NORTHWESTERN UNIVERSITY

Characterizing Pathways for Joining STEM in College and Beyond

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

David Isaac Miller

EVANSTON, ILLINOIS

June 2018

Abstract

Students often change majors during college, and most workers change jobs throughout their careers. Yet the diverse opportunities for entering natural science, technology, engineering, and mathematics (STEM) fields are often overlooked during college and beyond. This dissertation therefore analyzed four large nationally representative datasets to characterize the pathways for joining STEM among college students and workers from non-STEM backgrounds.

My theoretical framework contrasts sharply with the popular "leaky pipeline" metaphor, which has often pervaded debates about STEM talent in the United States. Since the 1980s, this metaphor has focused attention on plugging "leaks" in the pipeline (i.e., increasing STEM persistence) but away from more comprehensive strategies for addressing workforce needs. I argue that educators, employers, and policymakers should instead think of a network of pathways along which students and workers can take different routes to STEM competence. This pathways metaphor provides novel ways for thinking about developing STEM talent. For instance, late entry points into STEM offer new opportunities for recruiting women, especially given the large pool of women who start college as a non-STEM major and graduate with non-STEM degrees.

Across two studies, my dissertation studied pathways for joining STEM in the transitions from (a) beginning of college to graduation and (b) college graduation to the workforce. Nearly one fifth of STEM graduates started college as a non-STEM major, and one fifth of college-educated STEM workers had no bachelor's degree or higher in any STEM field. Analyses described joiners' educational history and job characteristics, identified predictors of STEM joining, and estimated the national impact of further widening joining pathways.

Study 1 found that joiners had similar high school STEM preparation compared to so-called "leaks" in the STEM pipeline (i.e., students who went from STEM to non-STEM majors) but had

weaker preparation than persisters (i.e., STEM graduates who started college in STEM). Nevertheless, compared to persisters, joiners achieved similar undergraduate success in terms of STEM course grades and rates of graduating college on time (i.e., within four years). Grades in introductory STEM courses strongly predicted STEM attrition but not STEM joining, indicating asymmetric pathways for leaving versus joining STEM. In contrast to grades, taking STEM courses early in college strongly predicted later earning of STEM bachelor's degrees among initial non-STEM majors, even after controlling for many theoretically relevant covariates. Impact analyses also found that, compared to "plugging the leaky pipeline" for female STEM majors, closing gender gaps in STEM joining would more potently increase women's representation among STEM graduates.

Study 2 studied pathways for joining the computing and engineering workforce among non-STEM college graduates. Results suggested that joiners used their non-STEM educational training by working on non-STEM job tasks such as finance and management at higher rates than persisters. Overall, pathways were far more open for joining the computing than engineering workforce. Careers in both these fields appeared to be unattractive to non-STEM graduates who valued benefitting society, consistent with common perceptions that STEM jobs lack career opportunities to help others. Impact analyses found that an additional 820,000 college graduates would have been computer scientists in 2015 if communally oriented workers had joined computing as often as other non-STEM graduates. Most of these additional computer scientists would have been women.

These results suggest new strategies for broadening participation in STEM. Based on Study 1, postsecondary educators and policymakers should evaluate how to facilitate STEM joining pathways by increasing the quality and quantity of STEM courses that non-STEM majors take early in college. Based on Study 2, employers should consider how to communicate to potential applicants, especially non-STEM graduates, that STEM careers offer ways to help others.

Acknowledgements

I thank my main research advisors, David Uttal and Alice Eagly, for their support and intellectual guidance since my first year of graduate school at Northwestern. Thank you, David, for taking me on as your student and giving me the freedom throughout graduate school to pursue research ideas that interested me. Thank you, Alice, for also becoming my advisor and providing input from a social psychological perspective, even though I was technically in another program area within the psychology department. I also thank my other dissertation committee member, Larry Hedges, for his expert advice on implementing advanced statistical methods and navigating the world of educational research and policy. I feel lucky to have had such an impressive and interdisciplinary group of advisors to help form this dissertation and launch my career. This research was supported by the Spatial Intelligence and Learning Center (National Science Foundation Grants SBE-0541957 and SBE-1041707) and a National Science Foundation Graduate Research Fellowship awarded to D. I. Miller (DGE-0824162 and DGE-1324585). This research was also based on data from the National Longitudinal Survey of Freshmen, a project designed by Douglas S. Massey and Camille Z. Charles and funded by the Mellon Foundation and the Atlantic Philanthropies.

Table of Contents

List of Tables

List of Figures

Chapter 1: An Overview

Many U.S. policymakers and employers have urgently called for more workers with skills and training in natural science, technology, engineering, and mathematics (STEM) fields to meet the needs of an economy that increasingly relies on technological innovation (National Science Board, 2018). The U.S. STEM workforce has grown much faster than the general U.S. workforce in past decades and is projected to do so over the next decade (Noonan, 2017). In addition, the demand for technological skills and STEM knowledge has been increasing in nearly all fields of employment, even for jobs not traditionally categorized as STEM jobs (President's Council of Advisors on Science and Technology [PCAST], 2012). Failing to address the need for more workers with STEM skills could threaten the United States' economic growth and capacity to innovate. As the National Science Board (2018) argued, such workers "are integral to a nation's innovative capacity because of their high skill level, their creative ideas, and their ability not only to advance basic scientific knowledge but also to transform advances in fundamental knowledge into tangible products and services" (p. 3-9).

To help address these national needs, many employers have sought to broaden the STEM talent pool by recruiting women (Saujani & Sweet, 2016). Although women are roughly half of the U.S. workforce, they are sparsely represented in several STEM fields. Among college-educated workers, for instance, women were 25% of computer science workers and 15% of engineering workers in 2015 (National Survey of College Graduates, 2015). In addition to addressing workforce needs, recruiting women into STEM is important for advancing gender equity. Compared to non-STEM jobs, STEM jobs offer women higher pay and lower unemployment rates, especially in computing and engineering (National Science of Board, 2018). In other words, the lack of gender diversity in STEM limits both employers' access to the broadest possible STEM talent pool and women's access to high-value jobs.

My dissertation analyzed four nationally representative datasets to inform policies and interventions for addressing both national issues of (a) need for more workers with STEM skills and (b) lack of gender diversity in STEM. More specifically, my analyses focused on characterizing pathways for joining STEM among non-STEM majors and non-STEM degree holders. This focus on late entry into STEM contrasts with the dominant focus on STEM attrition (i.e., leaving STEM) found in most other research on STEM transitions among adults. My dissertation shows that the pathways for becoming scientists and engineers are far less rigid and linear than commonly assumed. Taking advantage of the diverse entry points into STEM could offer new solutions for addressing workforce needs and recruiting women into STEM. In contrast, dominant theoretical frameworks that focus solely on STEM attrition overlook the varied trajectories for obtaining STEM degrees and becoming STEM workers.

History of the Leaky Pipeline Metaphor

The metaphor of a "leaky pipeline" has pervaded debates about the nation's limited supply of STEM talent for several decades. According to this metaphor, obtaining a STEM degree requires travelling a structured pipeline of key educational milestones such as taking calculus in high school, declaring a STEM major in college, and persisting until college graduation (Cannady, Greenwald, & Harris, 2014). Students with interest and aptitude for STEM often "leak" from the STEM educational pipeline at multiple time points creates a limited supply of STEM-trained workers. In her landmark report *Who Will Do Science?*, Berryman (1983) used this metaphor to explain the low numbers of women and underrepresented racial minorities among STEM Ph.D. earners. These demographic groups are sparsely represented at the Ph.D. level, she argued, because they leave the STEM educational pipeline at higher rates than other groups at multiple transitions between graduating high school and earning Ph.D. degrees.

Berryman (1983) has often been credited as the first to popularize this STEM pipeline metaphor, which has been widely used since her report (Xie & Shauman, 2003). For instance, Alper (1993) further popularized it by publishing an influential review article in *Science* provocatively entitled, “The Pipeline is Leaking Women All the Way Along.” By the early 2000s, the pipeline metaphor had become so widespread that it was “commonly accepted as the dominant, if not the standard, conceptual framework within which to organize studies of the science educational and career trajectory,” Xie and Shauman (2003) argued (p. 8). In addition to explaining the underrepresentation of women in STEM, the pipeline metaphor has been applied to describe overall shortages of STEM graduates. For instance, graphic illustrations in STEM education reports have often depicted a progressively smaller STEM pipeline as students “leak” from it between the beginning of high school to college graduation (e.g., National Wildlife Federation, 2013; New York Hall of Science, 2012; North Shore Technology Council, 2013). In this way, the “leaky pipeline” metaphor has become widely used to describe both the national issues of STEM workforce shortages and lack of gender diversity.

Despite its widespread use, the pipeline metaphor has also received criticism from a minority of scholars (e.g., Bennett, 2011; Cannady et al., 2014; Xie & Shauman, 2003). For instance, Xie and Shauman (2003) argued that the pipeline metaphor does not capture the complex, multifaceted processes of becoming a scientist and directs attention away from considering alternative educational and career trajectories such as late entry into STEM. Recent empirical research has also found fundamental inaccuracies in the metaphor’s application (Cannady et al., 2014; Miller & Wai, 2015). For instance, many U.S. students earn STEM degrees without having completed milestones that are often assumed to be prerequisites. One such example is taking calculus in high school. One nationally representative dataset found that, among STEM bachelor's degree earners who entered postsecondary education in 2003, almost half (49%) had not taken calculus in high school[1] (Beginning Postsecondary Students 04/09 Sample, 2016). In another national longitudinal study, many STEM bachelor's degree earners (39%) had not intended to pursue STEM when asked in either the 8th or 12th grade (Cannady et al., 2014). These statistics show that the processes for obtaining STEM degrees in the United States are much more diverse than what the STEM pipeline metaphor might suggest. In other words, new frameworks are needed that more fully and accurately capture the complex educational and career trajectories in STEM fields.

Pathways as a New Metaphor

Along with other scholars, I propose replacing the metaphor of a singular STEM pipeline with a network of multiple pathways into and out of STEM fields (Cannady et al., 2014; Mervis, 2012; PCAST, 2012; Xie & Shauman, 2003). For instance, a 2012 report by the President's Council of Advisors on Science and Technology (PCAST) argued that, "rather than a single pipeline that is prone to leakage…educators and policymakers should think of a network of pathways along which students can take different routes to STEM readiness and competency" (PCAST, 2012, p. 31). This perspective builds on prior structural and psychological theoretical perspectives. Structural factors such as institutional course requirements can create, facilitate, or constrain certain pathways (Charles & Bradley, 2009; Merolla, Serpe, Stryker, & Schultz, 2012). However, students will choose and travel these pathways differently because of psychological factors such as interests, social identities, and prior academic preparation (Diekman, Clark, Johnston, Brown & Steinberg,

[1] This statistic varied some across bachelor's degree field: biological and biomedical sciences (47%), computer science (64%), engineering (39%), mathematics and statistics (30%), physical sciences (46%),

2011; Eccles, 2011; Leslie, Cimpian, Meyer, & Freeland, 2015; Nix, Perez-Felkner, & Thomas, 2015; Wai, Lubinski, Benbow, & Steiger, 2010). Based on Eccles' (2011) expectancy-value theory, for instance, individuals make educational choices based on (a) expectations for success and (b) the value given to the perceived available options. This theory emphasizes that individuals make relative assessments of the benefits and costs of different pathways. For instance, among individuals with high math performance, those with weaker verbal performance are more likely to pursue STEM careers (Wang, Eccles, & Kenny, 2013). This result suggests individuals compare performance across domains when assessing their strengths and making occupational choices.

Compared to the leaky pipeline metaphor, this new metaphor of STEM pathways better represents developmental theories of identity formation such as Arnett's (2000) theory of emerging adulthood. Arnett proposed that the developmental period between late teens and early twenties is one in which, "independent exploration of life's possibilities is greater for most people than it will be at any other period" (p. 469). Consistent with this hypothesis, empirical data show that students frequently explore different academic pathways, regardless of their initial major. In one study, for instance, roughly one half of beginning bachelor's degree students left their initial intended major by either switching majors or leaving postsecondary education (Chen & Soldner, 2013). Surprisingly, attrition rates in STEM fields were typically equal or lower than those in other fields. STEM fields are therefore not especially "leaky"; the "leaks" instead reflect young adults' more general tendency to explore different academic identities. The pathways metaphor captures this phenomenon by recognizing this explorative nature of young adults. In contrast, the pipeline metaphor misrepresents identity formation by implicitly suggesting that attrition is especially high in STEM compared to other majors.

The pathways metaphor also offers novel approaches for addressing STEM workforce needs by directing attention to pathways for late entry into STEM such as joining STEM after starting college as a non-STEM major. These alternative pathways may be particularly important for increasing gender diversity. Few women start college intending to major in certain STEM fields such as computer science and engineering (Ceci et al., 2014). Because this initial pool of women is small, increasing the persistence of those female STEM majors may not generate many more female STEM graduates. Educators and policymakers could instead look to currently untapped sources of talent such as female non-STEM majors to more potently increase gender diversity. As the 2012 PCAST report argued, educators have "given much attention to 'off-ramps,' the drop-out and attrition patterns," but should also give "equal attention…to on-ramps, multiple routes to enter or re-enter STEM education" (PCAST, 2012, p. 30-31).

My Dissertation

Addressing Gaps in Prior Literature

My dissertation characterized pathways for joining STEM from non-STEM fields – hereafter called *STEM joining* – during the transitions from (a) the first year of college to college graduation and (b) college graduation to the STEM workforce in the United States. Analyzing both transitions allowed me to more comprehensively characterize the often overlooked educational and career trajectories in STEM and their implications for meeting pressing national workforce needs. Prior research on nontraditional pathways for entering STEM has typically focused on transitions before college or between high school and college, but not during or after college (e.g., Cannady et al., 2014; Xie & Killewald, 2012). This prior research therefore leaves unanswered questions about the openness of joining STEM as students enter emerging adulthood – a period of development in which students explore many other identities and worldviews (Arnett, 2000). Nearly all research on postsecondary transitions in STEM has focused on students leaving, not joining, STEM fields (e.g., Chen & Soldner, 2013; Graham et al., 2013; Miller & Wai, 2015; Watkins & Mazur, 2013). College educators may therefore be unaware of how to best leverage the potentially diverse postsecondary pathways into STEM and support students aiming to make that transition.

Furthermore, most prior research has also assumed that STEM degrees are needed to pursue STEM careers (Cannady et al., 2014, Figure 2; National Science Board, 2018). This assumption overlooks the myriad of ways that workers might learn STEM skills after formal education (e.g., through on the job training or gradually taking on more STEM-relevant job tasks). National statistics indicate that even the pathways for joining the STEM workforce are far more open and complex than often assumed. For instance, in 2015, approximately 1.2 million STEM workers were college graduates without a bachelor's degree or higher in any STEM field, which was one-fifth of the college-educated STEM workforce (National Survey of College Graduates, 2015). In other words, STEM employers already hire many workers without STEM degrees, but have little guidance from scholarly research on best practices for doing so. The aim of this dissertation was to address these critical gaps in prior literature by examining STEM joining pathways during college and beyond.

Defining STEM Joiners and STEM Fields

STEM joiners were defined as people who entered STEM from a non-STEM background. However, the exact operational definition varied across this dissertation's two studies. In the first study on postsecondary education transitions, "STEM joiners" referred to STEM bachelor's degree earners who intended a non-STEM major during students' first year at a four-year institution. In the second study on education to workforce transitions, "STEM joiners" referred to college-educated STEM workers who had no bachelor's or graduate STEM degree.

For both studies, the definition of "STEM majors" was based on the National Center for Education Statistics' classification of STEM fields, which included physical and life science, but excluded social science and health/nursing (Chen & Soldner, 2013). Under this broad category of "STEM," women's representation varied greatly by field. For instance, in 2015, women earned 59% of the U.S.'s bachelor's degrees in life science, but only 25% of bachelor's degrees in physical science, technology, engineering, and mathematics (pSTEM; WebCASPAR, 2018). For this reason, any gender-related analysis included at least some level disaggregation (e.g., separately analyzing life science vs. pSTEM majors).

