (between partners). Q__F__

Balancing Work and Family

Provide on-campus childcare centers. Q__F__
 Provide subsidies for on-campus or off-campus childcare services. Q__F__
 Allow unpaid sabbaticals and leave of absences for both genders without penalty, for family-related reasons such as elder caretaking and issues with children. Q__F__
 Offer family housing subsidies in regions where young families are priced out of the market. Q__F__
 Use technology to allow women and men with children to work and attend meetings from home. Q__F__
 Provide an academic role for women who have left professional positions to have children. Q__F__

Providing Leadership and Training Opportunities

Provide equal opportunities for women and men to lead committees and research groups. Q__F__
 Train department chairs on helping faculty manage work-life issues. Q__F__
 Develop mentoring programs to reduce isolation of female faculty. Q__F__
 Convene gender-equity workshops focusing on issues such as workplace climate and resource allocation. Q__F__

Supporting Greater Flexibility for Federal Grants and Funding

Support no-cost extensions for caregiving on grants and fellowships. Q__F__
 Support part-time fellowships and grants. Q__F__
 Support the deferred start of fellowships to allow for caregiving. Q__F__
 Endorse supplements to offset PI's productivity loss due to family-related absences. Q__F__
 Support conference and meeting grant supplements to cover cost of PI's dependent care travel (children's and childcare workers' expenses allowable). Q__F__
 Support grants for retooling after maternity leave. Q__F__
 Provide support to help faculty engaging in caregiving duties to catch up mid-career. Q__F__
 Endorse supplemental funding for hiring postdocs to maintain momentum during family leaves. Q__F__

TABLE 2 | Universities in sample.

Arizona State University, Brandeis University, Brown University, California Institute of Technology, Carnegie Mellon University, Case Western Reserve University, Colorado State University, Columbia University, Cornell University, Dartmouth University, Duke University, Emory University, Florida State University, Georgetown University, Georgia Institute of Technology, Harvard University, Indiana State University, Indiana University, Iowa State University, Johns Hopkins University, Kansas State University, Louisiana State University, Michigan State University, Massachusetts Institute of Technology, Montana State University, North Carolina State University, Northwestern University, New York University, Ohio State University, Oregon State University, Penn State University, Princeton University, Purdue University, Rensselaer Polytechnic Institute, Rice University, Rutgers University, Stanford University, SUNY Albany, SUNY Buffalo, SUNY Stony Brook, Texas A and M University, Tulane University, UC Berkeley, UC Davis, UC Denver, UC Irvine, UC Riverside, UC San Diego, UC Santa Barbara, UC Santa Cruz, UC Los Angeles, University of Alabama at Birmingham, University of Arizona, University of Cincinnati, University of Colorado at Boulder, University of Connecticut, University of Delaware, University of Florida, University of Georgia, University of Hawaii, University of Illinois at Chicago, University of Illinois at Urbana-Champaign, University of Iowa, University of Kansas, University of Kentucky, University of Maryland, University of Massachusetts—Amherst, University of Miami, University of Michigan, University of Minnesota, University of Missouri, University of Nebraska-Lincoln, University of New Mexico, University of North Carolina, University of Notre Dame, University of Pennsylvania, University of Pittsburgh, University of Rochester, University of South Carolina, University of South Florida, University of Southern California, University of Tennessee, University of Texas, University of Utah, University of Virginia, University of Washington, University of Wisconsin-Madison, Vanderbilt University, Virginia Polytechnic Institute, Washington State University, Washington University in St Louis, Wayne State University, Yale University, Yeshiva University.

that gender equity workshops were valuable ($M = 6.79$). On the topic of leaves for faculty becoming parents, respondents supported fully paid leave for adoption for women and men for 6 weeks ($M = 6.72$), as well as fully paid leave for giving birth for women, also for 6 weeks ($M = 6.72$). Also accommodating parents and those with travel and caretaking demands, we found support for the importance of allowing remote meeting attendance ($M = 6.61$). With children often comes a challenge to maintaining productivity, and we found support for the practice of ignoring family-related gaps in CVs ($M = 6.50$); that is, respondents agreed that someone with a total of 5 years on tenure track, one of which was spent on leave due to childcare, should be considered just 4 years on tenure track for purposes of setting the tenure clock. We also noted an endorsement of the policy of temporarily stopping the tenure clock for fathers ($M = 6.32$).

Note, again, that a rating of five signified neutral quality and a seven signified somewhat high quality. Having women chair searches was seen as a generally good quality strategy ($M = 6.24$), as was the awarding of part-time fellowships and grants to accommodate parents and academics with caregiving responsibilities ($M = 6.14$). Similarly, midcareer grants to faculty caregivers was rated above 6 ($M = 6.09$) as was the policy of allowing tenured faculty to go part-time for 1 year ($M = 6.06$). Rated just above 6 was the strategy of rewarding departments for hiring women ($M = 6.01$).