Answering Cross-Cutting Research Questions

In both studies, analyses were organized around answering three cross-cutting questions: (a) who are STEM joiners? (b) what predicts STEM joining? (c) what is the impact of widening joining pathways? Investigating these three questions helped describe the basic phenomenon of STEM joining as well as identify points of intervention and their potential impact on solving national workforce needs. These analyses can help form the basis for future research on developing policies and supports for leveraging the diverse pathways into STEM during college and beyond.

Who are STEM joiners? For each study, I first focused on developing a descriptive account of STEM joiners in terms of their prior educational history, demographics, and specific types of STEM majors and jobs they pursued. Answering this question involved comparing STEM joiners to other relevant groups such as people who persisted in STEM (e.g., intended a STEM major early in college and then later earned a STEM degree) on dimensions such as high school STEM preparation and job characteristics. Simple descriptive statistics such as means and percentages were used to answer this first question; no complex statistical modeling (e.g., multiple regression analysis) was needed.

What predicts STEM joining? Second, I identified factors that predicted which potential joiners (i.e., people from non-STEM backgrounds) later pursued STEM majors and jobs. For instance, in the first study on postsecondary transitions, analyses examined how two early-college factors – course-taking and grades – predicted later earning of STEM bachelor's degrees. Psychological theories on identity formation and goal pursuit guided the selection of predictors for analysis and interpretation of correlational results. Multivariable models such as multiple regression and propensity score matching were used to control for potential self-selection effects and other confounds. These analyses aimed to identify what factors had a causal effect on students' decisions; however, conclusions about causal effects were limited given the correlational nature of these data. Experimental data would have been needed to more definitively rule out alternative explanations.

What is the impact of widening joining pathways? Finally, I explored the potential national impact of increasing STEM joining rates. These analyses estimated the current national prevalence of STEM joining and quantified how much joining rates would need to change to meet projected STEM workforce growth (PCAST, 2012; Noonan, 2017). These analyses also considered the national impact of increasing women's joining rate to match that of men's. For instance, these analyses considered the effectiveness of closing the gender gap in STEM joining versus retention on increasing gender diversity among STEM graduates. Like the first research question, simple descriptive statistics such as percentages and sums of probability survey weights were used to answer these questions.

Chapter 2: Pathways for Joining STEM During College

For over twenty years, the "leaky pipeline" metaphor has pervaded national debates about producing natural science, technology, engineering, and mathematics (STEM) graduates (Alper, 1993). Based on this characterization, the U. S. suffers from a shortage of STEM graduates because students "leak" from the STEM pipeline at multiple points. Women are thought to leak from this pipeline at higher rates than men, which explains the low numbers of women among STEM Ph.D. earners and professors (Miller & Wai, 2015). Here we take a different approach by focusing on students who join STEM from non-STEM majors (hereafter, *STEM joiners*) and studying factors that facilitate that transition. Addressing gender gaps in joining STEM could also provide a novel strategy for increasing gender diversity among STEM bachelor's degree earners.

Many students substantially change their identities and worldviews during college. According to Arnett (2000), the late teens and early twenties is a period of *emerging adulthood* in which, "independent exploration of life's possibilities is greater for most people than it will be at any other period" (p. 469). Nationally representative studies confirm that college is an unstable time for many students. In one study, roughly one half of beginning bachelor's degrees students left their initial intended major by either switching majors or leaving postsecondary education (Chen & Soldner, 2013). Surprisingly, attrition rates in STEM fields were not higher than those for other fields, and in some cases, were lower. STEM fields are therefore not especially "leaky"; the "leak" instead reflects a more general tendency of young adults to "try on" multiple academic identities.

The STEM pipeline metaphor therefore misrepresents identity formation during emerging adulthood by suggesting that attrition is especially high in STEM compared to other majors. The metaphor also misrepresents why some groups such as women are underrepresented in STEM. Contrary to common perceptions, women and men now persist in physical science and engineering at comparable rates for many segments between college and achieving academic tenure (Ceci, Ginther, Kahn, & Williams, 2014; Miller & Wai, 2015). Gender gaps in STEM persistence exist before college, but the pathways for joining STEM are especially open at those earlier ages. In one national longitudinal study, for instance, many STEM bachelor's degree earners (39%) had not intended to enter STEM when asked in *either* 8th or 12th grade (Cannady et al., 2014). New frameworks are therefore needed for thinking about how to broaden participation in STEM fields.

We propose replacing the metaphor of a singular STEM pipeline with a network of multiple pathways into and out of STEM fields, along with other scholars (Cannady et al., 2014; Mervis, 2012; PCAST, 2012; Xie & Shauman, 2003). This perspective on multiple pathways builds on prior structural and psychological theoretical perspectives. Structural factors such as institutional course requirements can create, facilitate, or constrain certain pathways (Charles & Bradley, 2009; Merolla et al., 2012; Wang, 2013). However, students will choose and travel these pathways differently because of psychological factors such as interests, social identities, and prior academic preparation (Diekman et al., 2011; Eccles, 2011; Leslie et al., 2015; Nix et al., 2015; Wai et al., 2010).

Based on these theoretical perspectives, we hypothesized that taking STEM courses early in college would be especially important for two major reasons: structural and psychological. First, taking STEM courses early in college (e.g., during the first year) could create structural pathways for satisfying academic requirements for joining STEM. In contrast, taking STEM courses early in college could both foster interest in STEM and enable more advanced STEM course-taking in subsequent semesters. Such exposure can help students to realize how STEM fields align with some individuals' goals and interests (e.g., Diekman et al., 2011; Eccles, 2011; Merolla et al., 2012). Thus, similar to brief social psychological interventions (Walton, 2014), taking even just one additional STEM course early in college could initiate recursive processes that unfold over years, eventually changing students' majors.

We hypothesized also that earning high grades in introductory STEM courses would also contribute to decisions to enter STEM. According to Eccles' (2011) expectancy-value model, grades serve as objective cues shaping students' expectancies for success in a particular domain. Grades can also reflect students' motivations to persist on challenging domain-specific tasks. Much prior psychological literature has therefore understandably focused on grades as critical predictors and dependent variables in STEM fields (Chen & Soldner, 2013; Miller & Halpern, 2014).

The pre-college pathways of STEM graduates are considerably diverse (Cannady et al., 2014; Maltese & Tai, 2011; Xie & Killewald, 2012; Xie & Shauman, 2003), but the openness of pathways for joining STEM during college remains unclear (Crisp, Nora, & Taggart, 2009; Griffith, 2010; Whitten et al., 2007). We therefore analyzed three large nationally representative samples to characterize the pathways for earning a STEM bachelor's degree after intending a non-STEM major during students' first year of college. Analyses focused on three guiding research questions about STEM joining pathways:

(1) Who are STEM joiners in terms of their educational backgrounds and outcomes?

(2) What early-college factors help facilitate joining STEM?

(3) What is the potential national impact of increasing STEM joining rates?

The first and third research questions involved obtaining descriptive statistics (e.g., means, percentages), whereas the second question involved conducting inferential analyses that controlled for potential self-selection and other confounds. For the second question, we used pre-college variables such as high school STEM preparation as covariates when estimating the effects of early-college variables.

Method

We analyzed three nationally representative samples to characterize STEM joiners and the educational pathways they took. Each sample had its unique methodological strengths that collectively balanced the limitations of the others. For instance, one sample (Beginning Postsecondary Students) was best for estimating descriptive statistics about STEM joiners, whereas the other two samples were best for controlling for self-selection when estimating the causal effect of taking STEM courses early in college.

Data sources. The datasets were Beginning Postsecondary Students (BPS), National Longitudinal Survey of Freshman (NLSF), and Project TALENT. The BPS sample (n = 16,680) was a stratified national probability sample whose target population was all students who started postsecondary education for the first time during the 2003-04 academic year at any U.S. postsecondary institution (Wine, Janson, & Wheeless, 2011). Students in the BPS sample were surveyed at three time points between entering postsecondary education and six years after entering postsecondary education (spring 2004 to summer 2009). NLSF (n = 3,924) was a randomly selected sample of first-time undergraduates who entered college in fall 1999 at one of 28 selective 4-year public and private U.S. institutions (Charles, Fischer, Mooney, & Massey, 2009). Students in the NLSF sample were surveyed at five time points between entering college and four years after entering college (fall 1999 to spring 2003); the National Student Clearinghouse and the institutions' registrar offices provided data on students' six-year graduation status. The Project TALENT sample (n = 346,666) was a stratified national probability sample whose target population was all students who attended high school in 1960 (Wise, McLaughlin, & Steel, 1979). Students in Project TALENT were surveyed at four time points between high school and 11 years after graduating high school (spring 1960 to fall 1974).

Initial field of study. Students were asked slightly different questions about their initial major intentions across the three samples. In the first longitudinal wave, students in the BPS sample were asked, "Have you declared a major at [primary undergraduate school]?" and students in the NLSF sample were asked, "Have you already chosen a major?" If students answered "no," they were categorized as "undecided" and not asked any further questions about their intended major field of study. In contrast, students in the Project TALENT sample were asked, "In which of the following areas do you expect to specialize or 'major' in college? Mark ONE even if you haven't made up your mind definitely." Hence, no students were categorized as "undecided" in the Project TALENT sample because of the question's wording.

Participant inclusion criteria. Unless otherwise noted, our descriptive and inferential analyses focused on students who (i) started postsecondary education at a 4-year institution, (ii) intended a specific major during their first year of college, and (iii) earned a bachelor's degree within six years of entering college. Students who started postsecondary education at a 2-year institution were excluded to improve comparability across datasets (e.g., BPS vs. NLSF) and to focus our analyses. Only students who had intended a specific major were included because they were less ambiguous than undecided students. Undecided students could have been initially deciding between two STEM fields (e.g., physics or chemistry), a STEM and non-STEM field, two non-STEM fields, or other possibilities. Students labeled "undecided" in the BPS sample also could have chosen an initial major but simply not formally declared it yet when asked "Have you declared a major?" at the end of their first year in college. This ambiguity further supported our decision to include only students who initially intended a specific major. Only bachelor's degree earners were included because our theoretical focus concerned differences between attaining STEM versus non-STEM degrees. Non-degree earners were therefore excluded, but supplemental analyses investigated the robustness of our conclusions to the inclusion of non-degree earners. As shown in an interactive website (https://d-miller.shinyapps.io/joiningSTEM/), results were robust to inclusion of non-degree earners.

Descriptive analyses. Descriptive analyses investigated the prevalence of STEM joining, potential impact of increasing joining rates, and pre-college and undergraduate characteristics of STEM joiners. These analyses focused on the BPS sample because it was recent and representative of all U.S. postsecondary institutions. Government agencies routinely use BPS when forming policy decisions (PCAST, 2012). Data on students' graduation status and course taking were based on transcripts collected through institutional registrar offices (Wine et al., 2011). For all analyses, we used the weighting variable (WTB000) created by the National Center for Educational Statistics that accounted for student-level nonresponse bias and unequal sampling probabilities. An individual's survey weight indicates the approximate number of people that the individual represents in the population. Sums of these weights were therefore used to estimate the size of various subpopulations (e.g., the number of STEM joiners per year). Table S7 in the Supplemental Materials available online presents detail on the more complicated calculations (i.e., projections for the hypothetical scenarios shown in Figures 7 & 8).

Inferential analyses. Interpreting correlational data can be difficult, but multiple methods can help eliminate alternative explanations (Steiner, Cook, Shadish, & Clark, 2010). We therefore used multiple methods to (i) estimate the causal effect on early-college STEM course-taking on later STEM joining, (ii) consider the plausibility of relevant model assumptions, and (iii) evaluate how violating those assumptions would have affected our estimates. Because correlational data often has serious limitations, using such multiple methods is needed to rigorously balance the strengths and limitations of each approach (Ichino, Mealli, & Nannicini, 2008). And as explained further, the datasets also helped eliminate alternative explanations by being longitudinal and having a rich set of informative covariates.

We used multilevel logistic regression (Raudenbush & Bryk, 2002) and propensity score matching analyses (Caliendo & Kopeinig, 2008) to estimate causal effects by controlling for potential confounds (e.g., prior STEM course-taking, prior interests in STEM). These methods would have yielded unbiased causal estimates if the *conditional independence assumption* was true (also called the *unconfoundedness assumption*): conditional on observed covariates, the probability of receiving treatment was independent of the potential treatment outcome (Ichino et al., 2008). The conditional independence assumption is true for experimental designs because random assignment ensures that treatment status is independent of other factors that directly influence the potential treatment outcome. The plausibility of this assumption for non-experimental designs depends on controlling for a large and informative set of observed covariates, such as those contained in the NLSF and Project TALENT datasets (see Tables S3 and S4 for complete covariate lists). Controlling for such covariates can help reduce potential bias due to self-selection (Steiner et al., 2010).

We also used regression models of STEM course taking to help consider what magnitudes of self-selection were plausible among students intending a non-STEM major. We first estimated the magnitude of self-selection due to observable variables (e.g., prior STEM course-taking, perceived difficulty of STEM courses) and then compared this magnitude to that among undecided students. This comparison allowed us to test the hypothesis that self-selection due to observed variables was smaller among students intending a non-STEM major than among undecided students.

Finally, sensitivity analyses evaluated how unobserved confounders would have influenced our causal estimates (Ichino et al., 2008). We simulated unobserved confounders by varying the extent to which they might have related to STEM course taking (treatment status) and STEM joining (treatment outcome). Such simulations showed how sensitive our causal estimates were to unobserved confounders. This methodological approach aligns with Ichino and colleagues' (2008) argument that non-experimental studies have undeniable value but should "be put under the scrutiny of a sensitivity analysis…before being accepted as a guide for policy" (p. 325). Hence, testing causal claims with non-experimental data should be done cautiously, using multiple methods.

Covariates for inferential analyses. The validity of causal inference in quasi-experimental studies such as ours depends on controlling for a large and informative set of relevant pre-treatment covariates (Steiner et al., 2010). For this reason, inferential analyses focused on the NLSF and Project TALENT samples because they included diverse sets of pre-college covariates that were potentially related to self-selection into early-college STEM courses. Covariates included demographics, high school STEM preparation (i.e., high school STEM course-taking, grades, and standardized test scores), educational and occupational plans in high school, perceived difficulty of STEM courses, and self-reported and behavioral indicators of STEM interests (see Tables S1 and S2 in the supplemental materials for complete covariate lists). Prior research and theoretical frameworks have shown these covariates to be important to students' decisions to pursue STEM fields (e.g., Cannady et al., 2014; Eccles, 2011). Many of these pre-college covariates likely also influenced STEM joining, but out research goals focused more on estimating effects of early-college variables, especially early-college STEM course-taking. Unless otherwise noted, the BPS sample was not used for inferential analyses because its set of STEM-related, pre-college covariates was far less extensive.

Definition of STEM fields. The following fields were defined as STEM fields: biological/agricultural sciences, computer and information sciences, engineering, mathematics/statistics, and physical sciences. Social science was categorized as non-STEM for the specific purposes of our analyses. The practice-oriented fields of health and medicine were also categorized as non-STEM, consistent with categorization schemes used by the National Science Foundation (NSF) and other researchers (e.g., Chen & Soldner, 2013; PCAST, 2012; Xie & Killewald, 2012). However, supplemental analyses investigated the robustness of our conclusions to the inclusion of initial health majors. For double majors, we only analyzed the major that students listed as primary. We typically analyzed "STEM" as an aggregate group, but also repeated all analyses disaggregating STEM (e.g., compare students who joined life science vs. physical science fields); as explained in the Results section, disaggregating STEM fields was especially important to study gender differences.

Methodological strengths and limitations. All three samples had unique methodological strengths. The BPS sample provided a recent descriptive account of STEM joining that was representative of all U.S. postsecondary institutions. The NLSF sample focused on selective institutions, which have high rates for producing students who later earn STEM graduate degrees (Fiegener & Proudfoot, 2013). The Project TALENT sample provided the most rigorous controls for self-selection confounds. The diversity of these samples also established the robustness and generalizability of our results across four decades and institutional selectivity.

Although the strengths of each sample often balanced the limitations of another, some general limitations should be noted. Data on STEM course-taking level (e.g., calculus vs. pre-calculus) or pedagogical style (e.g., lecture-based vs. flipped classroom) were generally not available, for instance. Also if students answered they were undeclared (in the BPS sample) or hadn't yet chosen a major (in the NLSF sample), they were labeled "undecided" and were asked no further questions about their intended field of study. This methodological limitation meant that undecided students were an ambiguous group of students and therefore excluded from our analyses, unless otherwise noted. Future studies should examine the educational pathways of such students, but doing so was beyond the scope of this present study.