Between a neutral-quality rating of 5 and a slightly-high rating of 6, we found modest support for grants for retooling after maternity leave ($M = 5.90$), and funding fully paid leave for giving birth for women for 1 entire year ($M = 5.67$). Encouraging universities and colleges to hire faculty partners and spouses for non-professorial positions was also somewhat weakly endorsed ($M = 5.61$). Similarly, giving teaching relief to new parents for 1 year ($M = 5.40$) was weakly supported. Support was generally neutral for the practice of setting gender hiring goals ($M = 5.19$) and for offering fully paid leave for adoption to women and men faculty for 1 year ($M = 5.07$). Providing an academic role for mothers who used to be professors or who wish to participate in university life also was seen as a neutral strategy ($M = 5.03$).

Turning to a consideration of strategies deemed to be of relatively lower quality by the 334 respondents, six strategies

had mean quality ratings below 5.0 (and no gender interactions affecting the interpretation of the results). We describe these strategies below, ordered from relatively better to relatively worse. Interestingly, one frequently-mentioned strategy widely acknowledged in the literature as being especially beneficial for women was seen by our respondents as being of relatively low quality: Allowing the option of changing from full-time to part-time tenure-track, over the medium term of 2–5 years ($M = 4.90$). It has been argued that women may wish to work part-time for a few years, during which they can have and raise children, then later segue back to full-time work when their children begin school. Yet our data call into question the wisdom of this life plan, at least from the administrators' point of view. Setting gender quotas for important committees and administrative posts ($M = 4.17$), and allowing the option of changing from full-time to part-time tenure-track on a permanent basis ($M = 3.97$), were also relatively weakly endorsed.

In another surprise, administrators did not support changing the timing of tenure assessment to avoid peak fertility and childrearing demands ($M = 3.83$). This strategy has been broadly advocated as an essential way to reduce pressure on women scholars, who are expected to amass a tenurable portfolio during the exact same years as when they tend to have children—in their thirties. The fact that women in science experience the confluence of the tenure clock and the biological clock, but that men in science simply do not share these limitations, is an inescapable aspect of the dilemma faced by female scholars. What, exactly, the academy can do to ameliorate this problem for women scholars remains a pressing question. Other weakly-endorsed strategies included setting gender quotas for new tenure lines and calling for women-only lines until a critical mass of women is reached ($M = 3.62$). Finally, the worst strategy of all was seen as setting gender quotas (minimum thresholds) for promotion to higher levels of rank (e.g., full professor $M = 2.46$).

Gender Differences in Ratings of Strategy Quality or Feasibility

The most striking finding in this research was that, overall, women administrators were significantly more likely than men to rate all strategies higher in quality, on average [female mean

TABLE 3 | Strategies for increasing/retaining women in STEM professoriate, listed from highest to lowest quality ($n = 334$ faculty respondents; 44 strategies rated on 1-to-9 scale in which 1 = extremely low, 3 = somewhat low, 5 = neutral, 7 = somewhat high, and 9 = extremely high).