Of course, the datasets also did not measure all potentially relevant variables that could have been self-selection confounds. Because self-selection confounds can threaten the validity of causal inference, our analyses explicitly considered the potential impact of unobserved variables (see Results 4 and 6 in Table S7). Concerns about unobserved confounds were also somewhat mitigated because such confounds likely correlated with some of the observed variables. For instance, Eccles' (2011) expectancy-value theory posits that interests are proximal influences on educational decisions such as choosing a STEM major. According to this theoretical model, interests mediate the effects of more distal factors such as parental expectations and gender stereotypes. Controlling for proximal factors such as interests in STEM or prior STEM course taking could therefore partially control for more distal confounds (e.g., parental expectations).

Results

Who Are STEM Joiners?

Our first goal was to understand who STEM joiners are by using the BPS sample to characterize their initial majors, demographics, pre-college STEM preparation, and undergraduate success.

Initial majors of STEM graduates. STEM joiners accounted for 18% of the U.S.'s STEM bachelor's degree earners who entered college in 2003 (BPS sample). These joiners came from diverse non-STEM fields, especially health, business, social science, and education (Figure 2). Students who were initially undecided were another 23% of STEM graduates.

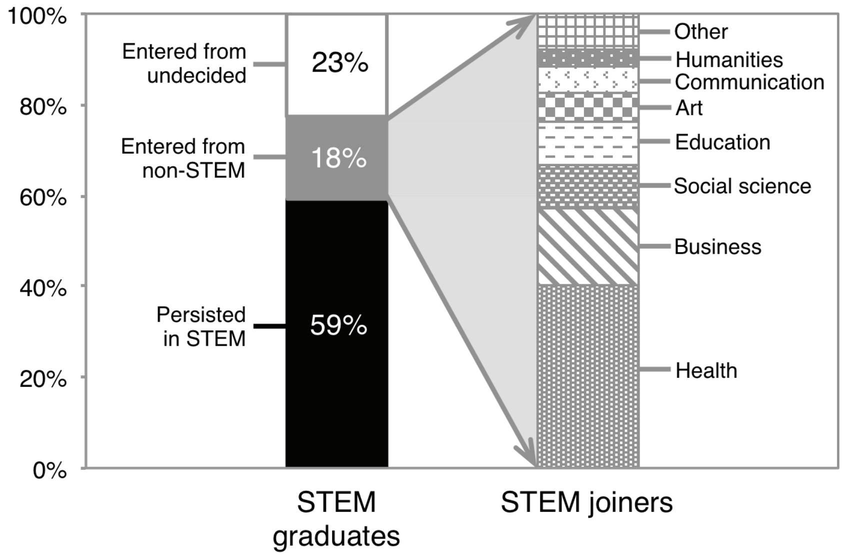

Figure 1. The initial intended majors of STEM graduates and STEM joiners (BPS sample).

Demographics. STEM joiners were gender-balanced (51% women), though women were more common among students who joined life science (58% women) than physical science, technology, engineering, and mathematics (pSTEM) fields other STEM fields (40% women). We elaborate on this finding further below. The representation of Black and Latino students among STEM joiners (14% of joiners) was not significantly different ($p > .250$) from their representation among other initial non-STEM majors who later earned a bachelor's degree (16%). In contrast, Asian students were twice as common among STEM joiners than other initial non-STEM majors (10% vs. 5%, $p = .029$).

High school STEM preparation. STEM joiners had similar high school STEM preparation compared to "leaks" in the STEM pipeline, that is, students who switched from STEM to a non-STEM field (hereafter, *STEM switch-outs*). For instance, compared to STEM switch-outs, joiners had similar SAT Mathematics scores (*M*s = 570 vs. 550, respectively, $p = .140$) and joiners were marginally more likely to have taken calculus in high school (49% vs.

38%, $p = 0.064$). We generally did not find differences in high school background between students who joined life science versus pSTEM fields (e.g., life science vs. pSTEM joiners had nonsignificantly different SAT Mathematics scores and high school STEM grades).

STEM joiners had slightly weaker high school STEM preparation compared to students who persisted in STEM from the beginning of college to graduation (hereafter, *STEM persisters*). For instance, joiners scored 40 points lower on the SAT Mathematics test ($p < .001$) and were marginally less likely to have taken high school calculus compared to STEM persisters (49% vs. 59%, $p = .056$); see Table S3 for other comparisons.

Undergraduate STEM success and course taking. STEM joiners, however, achieved similar undergraduate success compared to STEM persisters. For instance, the rate of graduating on time (i.e., within 4 years of entering college) was similar for STEM joiners and STEM persisters (49% vs. 51%, respectively, $p > .250$). Double majoring was more common among STEM joiners than persisters, though still infrequent for both groups (15% vs. 8%, respectively, $p = .015$). Grades in undergraduate STEM courses were also similar (Table S3). STEM joiners maintained interest in non-STEM fields by continuing to take non-STEM courses more frequently than STEM persisters (Figure 3). Hence, joiners could enrich STEM fields by injecting interdisciplinary training from outside fields.

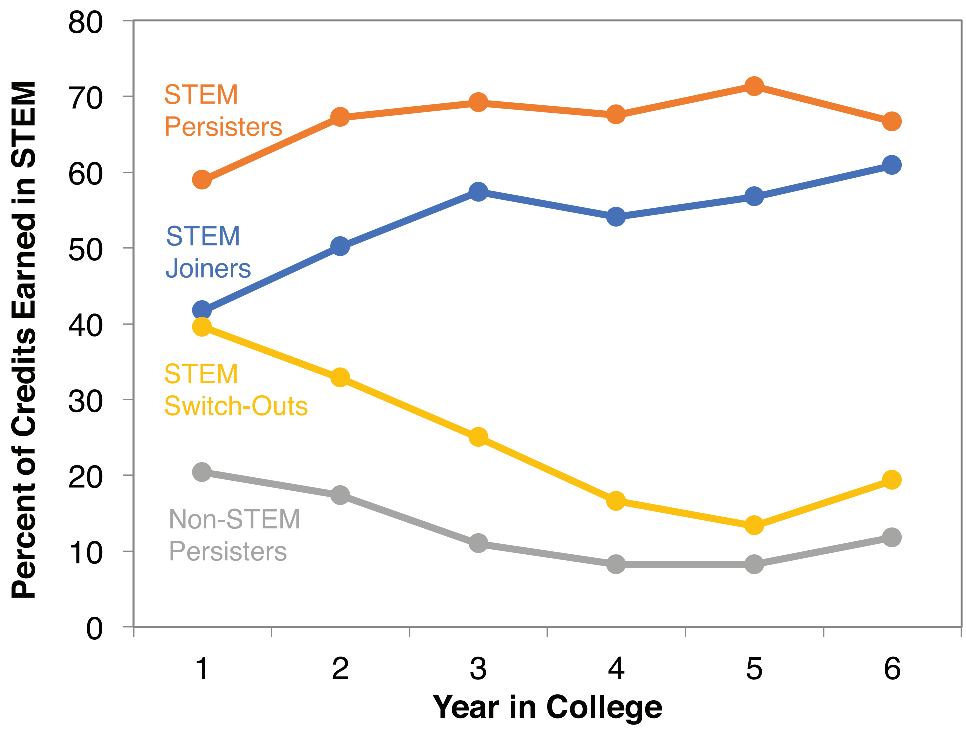

Figure 2. Average STEM course taking rates throughout college (BPS sample).

Summary. STEM joiners were nearly one-fifth of recent STEM graduates, came from many non-STEM fields, were diverse demographically, and had similar high school STEM preparation compared to "leaks" in the STEM pipeline (i.e., STEM switch-outs). Joiners had slightly weaker STEM preparation compared to STEM persisters, but achieved similar undergraduate success.

What Facilitates STEM Joining Pathways?

Next, we sought to understand the pathways for joining STEM by investigating how early-college STEM grades and course taking related to STEM joining. Consistent with expectancy-value theory (Eccles, 2011), we predicted that grades would form expectancies for success and therefore impact decisions to enter STEM. In addition, because this theory focuses on relative costs and benefits, we expected relative course performance (i.e., STEM minus non-STEM grades) would be especially informative. We also predicted that early-college STEM course taking would matter by initiating recursive processes that could change students' majors over time (Walton, 2014); these analyses used all three samples for reasons described later.

Early-college STEM grades. Contrary to predictions, early-college STEM grades *weakly* predicted STEM joining. For instance, in the BPS sample, grades in first-year STEM courses weakly predicted which initial non-STEM majors later earned STEM bachelor's degrees, $b = 0.13$, 95% CI [-0.08, 0.34]. As Figure 4b shows, similar results were found when analyzing relative course performance (i.e., STEM minus non-STEM grades). In fact, joiners' first-year grades in STEM courses were an average 0.26 grade points lower than in non-STEM course (Ms = 3.01 vs. 3.27, respectively). However, contrary to our prediction, we found that relative grades were not a predictor of STEM joining. In fact, only very weak first-year STEM grades (GPA < 2.00) or much higher grades in non-STEM courses (more than a full letter grade) predicted a slight decrease in STEM joining rates (Figure 4). Results for STEM *joining* contrasted with results for STEM *retention*: early-college STEM grades strongly predicted which initial STEM majors persisted and later earned STEM bachelor's degrees[2], $b = 0.86$, 95% CI [0.57, 1.14]. This difference in regression slopes for joining versus retention ($p < .001$) replicated in both the NLSF ($p = .003$) and Project TALENT ($p = .005$) samples. Results therefore showed that early-college grades had an asymmetric importance for STEM joining versus STEM retention pathways. Many students can join STEM even if their early-college STEM grades are at not at the top of the grades distribution.

[2] This difference might be expected if initial non-STEM major students took less challenging early-college STEM courses than initial STEM majors. However, the difference in regression slopes was also found among students who had taken calculus or a more advanced mathematics course in the first year of college, $p = .002$.

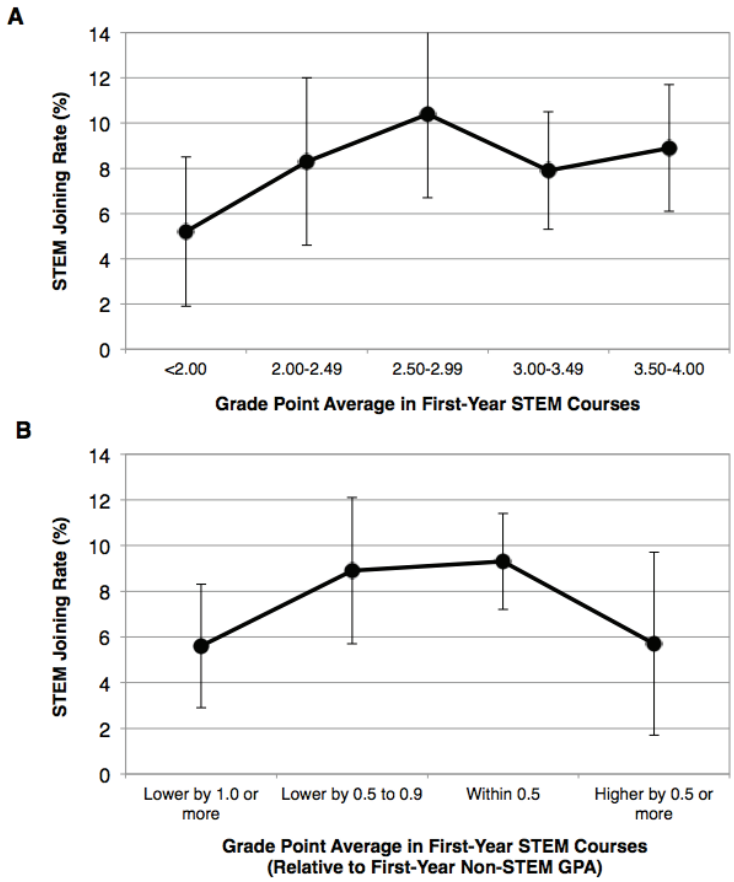

Figure 3. The relationship between first-year STEM grades and STEM joining (BPS sample). STEM grades were measured on an absolute scale (A) or relative to non-STEM grades (B). Error bars represent 95% CIs.

Early-college STEM course taking. In contrast to grades, early-college STEM *course taking* strongly predicted STEM joining. For instance, among students intending a non-STEM major, students who earned more than 30% of their first-year credits in STEM departments were 4.5 times as likely as others to later earn a STEM bachelor's degree ($p < .001$; BPS sample). This predictive effect of course taking was not significantly different for men and women. Moreover, results indicated a dose-response effect of this course taking (Figure 5A). However, students already interested in STEM could have self-selected into those courses, providing an alternate explanation for such results. Further analyses therefore investigated these course taking results by controlling for potential confounds (e.g., prior interests in STEM topics). As mentioned earlier, these analyses focused on the NLSF and Project TALENT samples because they included much more comprehensive sets of pre-college covariates than did the BPS sample.

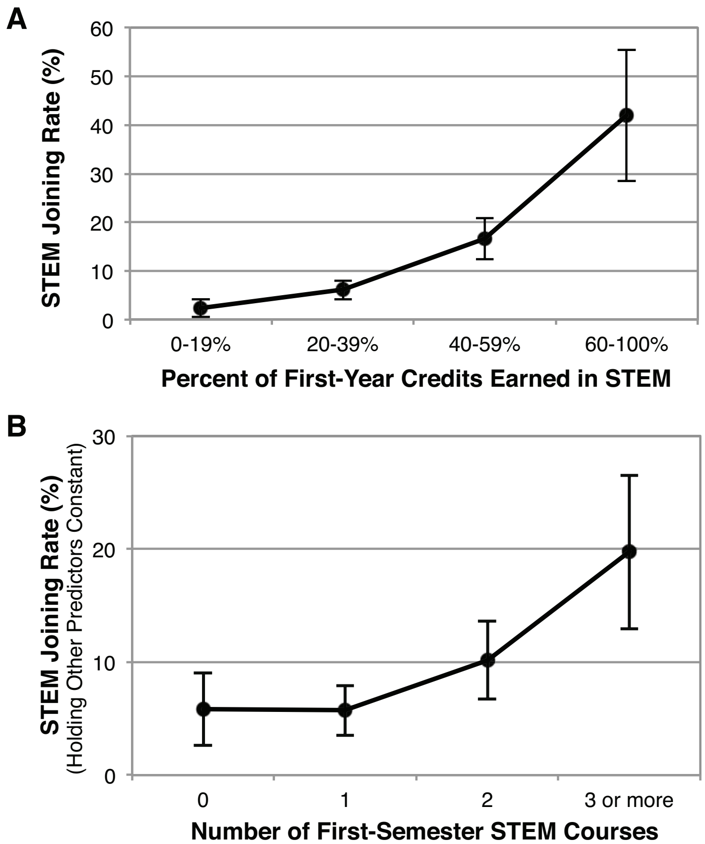

Figure 4. The dose-response relationship between early-college STEM course taking and STEM joining (NLSF sample). The first graph (A) shows this relationship descriptively by not controlling for other predictors (BPS sample). The second graph shows the estimated average marginal effects that held constant many other predictors (NLSF sample). Error bars represent 95% CIs.

Controlling for self-selection confounds. Even after controlling for many potential confounds, early-college STEM course taking predicted later earning of STEM bachelor's degrees among initial non-STEM majors. For instance, in the recent NLSF sample (Figure 5B), taking two or more first-semester STEM courses predicted a doubling of the average STEM joining rate relative to taking one such course or none (13% vs. 6%, respectively, when holding other predictors constant, $p < .001$; see Table S4 for more details). We found similar results with Project TALENT (Table S5), meaning that our results replicated across these diverse datasets that spanned four decades and levels of institutional selectivity.

Mechanistic insight into why early-college course taking matters. STEM course taking in the first semester predicted later earning of STEM bachelor's degrees years later because it predicted STEM course taking in the next two semesters (Figure 6, NLSF sample). This result was consistent with the hypothesis that early-college STEM course taking (e.g., in the first semester) creates structural pathways for joining and later persisting in STEM.