1	Provide on-campus childcare centers. (Q27, $M = 8.36$)
2	Provide equal opportunity for women and men to lead committees and research groups. (Q33, $M = 8.26$)
3	Develop mentoring programs to reduce isolation of female faculty. (Q35, $M = 7.92$)
4	Stop tenure clock for raising children for up to 1 year per child. (Q20, $M = 7.59$)
5	Provide fully-paid leave for giving birth (tenure-track women only): For 1 semester. (Q13, $M = 7.52$)
6	Allow unpaid sabbaticals and leave of absences for both genders without penalty, for family-related reasons such as elder caretaking and issues with children. (Q29, $M = 7.50$)
7	Train department chairs on helping faculty manage work-life issues. (Q34, $M = 7.40$)
8	Support the deferred start of fellowships to allow for caregiving. (Q39, $M = 7.20$)
9	Provide teaching relief for new tenure-track parents: 1 semester. (Q18, $M = 7.17$)
10	Support no-cost extensions for caregiving on grants and fellowships. (Q37, $M = 7.12$)
11	Explore/endorse couples-hiring. (Q5, $M = 7.05$)
12	Provide fully-paid leave for adoption/new parenthood (tenure-track women and men): For 1 semester. (Q15, $M = 7.03$)
13	Provide subsidies for on-campus or off-campus childcare services. (Q28, $M = 6.84$)
14	Convene gender-equity workshops focusing on issues such as workplace climate and resource allocation. (Q36, $M = 6.79$)
15	Offer family housing subsidies in regions where young families are priced out of the market. (Q30, $M = 6.77$)
16	Provide fully-paid leave for adoption/new parenthood (tenure-track women and men): For 6 weeks. (Q15, $M = 6.72$)
17	Provide fully-paid leave for giving birth (tenure-track women only): For 6 weeks. (Q12, $M = 6.72$)
18	Conduct (and disseminate) institutional research on gender equity. (Q11, $M = 6.68$)
19	Use technology to allow women and men with children to work and attend meetings from home. (Q31, $M = 6.61$)
20	Instruct search committees to ignore family-related gaps in CVs. (Q7, $M = 6.50$)
21	Stop the tenure clock for raising children for up to 1 year per child: For fathers. (Q21, $M = 6.32$)
22	Have a woman chair search committees whenever possible. (Q1, $M = 6.24$)
23	Endorse supplemental funding for hiring postdocs to maintain momentum during family leaves. (Q44, $M = 6.15$)
24	Support part-time fellowships and grants. (Q38, $M = 6.14$)
25	Provide support to help faculty engaging in caregiving duties to catch up mid-career. (Q43, $M = 6.09$)
26	Allow option of changing from full-time to part-time tenure-track: Short Term (up to 1 year). (Q23, $M = 6.06$)
27	Reward departments that hire women. (Q2, $M = 6.01$)
28	Support grants for retooling after maternity leave. (Q42, $M = 5.90$)
29	Support conference and meeting grant supplements to cover cost of PI's dependent care travel (children's and childcare workers' expenses allowable). (Q41, $M = 5.78$)
30	Provide fully-paid leave for giving birth (tenure-track women only): For 1 year. (Q14, $M = 5.67$)
31	Guarantee academic employment for professional spouses/partners. (Q6, $M = 5.61$)
32	Provide teaching relief for new tenure-track parents: 1 year. (Q19, $M = 5.40$)
33	Set gender goals for candidate pools. (Q3, $M = 5.19$)
34	Provide fully-paid leave for adoption/new parenthood (tenure-track women and men): For 1 year. (Q17, $M = 5.07$)
35	Provide an academic role for women who have left professional positions to have children. (Q32, $M = 5.03$)
36	Endorse supplements to offset PI's productivity loss due to family-related absences. (Q40, $M = 4.95$)
37	Allow option of changing from full-time to part-time tenure-track: Medium Term (2–5 years). (Q24, $M = 4.90$)
38	Support requests for shared tenure lines (between partners). (Q26, $M = 4.60$)
39	For promotion, increase value of teaching and service plus administration. (Q10, $M = 4.33$)
40	Set gender quotas for important committees and administrative posts. (Q9, $M = 4.17$)
41	Allow option of changing from full-time to part-time tenure-track: Permanent. (Q25, $M = 3.97$)
42	Change timing of tenure assessment to not coincide with peak fertility and childrearing demands. (Q22, $M = 3.83$)
43	Set quotas for new lines: women-only lines until critical mass reached. (Q4, $M = 3.62$)
44	Set gender quotas (minimum thresholds) for promotion to higher levels of rank (e.g., full professor). (Q8, $M = 2.46$)

$= 6.33$, male mean $= 5.99$, $t_{(332)} = -2.58$, $p = 0.01$]. This finding suggests that women see the issues of attracting and retaining women in the STEM professoriate as more salient or important than men do, on average. (Figure 3, which shows men's vs. women's mean quality ratings, reveals that women's

mean ratings were generally higher than men's.) The feasibility ratings (see Figure 4) did not show this same trend—women *did not* rate the average feasibility of the strategies higher than men did [female mean $= 5.51$, male mean $= 5.55$, $t_{(332)} = 0.34$, $p = 0.74$].