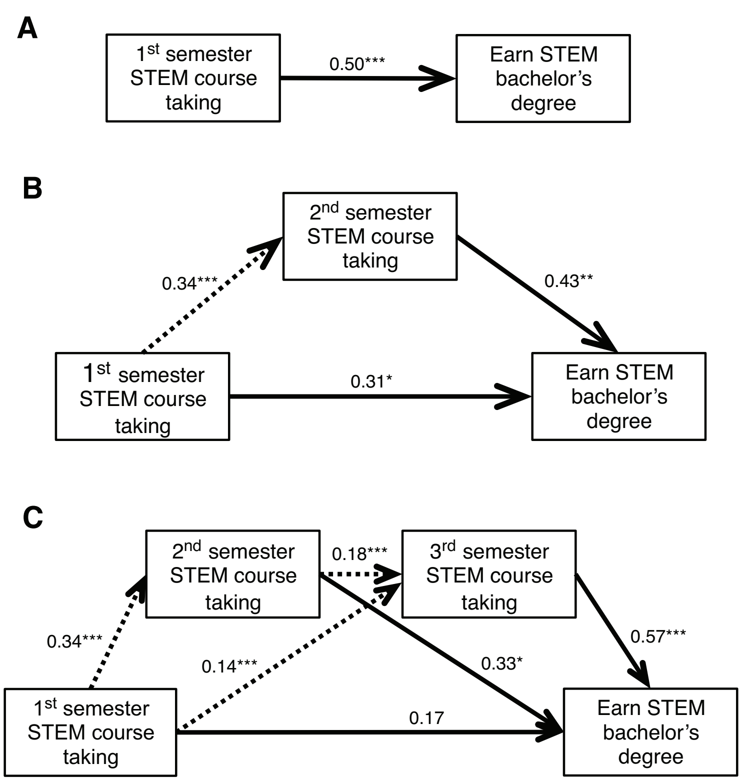

Figure 5. Mediational pathway by which first semester STEM course taking predicts later earning of STEM bachelor's degrees. In the NLSF sample, the predicted effect of first-semester STEM course-taking on later STEM joining (A) was partially mediated by second-semester STEM course-taking (B) and third-semester STEM course-taking (C). The numerical values are the unstandardized regression coefficients for STEM course-taking variables in multilevel logistic regression models of STEM joining (solid lines) or the unstandardized coefficients for course-taking variables in multilevel models of later STEM course-taking (dotted lines). All models controlled for pre-college variables and intending a pre-medicine major (see Table S1); for simplicity, regression coefficients for these other variables are not shown.
$^{***}p < .001$. $^{**}p < .01$. $^{*}p < .05$. $^{\dagger}p < .10$.

Institutional variability. Early course taking also helped explain why some institutions had particularly high STEM joining rates. For instance, one in five institutions in NLSF had a joining rate above 12% and another one in five had a rate below 6%. Multilevel analyses (Raudenbush & Bryk, 2002) found that this institutional variability related to the average early-college STEM course taking among institutions' non-STEM majors. In both the NLSF and Project TALENT samples, STEM joining rates were higher at institutions where non-STEM majors took more early college STEM courses (Table S3; Table S4). In both samples, the between-institution and within-institution relationships were similar in magnitude (see Supplemental Materials available online for information on how the variables were centered for between-institution analyses). The between-institution relationship was not statistically significant for the NLSF sample perhaps because of the small level-2 sample size (28 institutions), but was significant for the Project TALENT sample (see Tables S3 and S4 in the supplemental materials available online).

Further accounting for self-selection effects. Further analyses tested whether self-selection could have plausibly accounted for the course-taking results that controlled for potential confounds (e.g., results in Figure 5B). With correlational data, multiple methods and converging results are needed to rule out alternative explanations (Ichino et al., 2008). Four additional results, described in depth in Table S6, provided such evidence; we summarize these results below. Each result by itself had notable limitations, but using multiple methods collectively helps to balance the strengths and limitations of each approach.

Self-selection clearly explained some variance in undergraduate STEM course taking, but Results 1–3 in Table S6 collectively provided some evidence that the magnitude of this selection effect was likely small among non-STEM intenders taking STEM courses early in college. Self-selection due to observed pre-college variables was small among non-STEM intenders (Result 1). For instance, in the NLSF sample, advanced STEM course taking in high school weakly predicted STEM course taking early in the first semester of college, $r = .07$, 95% CI [.02, .13]. In contrast, these same variables were stronger predictors of STEM course taking among undecided students (Result 2). Hence, self-selection effects could be smaller among initial non-STEM majors than other students (e.g., undecided students). Moreover, compared to pre-college variables, institutional factors explained more variance in first-STEM course taking among initial non-STEM majors (Result 3). This institutional variability likely reflected in part the effects of some factors other than self-selection (e.g., such as course taking policies and availability of STEM courses for non-majors). Self-selection into early-college STEM courses was therefore likely small among undergraduates intending a non-STEM major, even if self-selection in other populations (e.g., undecided students) or other points in time (e.g., high school) could be larger (Maltese & Tai, 2011). Finally, sensitivity analyses showed that these causal estimates were robust to plausible unobserved confounders (Result 4); only hidden biases of implausibly large magnitudes would have substantially reduced causal estimates (e.g., by 50% or more).

Multiple methods and converging results therefore provided evidence against alternative explanations regarding self-selection. Each result by itself had notable limitations (e.g., results from multiple regression were limited by the range of observed variables), but the results collectively strengthened each other. For instance, simulation analyses (Result 4) provided some evidence for robustness to unobserved variables. Although covariates did not seem to account for our course-taking results, some covariates still independently predicted STEM joining. For instance, STEM joiners had more advanced pre-college STEM backgrounds compared to other students who initially intended a non-STEM major. In the BPS sample, 49% of STEM joiners took calculus in high school compared to 25% of other initial non-STEM majors ($p < .001$).

Summary. Taking STEM courses early in college robustly predicted STEM joining. STEM course taking was more important than grades. Multiple causal inference methods provided evidence that self-selection likely could not account for the course-taking results, though some caution should still be urged given the correlational nature of these data.

What Is The Impact Of Widening Joining Pathways?

This section estimates (a) the national prevalence of STEM joining and (b) the potential implications of increasing STEM joining rates using the BPS sample.

Prevalence of STEM joining. Among students who initially intended a non-STEM major in spring 2004, 6.8% later earned a STEM bachelor's degree by 2009, 95% CI [5.7%, 7.9%], excluding non-degree earners[3]. This low joining rate of ~7% meant that some other fields

[3] The STEM joining rate was 3.9%, 95% CI [3.3%, 4.5%], including non-degree earners. For further details on the effects of including non-degree earners, see this interactive website: https://d-miller.shinyapps.io/joiningSTEM/

were more effective at replenishing their supply of majors. For instance, social science and STEM fields had equally high attrition rates (Chen & Soldner, 2013) but distinctly different rates of replenishment. For every student who switched out of social science, 1.9 students entered from another discipline. This ratio is reversed in STEM fields: one student entered STEM for every 1.5 students who switched out. Plugging leaks (decreasing attrition) or widening joining pathways (increasing joining) could change this replenishment rate. But importantly, STEM fields typically have lower attrition rates than non-STEM fields (Chen & Soldner, 2013). The U.S. therefore suffers from a net loss of STEM graduates because of infrequent STEM joining, not because of an especially leaky pipeline in STEM compared to non-STEM fields.

Addressing workforce needs. A small increase in the STEM joining rate could substantially help the U.S. meet projected needs for more STEM graduates. For instance, we estimated that increasing the joining rate by just 5 percentage points would generate between 26,000 to 63,000 more STEM graduates per year, depending on the definition of potential STEM joiners (see Figure 7 for further details). Placing these numbers in context, a 2012 report to the President estimated that the U.S. needs to produce 100,000 more STEM graduates per year to match projected STEM workforce growth (PCAST, 2012). Hence, increasing joining rates by just 5 percentage points would generate between one-quarter to nearly two-thirds of these projected needs.

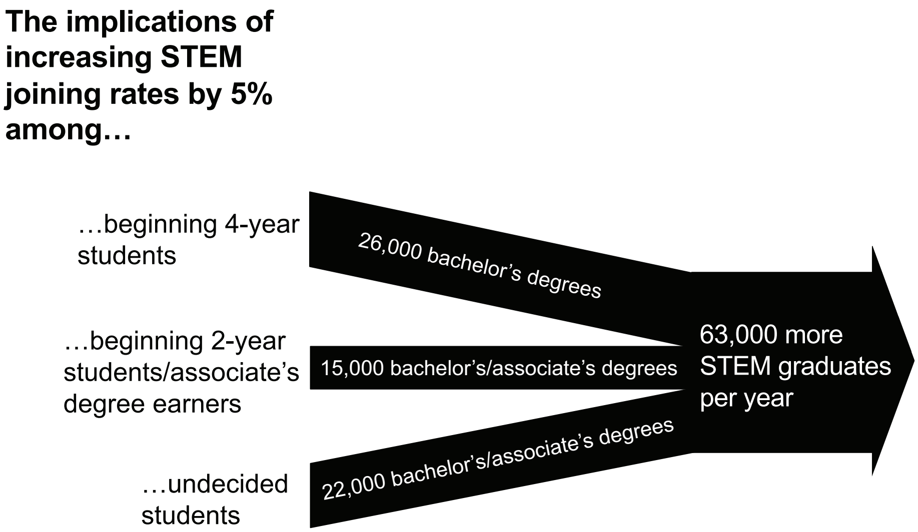

Figure 6. The effects of increasing STEM joining rates by 5% on the number of STEM graduates per year. Estimates based on BPS data (see Table S7 for more details).

Addressing gender disparities. Increasing STEM joining among women could substantially increase their representation in STEM. Men and women were equally likely to join life science from a non-STEM field (p = .190), but men were 2.8 times as likely as women (p < .001) to join other STEM fields such as physical science and engineering (pSTEM). In one hypothetical scenario, increasing women's pSTEM joining rate to match men's would generate 38% more female graduates in these male-dominated fields (Figure 8). As shown in Figure 8, equalizing gender differences in entering pSTEM from being undecided would have an even larger effect. But these initially undecided students were ambiguous as discussed earlier (see *Participant Selection Criteria*).

Gender differences in pSTEM retention were marginal. Among initial pSTEM majors who later earned bachelor's degrees, women were less slightly likely than men to persist in pSTEM (59% vs. 70%, respectively, p = .089). However, closing this persistence gap – the focus of many current efforts – would have little effect on degrees earned, as Figure 8 shows. For instance, women currently earn 25% of the U.S.'s pSTEM bachelor's degrees, and "plugging" the leaky pSTEM pipeline for female undergraduates would only increase this percentage to 27% (see also Xie & Shauman, 2003). Closing the persistence gap would have little effect because few women intend a pSTEM major (e.g., women were fourteen times as likely to intend a non-STEM than pSTEM major). The gender gap in undergraduate retention only explained 5% of the gap in earning pSTEM bachelor's degrees, whereas the gap in joining explained 19% of the gap in earning degrees. Finally, pre-college factors such as intentions in high school contribute even more to the later gap in bachelor's degrees (Ceci et al., 2014; Legewie & DiPrete, 2014).

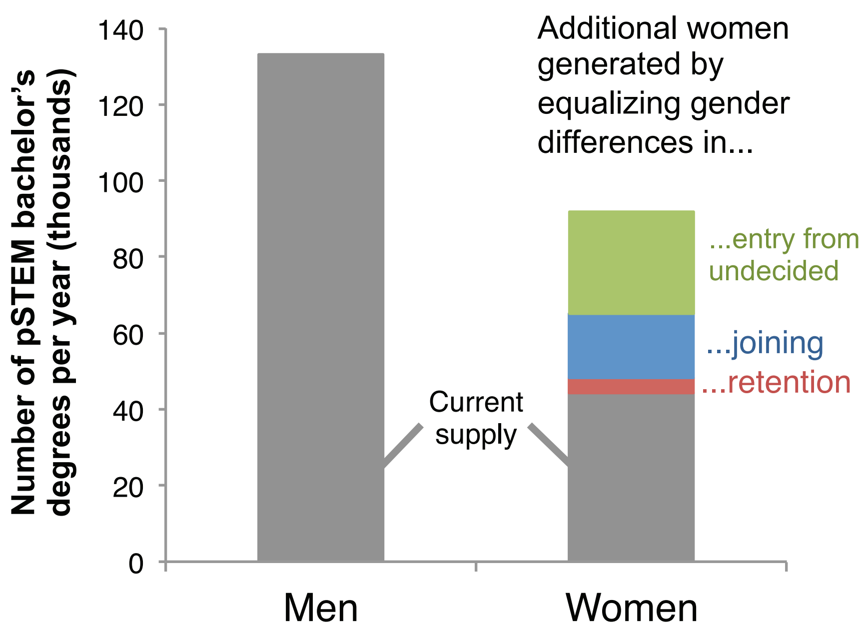

Figure 7. The effect of equalizing gender differences in undergraduate STEM transitions on the number of female STEM bachelor's degree earners per year (excluding life science). Estimates based on BPS data (see Table S7 for more details).

Summary. Increasing STEM joining rates could potentially help address STEM workforce needs and gender disparities in STEM.

Discussion

Aligning with broader literature on emerging adulthood (Arnett, 2000), results showed that college students frequently explore academic identities, regardless of initial major. Roughly one-fifth of STEM graduates came from initial non-STEM majors, demonstrating the multifaceted pathways for joining STEM. These joiners started with similar high school STEM preparation compared to "leaks" in the pipeline (i.e., STEM switch-outs), but then achieved similar undergraduate success compared to STEM persisters (e.g., on-time graduation rates were equal). Joiners came from diverse non-STEM fields and took more non-STEM courses than persisters did throughout college. Future research should investigate whether this interdisciplinary training could be an asset for jobs at the intersection of technology and society. For instance, innovation at companies like Facebook and Google not only requires technical expertise but also an understanding of what people want and need (PCAST, 2012). These results collectively demonstrate the multiple pathways, not a singular pipeline, to STEM success.

The pipeline metaphor focuses important attention on STEM retention, but unfortunately away from more comprehensive strategies for broadening participation in STEM. For instance, compared to "plugging the leaky pipeline," closing the gender gap in undergraduate joining would more potently increase women's representation in pSTEM. Yet many psychological studies and theories often overlook this large, important source of potential female STEM majors. Stereotype threat research, for instance, often uses pre-study surveys to exclude undergraduates who do not already strongly identify with mathematics (e.g., Murphy, Steele, & Gross, 2007).

Results showed how studying STEM joiners can modify conclusions about what psychological factors are most important to pursuing STEM. Much prior research has pointed to the importance of performance in introductory STEM courses, in particular. The harsh grading standards and demanding work in those courses are thought to create self-doubt, driving students away from STEM (see Ceci et al., 2014 for a review). This so-called "fear of the B-," however, seems to only apply to STEM retention, not joining. All three samples spanning diverse institutions consistently showed that STEM grades weakly predicted joining, despite strongly predicting retention. In fact, STEM joiners earned an average of 0.26 grade points lower in their first-year STEM courses, compared to non-STEM courses. These results suggest *asymmetries* in the psychological processes that facilitate STEM joining versus retention. Analogous to broader literature on social identity threat (e.g., Murphy et al., 2007), we suspect that low STEM course performance may less threatening to those that do not already identify with the domain (i.e., initial non-STEM majors). Nevertheless, joiners may find STEM subjects intrinsically interesting or important to future career prospects, consistent with expectancy-value theory (Eccles, 2011). Future research should test such hypotheses.

In contrast to grades, simply taking STEM courses early in college was crucial for joining STEM years later. Institutional differences in such course taking even helped to explain why some institutions had particularly high joining rates. These results remained after controlling for several theoretically relevant covariates and replicated across three nationally representative sample spanning four decades and diverse institutions. Multiple methods for causal inference helped to rule out self-selection effects, though experimental methods would more definitely rule out alternative explanations. Results indicated one potential mechanism for these early-college effects: they initiative recursive feedback loops that unfold throughout college (Walton, 2014). In particular, first-semester STEM course taking mattered because it predicted course taking in the next two semesters. This subsequent course taking mediated the predicted effect of first-semester courses on STEM joining. In other words, taking STEM courses early in college could allow students to start travelling pathways that later lead to earning STEM degrees. Not all students will choose to travel such pathways, but taking relevant courses in a timely manner could create opportunities to do so. These results emphasize that that successfully joining STEM during college requires persistence in paths that later lead to STEM degrees. Studying joiners could therefore help provide a more multifaceted understanding of STEM persistence (cf. Graham et al., 2013; Watkins & Mazur, 2013).