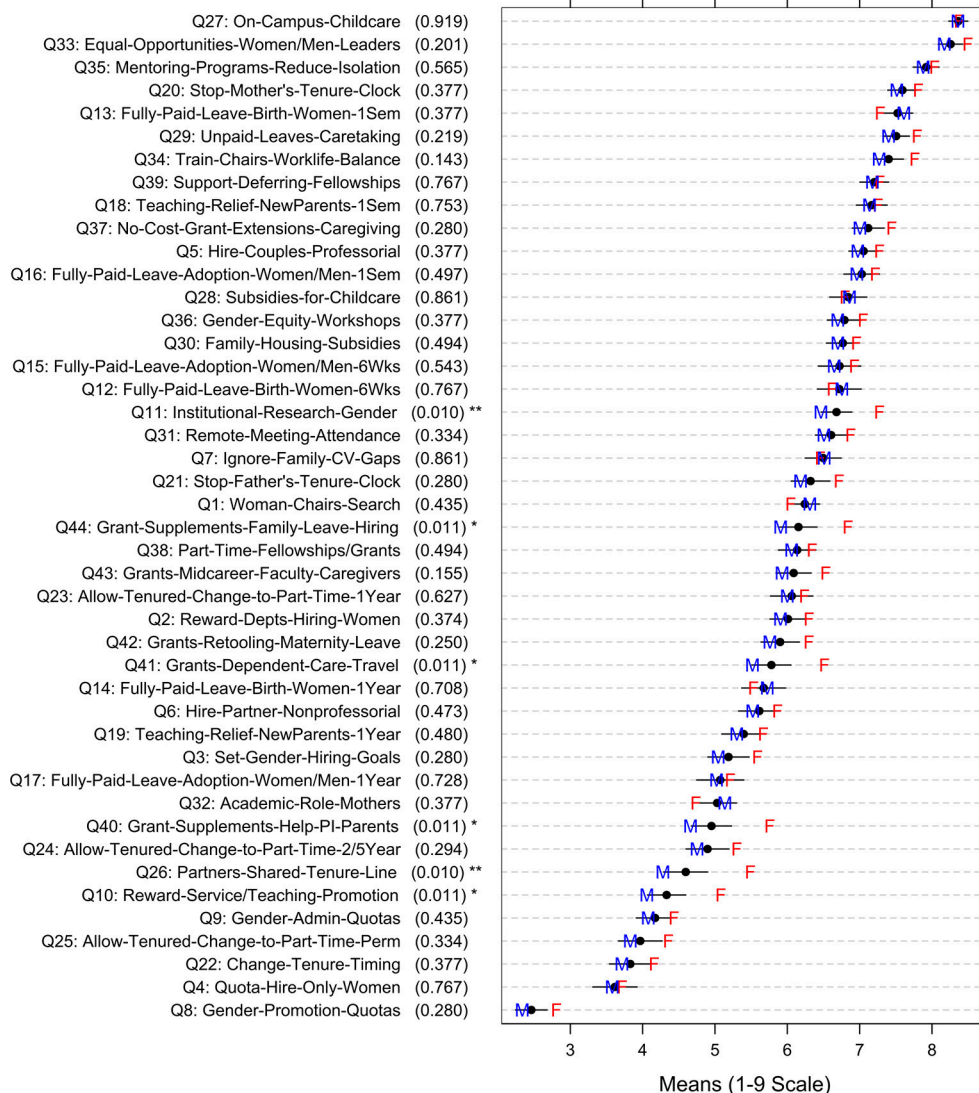


FIGURE 3 | Strategy Quality Ratings: Overall Means, Confidence Intervals, and Means by Gender (F = Female; M = Male; adjusted p -values in parentheses show significance level of comparison of item ratings by gender; * $p \leq 0.05$, ** $p \leq 0.01$).

To further explore the specific strategies most associated with higher ratings by women, the 44 strategies were analyzed individually to examine gender differences in ratings. Given that in this exploratory process we performed a large number of significance tests, we corrected the Type I error rate. Here, we use the Benjamini-Hochberg False Discovery Rate (FDR) correction, and set the FDR to 5%. In addition, we corrected all t -tests for potential violation of homogeneity of variance, and applied the Welch-adjustment to the degrees of freedom, thus making the p -values even more conservative. In what follows, we report actual t -values and degrees of freedom after Welch's adjustment for significance tests, but only report p -values after the FDR adjustment. After these conservative adjustments, gender differences in ratings of nine items remained significant. Six of these significant gender effects reflected gender

differences in ratings of strategy quality, and three reflected gender differences in ratings of strategy feasibility. We first discuss gender differences in quality ratings, turning next to a consideration of gender differences in feasibility ratings.

Gender Differences in Ratings of Strategy Quality

From the strategy category "Addressing Gender Biases After Hiring," we found that women more than men supported conducting (and disseminating) institutional research on gender equity [female mean = 7.27, male mean = 6.47, $t_{(197.18)} = 3.76$, $p = 0.010$]. Women's greater emphasis on the importance of gender-equity research is understandable, inasmuch as universities and colleges that conduct and then disseminate such information create an atmosphere in which women's issues

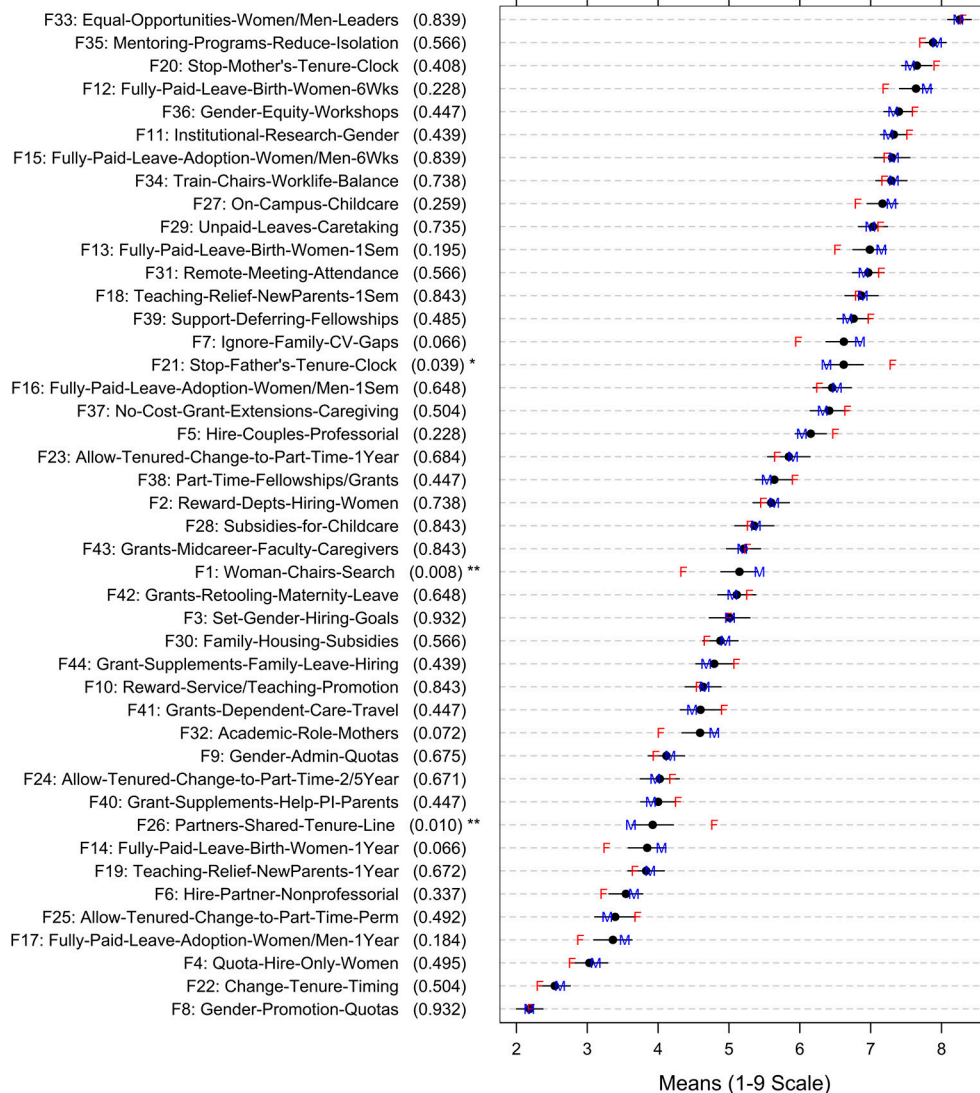


FIGURE 4 | Strategy Feasibility Ratings: Overall Means, Confidence Intervals, and Means by Gender (F = Female; M = Male; adjusted p -values in parentheses show significance level of comparison of item ratings by gender; * $p \leq 0.05$, ** $p \leq 0.01$).

are valued, studied, and (hopefully) meaningfully addressed. Obviously, knowing what the issues are is the critical first step in this process.

One cluster of gender differences concerned the role of federal-grant funding—specifically, federal policies, rules, and regulations as potential ways to address issues faced by researchers balancing family and work lives. Under the survey category, “Supporting Greater Flexibility for Federal Grants and Funding,” women rated as higher quality than did men the importance of endorsing supplemental funding for hiring postdocs to maintain momentum during family leaves [female mean = 6.84, male mean = 5.91, $t_{(169.26)} = 3.39$, $p = 0.011$]. Women also were more likely to rate as high quality the strategy of supporting conference and meeting grant supplements to cover cost of PI’s dependent care travel (with children’s and

childcare workers’ expenses allowable); female mean = 6.51, male mean = 5.52, $t_{(154.4)} = 3.27$, $p = 0.011$. In a similar vein, women also rated higher than did men the strategy of endorsing federal-grant supplements to offset Principal Investigators’ productivity losses due to family-related absences [female mean = 5.76, male mean = 4.66, $t_{(129.31)} = 3.27$, $p = 0.011$]. This group of federal-grant-related strategies reflects new and emerging thinking about how to redefine historical rules that impose highly-limiting restrictions, particularly upon women with children and caretaking responsibilities.

Another widely-cited strategy in the literature for accommodating families with childrearing needs is for universities to support requests for shared tenure lines. From the category, “Attaining Tenure and Maintaining Productivity,” women rated as higher quality than did men the strategy of

institutions supporting requests for shared tenure lines (between partners)—female mean = 5.49, male mean = 4.28, $t_{(150.18)} = 3.58$, $p = 0.01$. This finding may reflect greater endorsement by women of the need to truly *balance* family and work life, with implicit compromises affecting both the work and family portions of the balance.

We turn next to a key issue concerning women's retention in the professorial pipeline: Earning tenure, and more specifically, delineating the contents of a tenurable portfolio of work. Obviously, the precise nature of the types of work that are valued during tenure consideration is a critical aspect of the tenure decision, itself. The notion of expanding traditional definitions of what constitutes a tenurable portfolio—to accommodate and value women's styles of working within collaborative settings—also showed a sex difference in level of endorsement in our study. From the category, “Addressing Gender Biases After Hiring,” women administrators were more likely to support the concept of increasing the value of teaching and service plus administration when evaluating a candidate for promotion [female mean = 5.08, male mean = 4.06, $t_{(140.71)} = 3.59$, $p = 0.001$]. This gender difference reflected female administrators' greater desire (as compared to male administrators) to increase the value, during tenure and promotion evaluations, of those tasks undertaken and sometimes prioritized by women faculty.