Educators and policy makers can build on our findings to develop cost-effective actions for broadening participation in STEM. In particular, redesigning course requirements could widen these joining pathways. For instance, the University of Notre Dame requires that all students take two mathematics courses in the first year and two science courses by the end of the second year (University of Notre Dame, 2012). Several other institutions have STEM course requirements for all majors, but often do not specify *when* these courses must be taken (American Council of Trustees and Alumni, 2014). To facilitate STEM joining, university administrators could consider redesigning graduation requirements as early-college requirements. Redesigning the timing of such requirements could be done even without increasing the total number of undergraduate STEM requirements. Future research can build on our findings to identify specific course-taking policies that would most effectively take advantage of undergraduate entry points into STEM. Educators could also inexpensively redesign these introductory courses to make them more inviting to diverse students. Highlighting how scientific research involves working with and helping others could particularly encourage women to join STEM fields, for instance (Diekman et al., 2011; Valla & Ceci, 2014). However, we remain neutral whether the gender gap in STEM joining can be fully closed in reality, especially if factors such as different interests contribute to that gap (Ceci et al., 2014).

In conclusion, students who begin college in a non-STEM field can bring valuable interdisciplinary perspectives into STEM. However, the story of a singular STEM pipeline has directed thought away from leveraging this currently untapped pool of diverse talent. Studying the multiple pathways to STEM can help enrich understanding of when psychological factors such as grades and expectations contribute to students' academic decisions. In contrast to grades, simply taking STEM courses early in college was crucial to STEM joining, even after controlling for many theoretically relevant covariates. Policy makers and educators can therefore likely widen these joining pathways by increasing the quantity and quality of STEM courses that non-STEM majors take early in college.

Chapter 3: Pathways for Joining the Computing and Engineering Workforce

As discussed in the previous chapter, postsecondary educators and policymakers have called for an increase in the national supply of STEM college graduates to help meet projected STEM workforce growth. Part of this demand comes from the assumption that STEM jobs require STEM degrees. For instance, the National Science Board (2018) noted that STEM jobs "are generally assumed to require at least a bachelor's degree level of education in [a science or engineering] field" (p. 3-12). However, contrary to these common assumptions, nationally representative data show that STEM employers often hire workers without STEM degrees. For instance, in 2015, approximately 1.2 million STEM workers were college graduates without a bachelor's degree or higher in any STEM field (National Survey of College Graduates, 2015). They were one-fifth (20%) of the total college-educated STEM workforce in 2015, excluding social science[4]. Many workers with non-STEM degrees therefore already pursue STEM careers, indicating the openness of pathways for joining STEM even after graduating college.

Despite lacking formal STEM degrees, these workers may have gained technical skills through informal learning experiences such as gradually taking on technical job tasks (e.g., database administration) in non-STEM jobs (e.g., marketing analysts). In contrast, research publications (e.g., Cannady et al., 2014) and policy reports (e.g., PCAST, 2012) that discuss STEM careers typically emphasize the importance of formal postsecondary degrees but overlook informal learning experiences. Consequently, employers are left with little guidance from scholarly research on how to best train and manage the over one million college-educated STEM workers without STEM degrees. To address this critical gap in prior literature, I analyzed the

[4] These statistics were based on NSF's classification scheme of STEM jobs (National Science Board, 2018), which was modified for this study to exclude social science jobs for consistency with Study 1.

2015 National Survey of College Graduates (NSCG) to describe the pathways for joining STEM careers among college graduates without STEM degrees. The NSCG was a large national probability sample (*n* = 91,000) representative of college graduates who lived in the United States in 2015. In addition to characterizing joiners' jobs (e.g., how joiners used their non-STEM training), this study examined how workers' goals to benefit society related to decisions to pursue STEM employment.

My analyses focused on two large employment fields: computing and engineering. These two subfields were the largest ones under the broad category of STEM employment, collectively accounting for 78% of the college-educated STEM workforce and 88% of the college-educated pSTEM workforce in 2015 (National Survey of College Graduates, 2015). These fields were also important to study because they were more gender imbalanced than other STEM fields such as chemistry and mathematics (Cheryan, Ziegler, Montoya, & Jiang, 2017). Disaggregating results by specific STEM fields was appropriate because the educational requirements for entry-level jobs varied greatly by field. For instance, STEM joiners were far more common in the computer science than engineering workforce, as described in more detail later in this chapter. Analyzing STEM jobs as one aggregate category would have instead masked this variation by field.

This study considered how joiners used their non-STEM training in these fields. For instance, software developers often work in large teams to develop products that will be later marketed to consumers (Highsmith & Cockburn, 2001). Non-technical skills such as communication and organizational skills might therefore help these developers collaborate in teams and design profitable products. Consistent with this reasoning, in a nationally representative sample in 2016, software applications developers (*n* = 30) rated the following interpersonal and organizational activities as important to job performance: communicating with supervisors, peers, or subordinates; organizing, planning, and prioritizing work; coordinating the work and activities of others; developing objectives and strategies; establishing and maintaining interpersonal relationships (National Center for O*NET Development, 2018). The importance of all these activities was rated above the midpoint (i.e., "important") on a 1-5 scale ranging from "not important" to "extremely important."

More broadly, business-related skills are especially relevant because for-profit businesses employ most STEM workers (National Science Board, 2018). For instance, in 2015, 78% of computer scientists and 77% of engineers worked in the for-profit business/industry sector (National Survey of College Graduates, 2015). The need for non-technical skills may therefore be one reason why some STEM employers often hire workers with non-STEM degrees. The NSCG's questions were well suited to study how joiners used their non-STEM degrees. For instance, one set of questions asked about time spent on work activities such as "sales, purchasing, marketing" and "supervising people or projects" that might have leveraged joiners' non-STEM skills. Respondents also rated how related their job was to their highest degree, allowing me to examine joiners' perceptions of using their non-STEM training in computing and engineering careers.

Furthermore, the NSCG survey enabled me to study psychological factors that may have influenced employment decisions. For instance, one question asked respondents to rate the importance of "contribution to society" when "thinking about a job." Workers who valued benefiting society may have avoided computing and engineering if they saw those fields as lacking opportunities to help others (Diekman, Steinberg, Brown, Belanger, & Clark, 2017). I therefore investigated how the goal of benefitting society related to employment outcomes. Prior research has studied how communal goals to work with or help others has predicted career interests in STEM (for a review, see Boucher, Fuesting, Diekman, & Murphy, 2017). This research, however, has often been limited to convenience samples of psychology undergraduates and has not studied actual employment outcomes. To my knowledge, my study is the first to use nationally representative data to examine how communal goal endorsement relates to STEM employment.

Theoretical Framework

Goal Congruity Perspective

Americans have often perceived computer science and engineering as fields that do not provide opportunities to work with or help others. In one U.S. study of students in an introductory psychology course (*n* = 333), undergraduates perceived STEM jobs such as computer scientist and mechanical engineer to fulfill communal goals less than other jobs (Diekman, Brown, Johnston, & Clark, 2010). STEM jobs were viewed as lacking communal opportunities even when compared to male-stereotypic, non-STEM jobs such as lawyer and architect. These findings on perceived communal goal affordances have been replicated and extended in later research (Diekman et al., 2011; Matskewich & Cheryan, 2016; see Boucher et al., 2017 for a review). Computer scientists and engineers instead have often been seen as "nerdy, socially awkward men who love science fiction and video games" (Boucher et al., 2017, p. 2).

Given these perceptions, individuals who value communal goals such as helping others may opt out of computer science and engineering careers. Diekman et al.'s (2017) goal congruity perspective hypothesizes that communally oriented individuals may avoid these fields because they anticipate goal incongruity – a mismatch between their personal goals and perceived career opportunities. These considerations are important for both women and men because both sexes share the basic psychological needs for relatedness and belonging (Diekman et al., 2017, p. 147). However, women also typically endorse communal goals more strongly than men (e.g., Konrad, Ritchie, Lieb, & Corrigall, 2000; Su, Rounds, & Armstrong, 2009). This sex difference in communal goal endorsement may help partly explain the low numbers of women in computing and engineering careers (Cortes & Pan, 2017; Su & Rounds, 2015).

Joining Versus Persistence Pathways

Based on this goal congruity perspective, communal goal endorsement should negatively predict working in computer science or engineering among college graduates without STEM degrees. Such workers have limited exposure during formal education to the type of work opportunities available in computing and engineering. Consequently, they may base their career decisions on culturally shared stereotypes that portray those fields as not affording communal opportunities. Workers without STEM degrees may therefore avoid computing and engineering even if those fields offer ways to help others (i.e., even if those stereotypes are inaccurate).

Communal goal endorsement might also negatively predict persistence (e.g., engineering graduates working in engineering) if communally oriented workers struggle to find ways to help others in computing and engineering careers. For instance, among workers sampled for the nationally representative O*NET database, computer scientists and engineers rated caring for others as only somewhat important to job performance (National Center for O*NET Development, 2018). The average importance rating for "assisting and caring for others"[5] was 2.07 for computer scientists (n = 525) and 2.17 for engineers (n = 1,014) on a 1-5 scale ranging from "not important" to "extremely important"; the value of 2 corresponded to "somewhat important." In contrast, the average rating for all U.S. workers was 2.86 (n = 26,829). Other fields such as medicine or teaching therefore might be more attractive to communally oriented computer science and engineering graduates (Croft, Schmader, & Block, 2015).

[5]The survey defined assisting and caring for others as "providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients."

Nevertheless, several factors might attenuate the effect of communal goals on persistence from STEM degrees to STEM jobs. For instance, despite stereotypes suggesting otherwise, computing and engineering jobs could offer many ways to help others (e.g., using machine learning to improve non-profits' organizational practices, mentoring colleagues). The O*NET database finds that computer scientists and engineers must often give expert advice to others as part of their job. These workers rated the following activities as important to job performance (i.e., above the midpoint on a 1-5 scale): providing consultation and advice to others; training and teaching others; interpreting the meaning of information to others (National Center for O*NET Development, 2018). Communally oriented computer scientists and engineers therefore may not experience goal incongruity given these varied ways to help others in their jobs.

Furthermore, even if workers experience incongruity at first, they might align their jobs with their goals over time through role reconstruction (e.g., negotiate different working conditions) or role reconstrual (e.g., mentally reframe the nature of their work), as Diekman et al.'s (2017) goal congruity perspective would predict. For instance, even if an engineer struggles to find ways to directly help others, that person could think about how engineering benefits society more distally and contributes to economic growth; this thought process is an example of role reconstrual (Diekman et al., 2017, p. 145). Considering why STEM workers engage in tasks such as conducting scientific experiments could also help workers think of their jobs' broader implications to society. Consistent with this hypothesis, one recent study (n = 193) found that prompting participants to explain why, rather than how, scientists conduct experiments increased participants' tendency to say that science careers could fulfill communal goals (d = 0.42; p = .003; Steinberg & Diekman, 2018). The processes of role reconstruction and reconstrual might buffer against the effects of experienced incongruity, leading communally oriented STEM workers to persist in their careers.

Consistent with these hypotheses, most computer scientists (76%) and engineers (86%) in 2015 said they were somewhat or very satisfied with their job's contribution to society (National Survey of College Graduates, 2015). This satisfaction could reflect both the communal opportunities available in computing and engineering jobs as well as role reconstruction and reconstrual processes.

This Study's Hypotheses

I therefore predicted that wanting to help others would negatively relate to joining the computer science and engineering workforce (i.e., working in those fields without a STEM degree). In contrast, I had less strong predictions for whether communal goal endorsement would relate to persistence (e.g., working in computer science with a computer science degree), for reasons described earlier. One exploratory hypothesis was that communal goal endorsement would relate to persistence, but to a lesser extent than for joining.

I investigated these hypotheses using workers' valuing of contributing to society when thinking about a job. Workers also rated the importance of other career considerations such as "intellectual challenge" and "opportunities for advancement." These other responses helped me identify the unique link between communal goal endorsement and career decisions, controlling for other factors such as agentic goals (i.e., wanting to promote interests of the self, rather than others).

In addition, I estimated the national impact of these goal pursuit processes on the U.S. STEM workforce. Using methods described in the next section, I quantified how many more college graduates would have been computer scientists and engineers if communally oriented workers had pursued those fields at the same rate as other workers. These impact analyses were possible because the NSCG was a large national probability sample, permitting precise estimation of population sizes and nationally representative employment rates.

Method

The 2015 NSCG survey was a stratified national probability sample (n = 91,000) whose target population was individuals who lived in the United States in February 2015, had at least a bachelor's degree, were under 76 years of age, and were not institutionalized (see https://nsf.gov/statistics/srvygrads/). Survey questions asked about respondents' employment situation, principal employer, principal job, certifications and licenses, past employment, other work-related experiences, educational experiences, and demographic information (see https://www.nsf.gov/statistics/srvygrads/surveys/srvygrads-returnrespond2015.pdf). Individuals with STEM jobs and STEM degrees were oversampled.

Jobs and Degrees

Occupational field. The NSCG survey asked respondents to "choose the code that best describes the principal job you held during the week of February 1, 2015" and provided a list of specific job categories such as "database administrators" and "mechanical engineers." Postsecondary STEM teachers and professors were considered STEM workers, consistent with NSF's classification scheme of STEM jobs (National Science Board, 2018). However, precollege STEM teachers, managers (e.g., "engineering managers"), and technologists and technicians (e.g., "surveying and mapping technicians") were excluded, meaning that NSF's classification scheme provided a somewhat conservative definition of STEM workers. This study focused on workers who NSF classified as computer and information scientists (e.g., "information security analysts" and "computer support specialists") and engineers (e.g., "mechanical engineers").

Educational degrees. The survey asked respondents to list information about all degrees that respondents earned at the bachelor's level or higher, but did not ask about associate's degrees. Respondents were considered potential joiners if they did not earn any degree at the bachelor's level or higher whose major field of study or second major was a pSTEM field.

Other Survey Questions

Work activities. The survey asked, "which of the following work activities occupied at least 10 percent of your time during a typical work week on [your principal job]?" and listed activities such as "computer programming, systems or applications development" and "design of equipment, processes, structures, models."

Job's relationship to highest degree. One question asked, "to what extent was your work on your principal job related to your highest degree?" and provided three options: (1) closely related, (2) somewhat related, and (3) not related.

Career goals. The survey also asked, "when thinking about a job, how important is each of the following factors to you?" and listed nine goals: (1) salary, (2) benefits, (3) job security, (4) job location, (5) opportunities for advancement, (6) intellectual challenge, (7) level of responsibility, (8) degree of independence, and (9) contribution to society. Respondents were given four response options for each goal ranging from "not important at all" to "very important."

Definition of Joiners and Persisters

For this study, the term "joiners" referred to college-educated computer science or engineering workers with no bachelor's degree or higher in a pSTEM field. This definition included life and social science graduates because those fields were distantly related to the more math-intensive fields of computing and engineering (Ceci et al., 2014); analysis of the NSCG's question about job's relationship to highest degree also supported this conclusion. In contrast, physical science and mathematics graduates were excluded because those fields were more closely related to computing and engineering. Joiners were a larger percentage of the college-educated computer science workforce (32%) than engineering workforce (9%), indicating the career pathways were more open for joining computing than engineering. Persisters were defined as computer science or engineering workers who earned their highest degree in the same field (e.g., computer science persisters earned their highest degree in computer science). Based on this definition, persisters were 78% of engineers and 40% of computer scientists.

These definitions of joiners and persisters excluded other categories of workers (e.g., computer scientists with mathematics degrees) who were 28% of computer scientists and 12% of engineers. For instance, 19% of computer scientists earned their highest degree in engineering. These more ambiguous cases were excluded because those workers switched fields, but their highest degree was still highly related to their job. In addition, these definitions were based on degrees earned at the bachelor's level or higher, though some workers also had professional certifications and licenses. For instance, 18% of computer science joiners and 27% of engineering joiners had a formal professional certification or a state or industry license to work in their job; these percentages were 14% for computer science persisters and 25% for engineering persisters.

Analytic Strategy

NSF-created inverse probability survey weights were used in all analyses to adjust for unequal sampling probabilities and nonresponse bias. Consistent with the previous chapter, the analyses were organized around three guiding research questions about STEM joining pathways:

- **Descriptive:** Who are computer science and engineering joiners in terms of their work activities, job's relationship to highest degree, and detailed job category? Joiners were compared to persisters (e.g., engineering graduates working in engineering) and other workers on these dimensions for descriptive comparison.

- **Correlational:** What predicts STEM joining? More specifically, does communal goal endorsement relate to employment in computing or engineering among workers without STEM degrees? Does this relation remain after controlling for other factors such as agentic goals and does it also emerge for predicting persistence?
- **National impact:** What is the potential national impact of widening joining pathways? For instance, how many more computer scientists and engineers would potentially come from communally oriented workers joining those fields at the same rate as other workers?

Descriptive analyses used simple percentages to compare joiners to other workers. Correlational analyses used logistic regression to examine how communal goal endorsement related to the log odds of working in computing or engineering versus other careers. Multivariable logistic regression models estimated the unique link between communal goal endorsement and employment by controlling for eight other career goals (see previous subsection) and demographics.

Demographic variables included in multivariable models were sex (0 = female; 1 = male), marital status (0 = not married; 1 = married), family status (0 = not living with children; 1 = living with children), parents' highest level of education (0 = no college-educated parent; 1 = at least one parent with a bachelor's degree or higher), U.S. citizenship (0 = non-citizen; 1 = U.S. citizen), number of years since earning respondents' highest degree, highest degree type (three dummy codes to distinguish bachelor's, master's, professional, and doctorate degrees), and race/ethnicity (three dummy codes for being African American, Asian, or Hispanic).

Lastly, impact analyses considered different hypothetical scenarios such as communally oriented workers joining computing and engineering at the same rate as other workers. These analyses used survey weight sums to estimate the size of different populations such as communally oriented U.S. workers without pSTEM degrees. For the NSCG, a respondent's survey weight was the approximate number of people who that respondent represented in the population (see https://nsf.gov/statistics/srvygrads/). Hence, sums of these survey weights provided estimates of population sizes (e.g., NSF has often estimated the size of the STEM workforce using NSCG survey weight sums; National Science Board, 2018, Chapter 3).

Results

Who Are STEM Joiners?

My first analytic goal was to understand who joiners are by characterizing their work activities, job's relationship to highest degree, and detailed job category. These descriptive analyses used persisters and other workers as comparison populations.

Work activities. Figure 8 compares the work activities of computer science and engineering joiners (left and middle red bars) to computer science persisters (blue bars), engineering persisters (green bars), and other workers with no pSTEM degree (right red bars).

One central finding was that joiners spent somewhat less time than persisters on the two key STEM activities of computer applications and design (see the top two panels in Figure 8). For instance, 63% of computer science joiners spent at least 10% of their time during a typical work week on "computer programming, systems or applications development," compared to 87% of computer science persisters. Similarly, 51% of engineering joiners spent at least 10% of their work time on "design of equipment, processes, structures, models," compared to 70% of engineering persisters. Nevertheless, as also shown in Figure 8, joiners spent far more time on these STEM activities than their counterparts with no pSTEM degree not working in computer science or engineering (compare to right red bars).

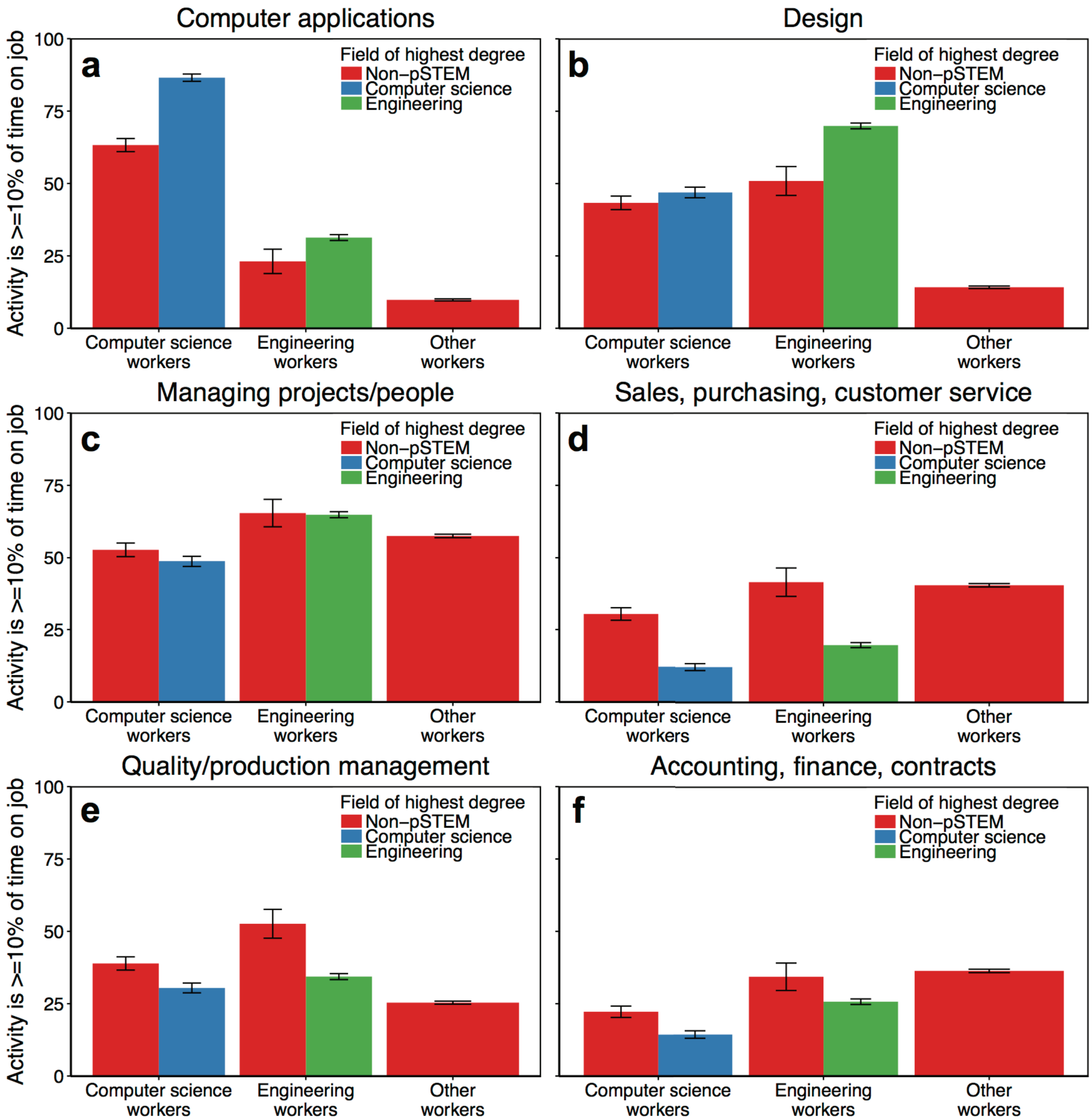

Figure 8. The percentage of workers who spend at least 10 percent of their time during a typical work week on different activities. These activities include (a) "computer programming, systems or applications development," (b) "design of equipment, processes, structures, models", (c) "managing or supervising people or projects," (d) "sales, purchasing, marketing, customer

service, public relations," (e) "quality or productivity management," and (f) "accounting, finance, contracts." Error bars represent standard errors.

Another key finding from Figure 10 was that joiners spent somewhat more time than persisters on non-STEM activities such as management and marketing (see panels c-f). For instance, 30% of computer science joiners spent at least 10% of their time during a typical work week on "sales, purchasing, marketing, customer service, public relations," compared to 12% of computer science persisters (panel d). Joiners also spent substantially more time than persisters on "quality or productivity management" and "accounting, finance, contracts" activities (panels e and f).

Job's relationship to highest degree. One potential explanation for the work activity results was that joiners selected work activities that leveraged their formal non-pSTEM training. Consistent with this hypothesis, most computer science joiners (63%) and engineering joiners (85%) said their work on their principal job was "closely related" or "somewhat related" to their highest degree, even though that degree was in a non-pSTEM field. This percentage was especially high among computer science joiners (84%) and engineering joiners (91%) whose highest degree was in business; this result suggests that many joiners used business skills such as team management and marketing in computing and engineering jobs. Such business skills would be relevant for the for-profit technology and engineering companies that commonly employ computer scientists and engineers. Like persisters, most joiners in computer science (71%) and engineering (81%) worked in the for-profit business/industry sector.

Detailed job category. Table 1 shows the detailed job categories of computer science joiners and persisters which provided another way to characterize joiners' jobs. Compared to persisters, joiners worked less often in the programming-intensive jobs of computer software engineer and software developer, consistent with joiners spending less time on programming activities (Figure 10a). Joiners instead worked more often in the less programming-intensive jobs of computer support specialist and database administrator. However, rates of working in several other roles such as computer system analyst and network systems administrator did not significantly differ between joiners and persisters.

Table 1

Detailed Job Category of Computer Science Persisters and Joiners

	Persisters		Joiners		
Computing job category	*M* (%)	*SE*	*M* (%)	*SE*	*p*
Computer engineers - software	24.8	1.6	5.7	1.1	<.001
Software developers - applications and systems software	24.7	1.6	14.2	1.6	<.001
Computer system analysts	11.8	1.2	13.6	1.6	0.362
OTHER computer and information science occupations	7.8	1.0	20.8	1.9	<.001
Network and computer systems administrators	7.2	1.0	7.1	1.2	0.936
Computer support specialists	7.0	0.9	13.1	1.6	<.001
Web developers	4.6	0.8	11.2	1.5	<.001
Database administrators	3.3	0.7	9.4	1.4	<.001
Information security analysts	3.1	0.6	1.8	0.6	0.161
Computer & information scientists, research	2.6	0.6	1.4	0.6	0.137
Computer network architect	2.2	0.5	0.9	0.4	0.046
Postsecondary Teachers: Computer Science	0.8	0.3	0.8	0.4	0.999

Table 2 compares the detailed job categories of engineering joiners versus persisters based on their detailed job category. Joiners were less likely than persisters to work as mechanical or electrical engineers, but more likely to work as sales engineers or other engineers. These findings align with the work activity results presented earlier; for instance, joiners working more often as sale engineers aligns with them spending more time than persisters on “sales, purchasing, marketing, customer service, public relations.” Rates of working in several other roles such as civil and industrial engineer did not significantly differ between joiners and persisters.

Table 2

Detailed Job Category of Engineering Persisters and Joiners

	Persisters		Joiners		
Engineering job category	*M* (%)	*SE*	*M* (%)	*SE*	*p*
Mechanical engineers	21.8	7.3	7.0	2.5	0.055
Electrical and electronics engineers	18.9	6.9	3.7	1.9	0.035
Civil, including architectural/sanitary engineers	15.4	6.4	14.3	3.5	0.883
OTHER engineers	8.6	5.0	23.6	4.2	0.021
Aeronautical/aerospace/astronautical engineers	5.6	4.1	2.0	1.4	0.399
Chemical engineers	5.2	3.9	0.9	0.9	0.286
Industrial engineers	4.2	3.6	9.7	3.0	0.238
Postsecondary Teachers: Engineering	3.5	3.3	1.2	1.1	0.494
Sales engineers	3.5	3.3	15.0	3.6	0.018
Environmental engineers	3.0	3.0	7.5	2.6	0.269
Computer engineer - hardware	2.9	3.0	8.3	2.8	0.188
Materials and metallurgical engineers	1.9	2.4	1.5	1.2	0.890
Nuclear engineers	1.7	2.3	0.8	0.9	0.723
Bioengineers or biomedical engineers	1.5	2.2	1.4	1.2	0.979
Petroleum engineers	0.9	1.7	1.4	1.2	0.813
Marine engineers and naval architects	0.6	1.4	0.8	0.9	0.929
Agricultural engineers	0.4	1.1	0.7	0.8	0.807
Mining and geological engineers	0.3	1.0	0.3	0.5	0.967

Summary. In summary, compared to persisters, joiners spent somewhat less time on STEM work activities such as applications development and more time on non-STEM work activities such as marketing and finance. These differences may reflect that joiners choose work activities that leveraged their non-STEM educational training. Joiners with business degrees were especially likely to say that their job was somewhat or closely related to their educational training. In other words, business skills may be especially relevant to success in computing and engineering, consistent with the goals of for-profit technology and engineering companies. The differences in work activities were also reflected in the detailed job categories.

What Predicts STEM Joining?

Next, I sought to study how communal goal endorsement related to pathways for joining computing and engineering among workers without pSTEM degrees ($n = 44{,}864$). These workers generally cared about the societal implications of their work; most rated “contribution to society” as somewhat important (36%) or very important (55%) when thinking about a job. Only a minority said that contributing to society was somewhat unimportant (7%) or not at all important (2%). Hence, most of these workers (91%) were communally oriented, which I defined as selecting the somewhat or very important option; this definition was used throughout this study's analyses.

Computer science joining. As expected, communal goal endorsement negatively related to joining the computing workforce (see Figure 9). For instance, communally oriented workers (i.e., rated contribution to society as somewhat or very important) were 52% less likely to join the computing workforce than other workers without pSTEM degrees (2.2% vs. 4.6%; $p < .0001$). The extent of goal endorsement even mattered among communally oriented workers. Workers who rated contribution to society as "very important" joined the computing workforce 60% less often than workers who selected the "somewhat" important option (1.4% vs. 3.5%; $p < .0001$).

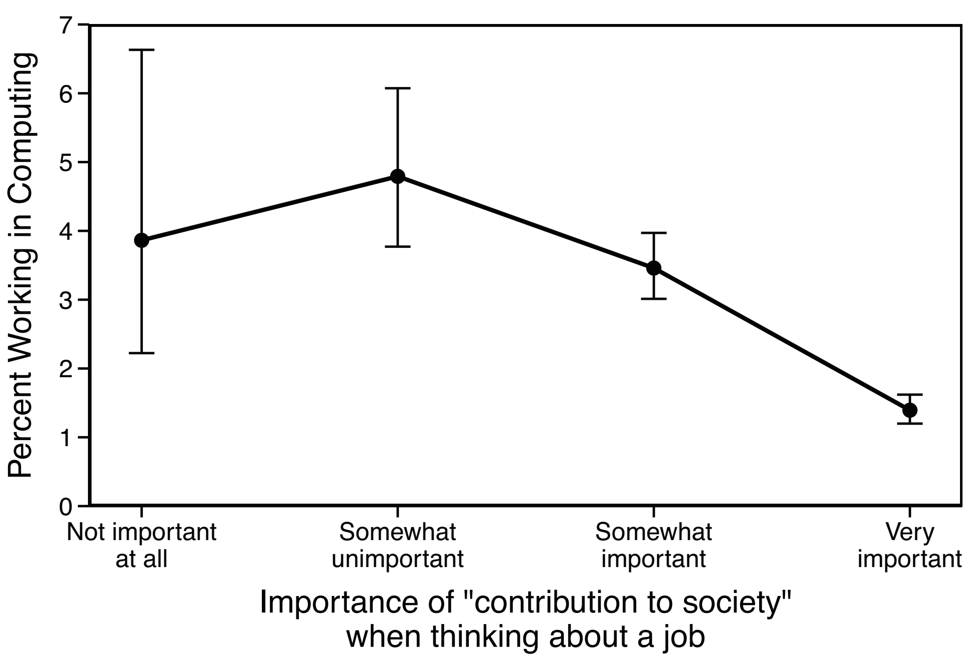

Figure 9. Computer science joining rates by communal goal endorsement among workers without pSTEM degrees (n = 44,864). Error bars represent 95% confidence intervals.

Engineering joining. Workers without pSTEM degrees were much more likely to join the computing than engineering workforce (2.4% vs. 0.4%, on average). Nevertheless, communal goal endorsement also negatively related to joining engineering (see Figure 10). Among workers without pSTEM degrees, those who rated contribution to society as "very important" joined engineering 64% less often than workers who selected the "somewhat" important option (0.2% vs. 0.7%; $p < .0001$). As a combined group, these communally oriented workers were collectively 43% less likely to join the engineering workforce than other workers without pSTEM degrees (0.4% vs. 0.7%), though that difference was not statistically significant ($p = .060$).