Gender Differences in Ratings of Strategy Feasibility

Turning to a consideration of the *feasibility of strategies* (as opposed to simply their quality), three strategies emerged as being seen as more feasible by one gender than the other. From the “Attaining Tenure and Maintaining Productivity” category, female administrators saw it as more feasible than did male administrators for *male faculty* to stop the tenure clock for raising children for up to 1 year [female mean = 7.31, male mean = 6.38, $t_{(155.92)} = 3.05$, $p = 0.039$]. Once again, this gender difference reveals that women and men perceive differently men's ability and willingness to delay career advancement in order to prioritize the needs of young children and partners or spouses.

From the category “Addressing Gender Biases During Hiring,” women saw as less feasible the strategy of having a woman chair search committees whenever possible, while men saw this strategy as more feasible [female mean = 4.36, male mean = 5.43, $t_{(171.03)} = 3.83$, $p = 0.008$]. Female administrators' lower-feasibility ratings probably represented an acknowledgment of the sometimes-onerous service and administrative demands placed upon women faculty, particularly in departments in which women are underrepresented.

Echoing an earlier finding, from the category, “Attaining Tenure and Maintaining Productivity,” women administrators saw it as more feasible to support requests for shared tenure lines (between partners) than did male administrators [female mean = 4.79, male mean = 3.62, $t_{(145.91)} = 3.59$, $p = 0.010$]. Note that above we reported that female administrators also saw shared tenure lines as being a higher-quality strategy than did male administrators. Thus, repeatedly we found that female and male administrators held differing views on both the quality and the feasibility of partners sharing work and family duties: Administrators' gender predicted how they rated both

the effectiveness of the tenure-line-sharing approach, and the potential for actually accomplishing this approach in the real world.

GENERAL DISCUSSION

Women administrators in our sample believed that the 44 strategies for attracting and retaining women faculty were significantly higher in quality overall than was perceived by male administrators. Thus, our findings provide empirical support for the importance of women in administrative roles, since real-world resources are limited and women administrators deem women's recruitment and retention strategies to be generally high in quality and thus more worthwhile than men deem them to be. As can be seen in **Figures 3** and **4**, there was broad general agreement regarding the *relative quality and relative feasibility of strategies*, with administrators of both genders agreeing in general on the ranking of strategies by quality, or by feasibility. However, women believed the strategies were *higher in quality overall* than men did, although women did not see the strategies as being more *feasible* overall than men did. (In other words, women did not simply use the scale differently from men—women were not more positive in general in all of their ratings, compared with men).

At the level of individual strategies, women and men administrators rated the quality of certain strategies differently, with women rating the following policies as significantly higher in quality than men did: (a) various forms of flexibility with federal-grant funding designed to accommodate women with young children and keep these women in the game; (b) increasing the value of teaching, service, and administrative experience in the tenure/promotion evaluation process; (c) devoting university resources to conducting and disseminating gender-equity research; and (d) supporting requests from partners for shared tenure lines that enable couples to better balance work and personal/caretaking roles. Regarding feasibility of strategies, women administrators saw it as more feasible than men did for men to stop the tenure clock for 1 year due to childrearing demands, and for universities to support requests for shared tenure lines (between partners). But women administrators saw it as less feasible for women to chair search committees, presumably an acknowledgment of the potentially-onerous nature of service demands placed upon women.

What do these findings mean for the debate about how to attract and retain more women in academic science? Our national survey revealed that women administrators think differently from their male counterparts about certain key approaches to attracting and retaining women. Because women administrators value pro-women strategies more than men do overall, and value some individual strategies in particular more than men do, the resources lobbied for and allocated by these women administrators may be deployed more often toward the strategies they endorse, although we offer no specific evidence confirming this. Endorsement of a strategy in a survey may not necessarily translate into action. Likewise, the opinion of an administrator does not necessarily mean that the effectiveness (should the actual strategy be implemented in policy) is proven; the present data consisted of ratings by administrators. The current survey was not designed to address

whether an administrator actually implemented these strategies or how successful they were.

Women administrators' views that the strategies are higher in quality overall than men perceive them to be could result in women administrators spending relatively more of their limited budgets than men would on women-in-science issues. It has been argued that women and people of color in academic administrative posts bring different perspectives to their jobs (Smith et al., 2004), and our data support this position, at least with regard to beliefs about the quality and feasibility of strategies for attracting and retaining women in the STEM professoriate.

It is worth noting, however, that men and women agreed much of the time about the relative ranking of strategies—in other words, both genders basically agreed on what constituted the best vs. worst strategies among the 44 presented for evaluation. Women in general endorsed the strategies as being higher in quality overall than did men, and women and men disagreed about the quality of some strategies, but there was general agreement about the overall quality of one strategy relative to all the others, as seen by the similar rank ordering of strategies by both sexes. Administrators basically agreed on what represented higher- vs. lower-quality strategies. This is heartening news, since agreement about what constitutes a good strategy generally makes it simpler to get the strategy actually introduced as a policy. But women administrators were more supportive of strategies to attract and retain women in STEM, overall—and furthermore, there were some specific strategies that women endorsed at a significantly higher level than their male counterparts did.