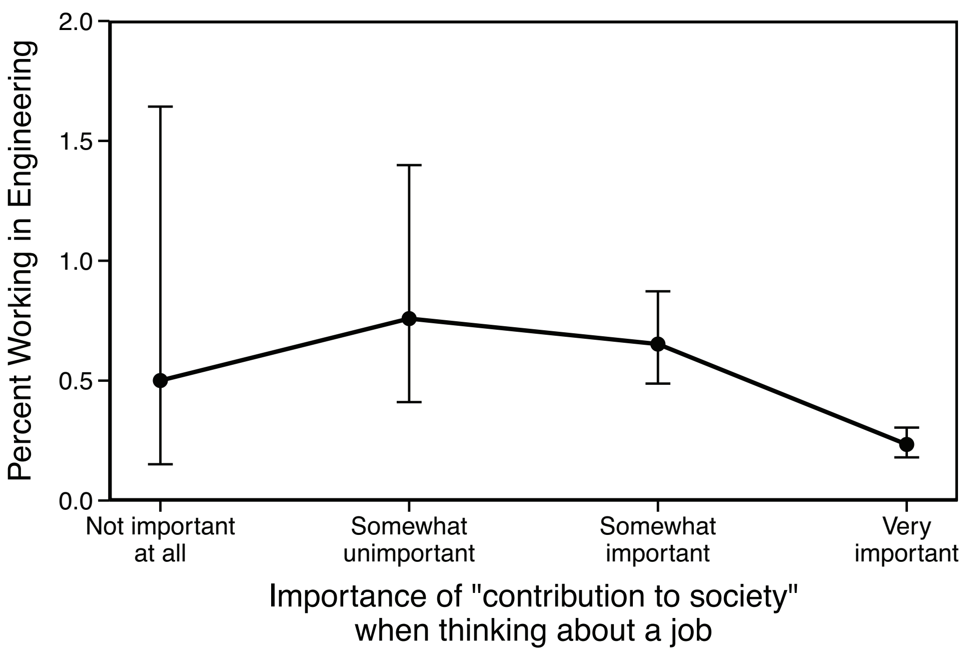

Figure 10. Engineering joining rates by communal goal endorsement among workers without pSTEM degrees (n = 44,864). Error bars represent 95% confidence intervals.

Regression models. Logistic regression analyses further extended these results. For instance, in simple regression models, each unit increase in communal goal endorsement (e.g., "somewhat unimportant" vs. "somewhat important") predicted an average 42% decrease in the odds of joining the computing workforce (b = -0.54; $p < .0001$) and a 40% decrease for joining the engineering workforce (b = -0.51; $p < .0001$). Multivariable analyses estimated the unique link between communal goal endorsement and employment by controlling for 20 other predictors (i.e., 8 other career goals and 12 demographic variables; see "Analytic Strategy" in the Methods section). Even after controlling for these other predictors, communal goal endorsement continued to significantly relate to joining the computer science workforce (b = -0.54; $p < .0001$), though not the engineering workforce (b = -0.22; p = .077). The left side of Figure 11 summarizes these logistic regression analyses for predicting joining the computer science and engineering workforce.

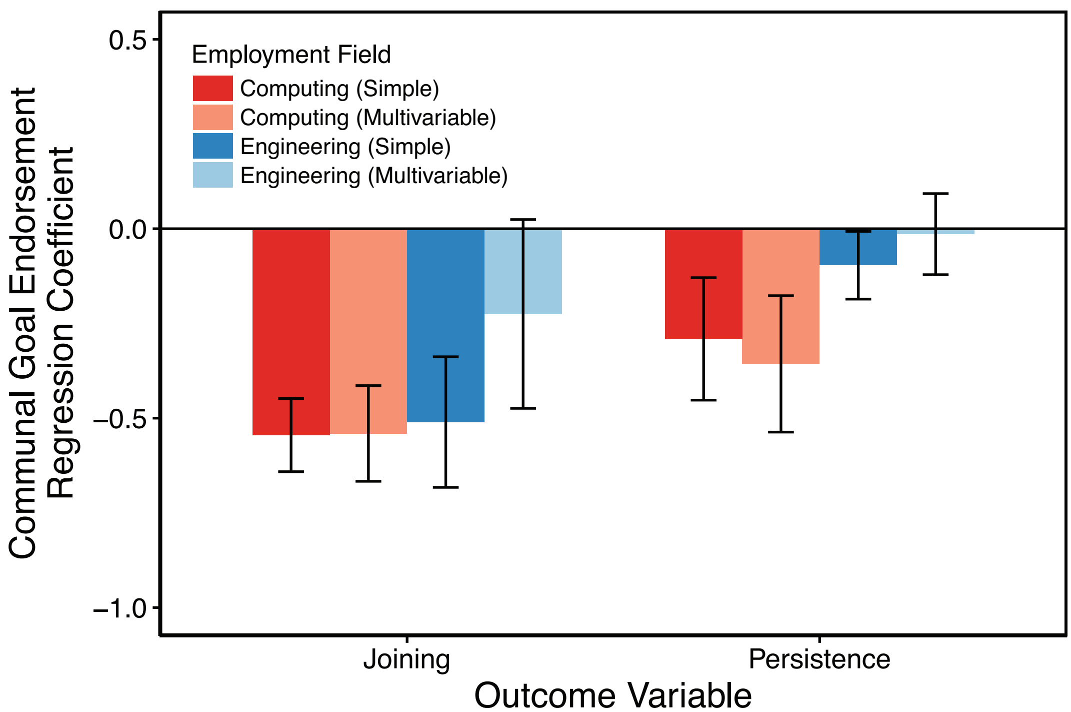

Figure 11. Regression coefficients for communal goal endorsement predicting joining and persistence in computer science and engineering. The coefficients represent the average change in log odds in working in computer science or engineering per unit increase in communal goal endorsement (e.g., rating contribution to society as "somewhat unimportant" vs. "somewhat important"). Error bars represent 95% confidence intervals.

Predicting persistence. The right side of Figure 11 shows the results for predicting persistence (e.g., predicting computer science employment among workers whose highest degree was in computer science). Those results were based on workers whose highest degree was in computer science (n = 4,656) or engineering (n = 16,755).

The regression coefficients were generally smaller in magnitude for predicting persistence than joining. For instance, among workers whose highest degree was in computer science, each unit increase in communal goal endorsement predicted an average 25% decrease in the odds of computer science employment ($b = -0.29$; $p = .0004$). This coefficient for predicting persistence was significantly smaller in magnitude ($p = .008$) than for predicting computer science joining ($b = -0.54$; $p < .0001$). For engineering employment, the regression coefficient was also significantly smaller in magnitude ($p < .0001$) for predicting engineering persistence ($b = -0.10$; $p = .034$) than joining ($b = -0.51$; $p < .0001$). However, the differences in persistence versus joining coefficients were no longer significant in multivariable models ($p = .102$ for computing; $p = .128$ for engineering), perhaps due to the greater imprecision of regression coefficients in those models.

These results regarding persistence can also be interpreted by comparing communally oriented workers (who rated contribution to society as somewhat or very important) to other workers. For instance, among computer science graduates, communally oriented workers were 13% less likely to work in computer science than other workers (58% vs. 67%; $p = .019$). In contrast, among non-pSTEM graduates, communally oriented workers were 52% less likely to join the computing workforce than other workers (2.2% vs. 4.6%). Communal goal endorsement therefore tended to predict the odds of employment outcomes more strongly for joining than persistence pathways in computer science and engineering.

Sex differences. Consistent with prior research (e.g., Konrad et al., 2000; Su et al., 2009), women endorsed communal goals more strongly than men. For instance, among workers without pSTEM degrees, 61% of women rated contribution to society as "very important," compared to 48% of men ($p < .0001$). In addition, 95% of women were communally oriented (i.e., selected the somewhat or very important option), compared to 87% of men ($p < .0001$). Hence, most women and men cared about the societal implications of their work, but women did so more than men. I therefore examined whether sex moderated the predicted effect of communal goal endorsement.

Exploratory moderation analyses generally found that communal goal endorsement related to employment outcomes more strongly for women than men. For instance, communal goal endorsement related to joining the computing workforce more strongly for women (b = -0.74; $p < .0001$) than men (b = -0.34; $p < .0001$). This difference in regression coefficients was statistically significant (i.e., sex significantly interacted with goal endorsement when predicting computer science employment; $p = .0001$). For instance, communally oriented women joined the computing workforce 68% less often than other women without pSTEM degrees (1.4% vs. 4.3%; $p < .0001$). In contrast, communally oriented men joined computing only 29% less often than other men (3.4% vs. 4.8%; $p = .019$). The regression coefficients for predicting joining the engineering workforce were also larger in magnitude for women (b = -0.56; $p = .044$) than men (b = -0.31; $p = .0009$), though this difference was not significant ($p = .402$). In addition, the coefficients for predicting computer science and engineering persistence tended to be larger in magnitude for women than men, though these differences were not significant (ps > .13).

Communal goal endorsement therefore tended to matter as a psychological consideration somewhat more for women than men. Nevertheless, most men still valued contributing to society, which negatively predicted their employment in computing and engineering. Among men, communal goal endorsement negatively related to computer science joining (b = -0.34; $p < .0001$), engineering joining (b = -0.31; p = .0009), and computer science persistence (b = -0.18; p = 0.049), though not engineering persistence (b = -0.08; p = .115). In other words, despite sex differences, communal goal endorsement still mattered for both women and men.

These sex differences in communal goal endorsement also helped interpret sex differences in joining the computing and engineering workforce. Among workers without pSTEM degrees, women were 57% less likely than men to work in computer science (1.5% vs. 3.6%; $p < .0001$) and 84% less likely to work engineering (0.1% vs. 0.8%; $p < .0001$); see Figure 12. Women were 56% of workers without pSTEM degrees, but only 35% of computer science joiners and 18% of engineering joiners.

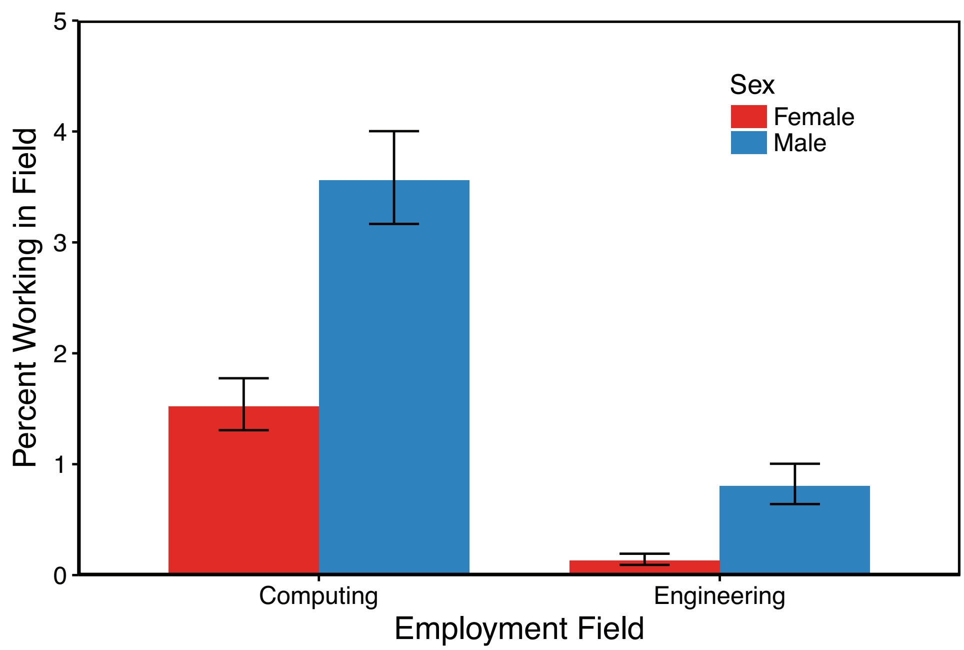

Figure 12. Sex differences in computer science and engineering employment among workers without pSTEM degrees (n = 44,864). Error bars represent 95% confidence intervals.

These sex differences in joining rates, however, tended to be smaller among the minority of workers (9%) who rated contribution to society as “somewhat unimportant” or “not at all important." Among those workers who were not communally oriented, women were only 10% less likely than men to work in computer science (4.3% vs. 4.8%; p = .690) and 61% less likely to work in engineering (0.4% vs. 0.9%; p = .264). Neither of those sex differences were significant, though the lack of statistical significance might also reflect the reduced sample size of workers without pSTEM degrees who were not communally oriented (n = 3,131).

As noted earlier, the interaction between sex and communal goal endorsement was significant for predicting computer science employment (p = .0001). That interaction also meant that the sex difference in joining computer science was significantly smaller among less communally oriented workers. In other words, women and men joined the computing workforce at more equal rates among workers who said contributing to society was less important (e.g., compare the left vs. right side of Figure 13).

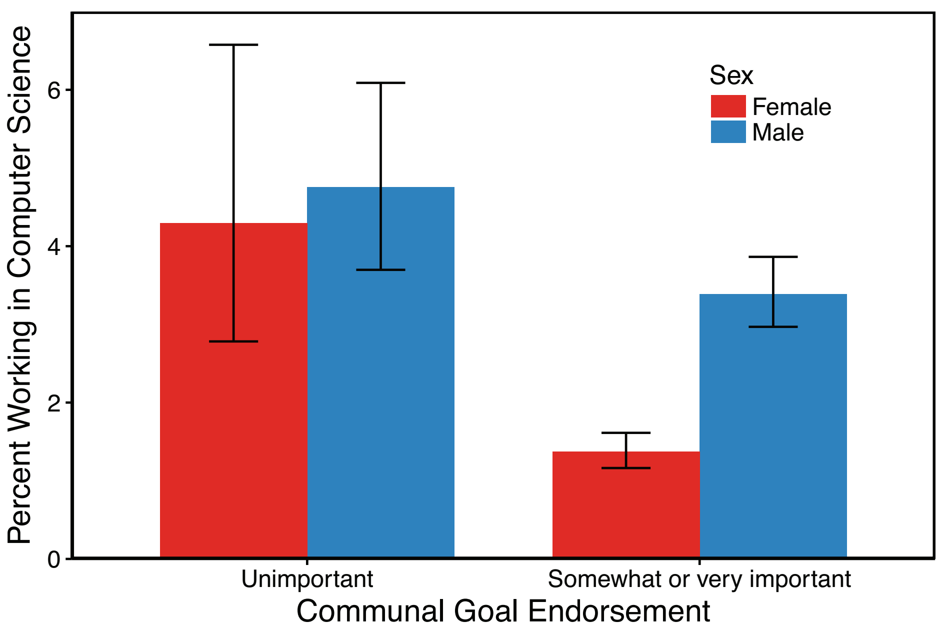

Figure 13. Sex differences in joining computer science by communal goal endorsement. The left side represents data from workers without pSTEM degrees who rated contribution to society as "somewhat unimportant" or "not at all important" (n = 3,131) and the right side represents

workers who selected the “somewhat important” or “very important” options (n = 41,733). Error bars represent 95% confidence intervals.

Summary. In summary, communal goal endorsement negatively related to joining computing and engineering, based on analysis of nearly 45,000 workers without pSTEM degrees. For instance, communally oriented workers were 52% less likely to join the computing workforce than other workers. Simple and multivariable logistic regression models showed this relation was empirically robust with only one exception (i.e., the multivariable model for engineering; p = .077). In contrast, communal goal endorsement related less strongly to the odds of persistence than joining (e.g., among computer science graduates, communally oriented workers were only 13% less likely to work in computer science than other workers).

Goal endorsement also tended to relate to employment outcomes more strongly for women than men, though this difference was only significant for predicting computer science joining (p = .0001). Conversely, the sex difference in computer science joining was smaller among less communally oriented workers. Despite sex differences, most men also valued contributing to society, which negatively related to men’s employment in computing and engineering.

What is the National Impact of Widening Joining Pathways?

The results presented so far suggest that communally oriented workers avoided computing and engineering careers more often than other workers. In this section, I estimated the national impact of these fields’ apparent lack of attractiveness to these workers who valued helping others.