Our data suggest that there are a substantial number of highly-rated strategies, and call into question the potential endorsement of the low-rated strategies. The highest-quality strategy was to provide on-campus childcare centers (rated $M = 8.36$ out of 9). Providing equal opportunity for women and men to lead committees and research groups was next ($M = 8.26$), followed by developing mentoring programs to reduce isolation of female faculty ($M = 7.92$) and stopping the tenure clock for raising children for up to 1 year per child ($M = 7.59$). There was broad support for providing fully-paid leave for giving birth for tenure-track women only for one semester ($M = 7.52$) and for allowing unpaid sabbaticals and leave of absences for both genders without penalty, for family-related reasons such as elder caretaking and issues with children ($M = 7.50$). Training department chairs on helping faculty manage work-life issues ($M = 7.40$) was seen as a high-quality strategy, as was supporting the deferred start of fellowships to allow for caregiving ($M = 7.20$), and providing teaching relief for new tenure-track parents for one semester ($M = 7.17$). Additional high-quality strategies involved supporting no-cost extensions for caregiving on grants and fellowships ($M = 7.12$), and exploring and endorsing couples-hiring ($M = 7.05$). Providing fully-paid leave for adoption/new parenthood (for tenure-track women and men), for one semester, was also seen as valuable ($M = 7.03$).

The two lowest-rated strategies involved use of gender quotas for hiring ($M = 3.62$) and promotion ($M = 2.46$). Interestingly, one strategy that has been widely recommended for its potential

to alleviate the conflict between women's biological clock and the tenure clock—changing the timing of tenure assessment not to coincide with peak fertility and childrearing demands—was also rated as a relatively poor idea ($M = 3.83$), and this was true for administrators of both genders. Similarly, allowing professors to change from full-time to part-time, permanently, on the tenure track—another strategy often acknowledged as being potentially beneficial to women—was also rated as low in quality ($M = 3.97$). It is fascinating that so many strategies widely written about and discussed as being potentially helpful were nevertheless viewed by active administrators as being low in quality.

Overall, the take-home message of this national empirical study was that (a) female administrators perceive strategies to retain women STEM professors as being higher in quality overall—i.e., more important and worthy—than do male administrators, (b) women vs. men administrators perceive some strategies differently; i.e., women and men disagree about the quality of certain strategies, and (c) women and men administrators agree in general regarding which strategies are higher vs. lower in quality. Thus, the belief that women in administrative roles will place greater emphasis than men will on strategies to retain women STEM professors was supported. A hopeful result was that men and women agree in general about better vs. worse approaches—thus suggesting that committees comprised of people of both genders will be able to find common ground for selecting and funding potential strategies. However, there were important exceptions; for example, women's greater endorsement of the need to more heavily weigh teaching, service, and administration in tenure-decision-making, and women's greater support of shared tenure lines (between partners) to enable broader sharing of childrearing and work activities and goals within a family. Areas of disagreement regarding strategy quality are important focuses of future policy and planning, because female administrators may have insights into how to retain women professors, by virtue of their personal experiences, that male administrators do not share.

ETHICS STATEMENTS

Cornell University IRB approved this project. Fully informed consent was given.

AUTHOR CONTRIBUTIONS

WW designed the study, supervised data collection and analysis, interpreted findings, and wrote the manuscript. AM collected data and did preliminary analyses and write-up. SB assisted with survey content/development. BC assisted with follow-up data collection. FT and FV analyzed and interpreted data, and FT created the figures. SC helped design the study, interpret findings, and write the manuscript.

FUNDING

This research was supported by a grant from the NIH Grant 1R01NS069792-01.

REFERENCES

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300. doi: 10.2307/2346101
- Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236
- Ceci, S. J., and Williams, W. M. (2010a). *The Mathematics of Sex: How Biology and Society Conspire to Limit Talented Women and Girls*. New York, NY: Oxford University Press.
- Ceci, S. J., and Williams, W. M. (2010b). Sex differences in math-intensive fields. *Curr. Direct. Psychol. Sci.* 19, 275–279. doi: 10.1177/0963721410383241
- Ceci, S. J., and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3157–3162. doi: 10.1073/pnas.1014871108
- Ceci, S. J., and Williams, W. M. (2015). Women have substantial advantage in STEM faculty hiring, except when competing against more accomplished men. *Front. Psychol.* 6:1532. doi: 10.3389/fpsyg.2015.01532
- Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412
- Ceci, S. J., Williams, W. M., and Kahn, S. (eds.). (2017). Underrepresentation of women in science: international and cross-disciplinary evidence and debate. *Front. Psychol. Special Issue*. <http://journal.frontiersin.org/researchtopic/2794/underrepresentation-of-women-in-science-international-and-cross-disciplinary-evidence-and-debate>
- Committee on Maximizing the Potential of Women in Academic Science and Engineering, National Academy of Sciences, National Academy of Engineering, Institute of Medicine (2007). *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Washington, DC: The National Academies Press.
- Dunn, D., Gerlach, J. M., and Hyle, A. E. (2014). Gender and leadership: reflections of women in higher education administration. *Int. J. Leadersh. Change* 2, 9–18. Available online at: <http://digitalcommons.wku.edu/ijlc/vol2/iss1/2>
- Eagly, A. H., and Carli, L. L. (2007). *Through the Labyrinth: The Truth About How Women Become Leaders*. Boston, MA: Harvard Business School Press.
- Ecklund, E. H., and Lincoln, A. E. (2011). Scientists want more children. *PLoS ONE* 6:e22590. doi: 10.1371/journal.pone.0022590
- Ginther, D. K., and Kahn, S. (2009). "Does science promote women? Evidence from Academia 1973–2001," in *Science and Engineering Careers in the United States (National Bureau of Economic Research Conference Report*, eds R. B. Freeman and D. F. Goroff (Chicago, IL: University of Chicago Press), 163–194.
- Ginther, D. K., and Kahn, S. (2015). "Women's careers in academic social science: progress, pitfalls, and plateaus," in *The Economics of Economists*, eds A. Lanteri and J. Vroman (Cambridge: Cambridge University Press), 285–315.
- Gorman, S. T., Durnowicz, M. C., Roskes, E. M., and Slattery, S. P. (2010). *Women in the Academy: Female Leadership in STEM Education and the Evolution of a Mentoring Web*. Forum on Public Policy Online, v2010 n2 2010. Available online at: <http://eric.ed.gov/?id=EJ903573>
- Goulden, M., Frasch, K., and Mason, M. (2009). *Staying Competitive: Patching America's Leaky Pipeline in the Sciences*. The Center for American Progress.
- Hill, C., Corbett, C., and St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: American Association of University Women.
- Hough, M. A. (2010). *Leadership Styles of Effective Female Administrators in Higher Education*. Unpublished Doctoral dissertation: ERIC Number: ED535067. Available online at: <http://search.proquest.com/docview/902033280> ISBN-978-1-1249-8206-9
- Jones, R. D., Griffith, K. A., Ubel, P. A., Stewart, A., and Jagsi, R. (2016). Mixed-methods investigation of the motivations, goals, and aspirations of male and female academic medical faculty. *Acad. Med.* 91, 1089–1097. doi: 10.1097/ACM.0000000000001244
- National Research Council (2009). *Gender Differences at Critical Transitions in the Careers of Science, Engineering and Mathematics Faculty*. Washington, DC: National Academies Press.
- Raj, A., Carr, P. L., Kaplan, S. E., Terrin, N., Breeze, J. L., and Freund, K. M. (2016). Longitudinal analysis of gender differences in academic productivity among medical school faculty across 24 medical schools in the United States. *Acad. Med.* 91, 1074–1079. doi: 10.1097/ACM.0000000000001251
- Smith, C. S. (2014). "Assessing academic STEM women's sense of isolation in the workplace," in *Alliances for Advancing Academic Women*, eds P. J. Gilmer, B. Tansel, and M. H. Miller (Rotterdam: Sense Publishers), 97–117.
- Smith, D. G., Turner, C. S. V., Osei-Kofi, N., and Richards, S. (2004). Interrupting the usual: successful strategies for hiring diverse faculty. *J. High. Educ.* 75, 133–160. doi: 10.1353/jhe.2004.0006
- Williams, J. C., Phillips, K. W., and Hall, E. W. (2014). *Double Jeopardy: Gender Bias Against Women of Color in Science. Tools for Change*. Available online at: <http://www.uchastings.edu/news/articles/2015/01/double-jeopardy-report.pdf>
- Williams, W. M., Barnett, S. M., and Sumner, R. A. (2013). "Where are all the women in academic science, technology, engineering, and mathematics fields?" in *Handbook of Research on Promoting Women's Careers* (London: Edward Elgar), 326–338.
- Williams, W. M., Barnett, S. M., and Wethington, E. (2015). "What women in science need to know about work-life balance" in *Success Strategies of Women in Science: A Portable Mentor*, ed P. Pritchard (New York, NY: Elsevier), 391–411.
- Williams, W. M., and Ceci, S. J. (2012). When scientists choose motherhood. *Am. Sci.* 100, 138–145. doi: 10.1511/2012.95.138
- Williams, W. M., and Ceci, S. J. (2015). National Hiring experiments reveal 2-to-1 faculty preference for women on stem tenure track. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5360–5365. doi: 10.1073/pnas.1418878112
- Xie, Y., and Shauman, K. A. (2003). *Women in Science: Career Processes and Outcomes*. Cambridge, MA: Harvard University Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Williams, Mahajan, Thoenes, Barnett, Vermeylen, Cash and Ceci. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