Analytic approach. Given its precisely defined sampling frame, the nationally representative NSCG survey allowed me to quantify how many more college graduates would have been computer scientists and engineers in different hypothetical scenarios. For instance, as noted earlier, 2.2% of communally oriented workers without pSTEM degrees were computer scientists in 2015. Analyses considered the national impact of raising this joining rate from 2.2% to 4.6%, which was the computer science joining rate for other workers without pSTEM degrees. This percentage point difference was multiplied by the estimated population size of communally oriented U.S. workers without pSTEM degrees (see “Analytic Strategy” in the Methods section) to estimate the number of additional computer scientists in this scenario. The numbers in the following paragraphs strictly reflected college graduates because the NSCG survey only sampled people living in the United States in 2015 who had earned a bachelor’s degree or higher.

Computer science employment. Impact analyses found that an additional 820,000 college graduates would have been computer scientists if communally oriented workers had joined computer science as often as other workers without pSTEM degrees. This change would have increased the size of the U.S. college-educated computing workforce by 29% in 2015. This large increase comes in part from the large population size of communally oriented workers without pSTEM degrees, which was estimated to be 34.4 million in 2015. Given this large pool of potential joiners, increasing communally oriented workers’ joining rate from 2.2% to 4.6% had a large estimated population effect (over 800,000 more computer scientists), even though the change in percentage point units (2.4 points) was small.

In contrast, compared to closing the joining gap, closing the persistence gap between communally oriented versus other workers would have yielded fewer computer scientists. Among communally oriented workers who earned their highest degree in computer science, 62% persisted by working in computer science in 2015. Increasing this persistence rate to 73%, the rate for other computer science graduates, would have generated an additional 140,000 computer scientists. In other words, closing the joining gap would have generated 5.9 times as many computer scientists as closing the persistence gap.

The joining gap was more consequential than the persistence gap because potential joiners far outnumbered potential persisters. For instance, among communally oriented workers, an estimated 34.4 million had no pSTEM degree, but only 1.62 million had earned their highest degree in computer science. In other words, communally oriented potential joiners outnumbered communally oriented potential persisters by a ratio of 21 to 1. Consequently, closing the joining gap mattered more, even though the percentage point change in computer science employment was smaller for potential joiners (2.2% to 4.6%) than potential persisters (62% to 73%). The left side of Figure 14 summarizes these impact analyses for computer science employment.

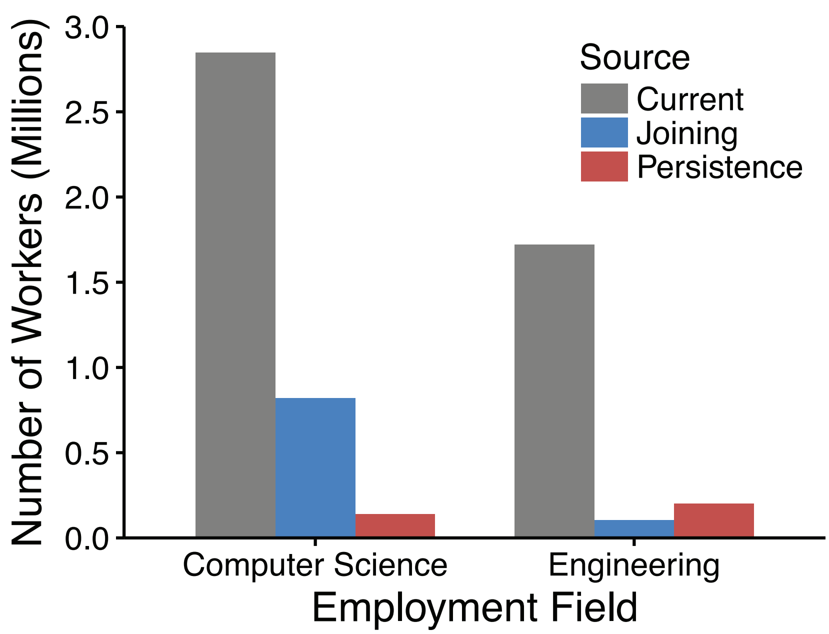

Figure 14. Impact analyses for increasing computer science and engineering employment rates among communally oriented workers. The grey bars represent estimated workforce sizes in 2015. The blue bars represent the number of additional workers generated if communally oriented workers had joined computer science and engineering as often as other workers without pSTEM

degrees. The red bars represent analogous results for persistence (e.g., if communally oriented engineering graduates had worked in engineering as often as other engineering graduates).

Engineering employment. The right hand of Figure 14 repeated these impact analyses for engineering employment. The engineering joining rate was 0.4% for communally oriented workers, compared to 0.7% for other workers without pSTEM degrees. As shown in Figure 14, closing this engineering joining gap would have had a smaller national impact than closing the computer science joining gap (compare the blue bars). If communally oriented workers had joined engineering as often as other workers without pSTEM degrees, an additional 105,000 college graduates would have been engineers, which would have increased the size of the engineering workforce by 6% in 2015. The joining gap was smaller in percentage point units for engineering (0.4% vs. 0.7%) than computer science (2.2% vs. 4.6%) and therefore also less consequential for engineering. In addition, the right red bar in Figure 14 shows results for engineering persistence. Among the 3 million communally oriented workers who earned their highest degree in engineering, 37% persisted by working in engineering in 2015, compared to 44% for other engineering graduates. Closing this persistence gap would have yielded 200,000 more engineers.

Sex differences in computer science employment. Impact analyses for computer science employment were repeated separately for women and men because of the sex differences identified earlier (e.g., sex interacted with communal goal endorsement when predicting computer science joining). For instance, the computer science joining rate was 1.3% for communally oriented working women, compared to 4.3% of other working women without pSTEM degrees. Analyses estimated the impact of closing such gaps separately for women and men.

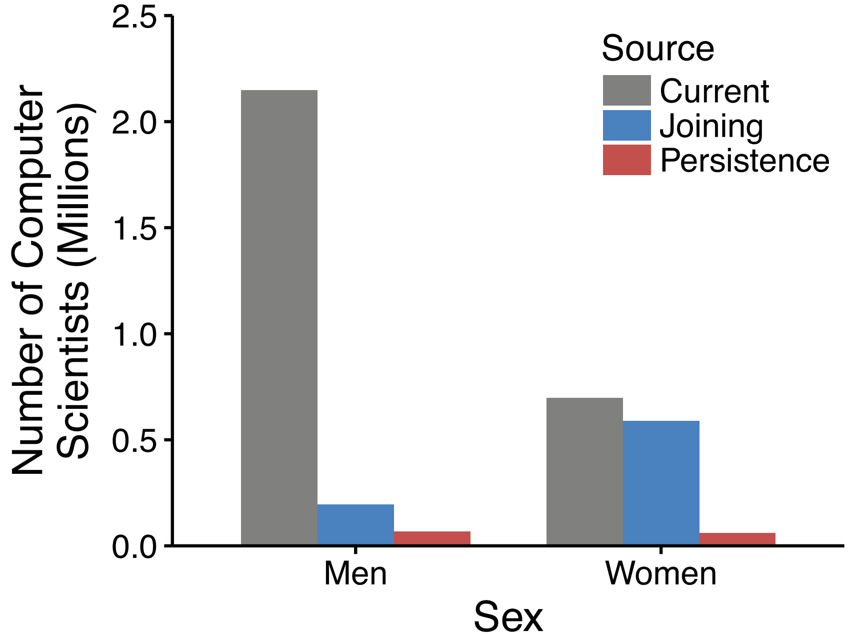

Figure 15. Impact analyses separated by sex for increasing computer science employment rates among communally oriented workers. The grey bars represent estimated workforce sizes in 2015. The blue bars represent the number of additional workers generated if communally oriented workers had joined computer science as often as other workers without pSTEM degrees. The red bars represent analogous results for persistence.

Figure 15 displays results for these sex-disaggregated impact analyses. Perhaps the most striking result is that closing the joining gap for women would have nearly doubled the number of female computer scientists. More than an additional half million women (588,000) would have been computer scientists if communally oriented women had joined the computer science workforce as often as other women without pSTEM degrees. This change would have increased the number of U.S. female computer scientists by 84% in 2015. In contrast, closing the joining gap for men would have yielded far fewer male computer scientists (195,000). In this scenario of closing the joining gap for both women and men, women's representation among computer scientists would have increased from 24.5% to 35.4% in 2015. Lastly, closing the persistence gap would have also produced even fewer male and female computer scientists (67,000 and 60,000, respectively).

Summary. In summary, impact analyses suggested that the United States would benefit from many more computer scientists if communally oriented workers found computing more attractive. An additional 820,000 college graduates would have been computer scientists if communally oriented workers had joined computer science as often as other workers without pSTEM degrees. Sex-disaggregated results showed that most of these additional computer scientists would have been women. The number of female computer scientists would have nearly doubled if communally oriented women had joined computing as often as other women without pSTEM degrees.

In contrast, compared to closing the joining gap, closing the persistence gap between communally oriented versus other workers would have yielded far fewer computer scientists (140,000). This result reflects that potential joiners (i.e., college graduates without pSTEM degrees) far outnumbered potential persisters (i.e., computer science graduates). Consequently, communal goal pursuit appears to matter far more in practical terms for joining than retention pathways in computing. Lastly, results also indicated differences between computing and engineering. Closing the joining gap would have yielded far fewer engineers (105,000) than computer scientists (820,000), reflecting that workers without pSTEM degrees joined engineering far less often than computing.

Discussion

This study used the nationally representative NSCG survey to characterize pathways for joining the computing and engineering workforce among college graduates without pSTEM degrees. By focusing on computing and engineering, this study answered research questions with field-specific granularity while still accounting for most (78%) of the college-educated STEM workforce in 2015. Descriptive analyses provided insight on what joiners did in their jobs, and analyses focused on communal goals helped reveal psychological considerations potentially important to individuals' career decisions.

Job Characteristics

Descriptive analyses suggested that many computer science and engineering joiners used their educational training by working on non-STEM job tasks such as finance and management at higher rates than persisters. Joiners with business degrees were especially likely to say that their job was related to their educational training, despite joiners lacking computer science or engineering degrees. This result could reflect that business knowledge and skills are important to success in computing and engineering careers, especially considering that for-profit companies employ most workers in these fields (National Science Board, 2018). In contrast, joiners spent somewhat less time than persisters on STEM job tasks such as applications development and equipment design. These work activity results were also reflected in the detailed job categories (e.g., computer science joiners worked as computer software engineers less often than computer science persisters).

The differences between joiners and persisters, however, were modest compared to the larger differences between joiners and other workers without pSTEM degrees. For instance, 63% of computer science joiners spent at least 10% of their job on "computer programming, systems or applications development," compared to 10% of other workers without pSTEM degrees (the percentage for computer science persisters was 87%). In other words, joiners did not work only in jobs with a trivial demand for STEM skills (e.g., joiners did not work only in customer service roles). For instance, only a minority of both computer science joiners (13%) and computer science persisters (7%) worked as computer support specialists. Future research should therefore study how joiners learned their technical skills, despite lacking formal pSTEM degrees at the bachelor's level or higher.

Goal Pursuit Processes

In addition to characterizing joiners' jobs, my analyses provided the first nationally representative study of how communal goal endorsement relates to STEM employment. Results clearly indicated the importance of goal pursuit processes, especially for joining the computer science workforce. For instance, among non-pSTEM college graduates, workers who rated benefitting society as somewhat or very important were 52% less likely to join the computing workforce than other workers. If these communally oriented workers had joined computing as often as other workers, over 800,000 more college graduates would have been computer scientists in 2015. These claims about national impact were justified because the NSCG was a large national probability sample, allowing me to estimate population sizes and nationally representative employment rates.

In multiple ways, communal goal pursuit processes mattered more for joining than persistence pathways in computing. First, communal goal endorsement predicted the log odds of computer science employment more strongly among non-pSTEM graduates than computer science graduates. This result could have reflected communally oriented computer scientists persisting in their careers by finding opportunities to help others or engaging in role reconstruction and reconstrual processes, as noted in the introduction. Furthermore, goal pursuit mattered more for joining than persistence pathways in practical terms as well. Communally oriented non-pSTEM graduates (i.e., potential joiners) outnumbered communally oriented computer science graduates (i.e., potential persisters) by a ratio of 21 to 1. Consequently, even small changes in joining rates could generate many more computer scientists. Impact analyses showed that, compared to closing the persistence gap, closing the joining gap between communally oriented versus other workers would have yielded 6 times as many computer scientists. In other words, both theoretically and practically, communal goal endorsement was more important for computer science joining than persistence pathways.

In contrast, the importance of communal goal endorsement was more mixed for engineering than computer science joining pathways. Non-pSTEM graduates worked in engineering far less often than computing overall (0.4% vs. 2.4%), indicating the pathways were less open for joining engineering than computer science. Given these base rates, the differences between communally oriented versus other non-pSTEM graduates mattered less for the engineering than computer science workforce. Nevertheless, communal goal endorsement still predicted large decreases in the log odds of joining engineering (see Figure 11). In other words, communal goal pursuit appeared to still be an important psychological consideration for joining engineering, even though those pursuit processes had a comparatively small impact on the number of engineers in the United States.

Sex Differences

Consistent with prior research (e.g., Konrad et al., 2000), women valued benefitting society more than men in the NSCG sample. Variation in communal goal endorsement also related to the log odds of computer science joining more strongly for women than men. Furthermore, increasing joining rates among communally oriented non-pSTEM graduates would have mainly generated more female, rather than male, computer scientists. If communally oriented women had joined computing as often as other women, the number of female computer scientists would have almost doubled in 2015. Communal goal pursuit processes therefore were more important for women than men when joining computer science.

These sex differences in goal pursuit can help interpret sex differences in joining rates. For instance, women were overall 57% less likely than men to join computing, but this sex difference was not significant among the subset of workers who rated benefitting society as unimportant. The sex difference in joining computing was significantly smaller among less versus more communally oriented workers (p = .0001). In other words, women and men joined at more equal rates when pursuit of communal goals was a less important psychological consideration. These results therefore add to other evidence that sex differences in goal pursuit processes may contribute to women's underrepresentation in computer science (Cortes & Pan, 2017; Diekman et al., 2017; Su & Rounds, 2015). For instance, if communally oriented women and men had joined computing as often as other women and men, women's representation among computer scientists would have increased from 25% to 35% in 2015.

The contribution of goal pursuit processes to women's underrepresentation in engineering careers was less clear from this study. For instance, the interaction between sex and communal goal endorsement was not significant for predicting engineering joining (p = .402), though it was significant for computer science joining (p = .0001). This result might reflect statistical limitations because joining engineering was a rare event, which could have compromised power for detecting interactions. Regardless, the sex difference in engineering joining had limited impact on aggregate gender diversity in engineering because joining rates were so low. Goal pursuit processes could still impact diversity in engineering by limiting the number of women who earn engineering degrees before entering the workforce (Diekman et al., 2017). However, studying such pre-workforce pathways was beyond the scope of this study, and datasets included in Study 1 of my dissertation did not measure communal goals.

Despite sex differences, goal pursuit also mattered for men. For instance, among non-pSTEM graduates, most men rated contributing to society as somewhat important (40%) or very important (48%). Variation in communal goal endorsement also negatively related to men's employment in computer science and engineering. For instance, an additional 195,000 men would have been computer scientists if communally oriented men had joined computing as often as other men without pSTEM degrees. In other words, wanting to benefit society also related to men's employment outcomes, even though women valued that career goal more on average.

Limitations

Several limitations of this research should be noted. First, this study's cross-sectional, correlational design limits causal claims about the effect of communal goal endorsement on employment outcomes. For instance, reverse causation is possible; computer scientists and engineers may have disengaged from their goals if they could not find ways to help others in their jobs. Hence, employment outcomes could influence communal goal endorsement, which could reflect available career opportunities. This alternative explanation is unlikely because communal goals are often fundamental motives that may be difficult to abandon (Diekman et al., 2017). Nevertheless, longitudinal data are needed to evaluate this hypothesis.

Second, this study's measurement of communal goals was limited to a single item that asked respondents to rate the importance of "contribution to society" when "thinking about a job." Respondents may have interpreted the term "contribution to society" in different ways given this vague wording. In contrast, other surveys like O*NET's often provide concrete definitions of terms such as describing "assisting and caring for others" as "providing personal assistance, medical attention, emotional support, or other personal care to others such as coworkers, customers, or patients" (National Center for O*NET Development, 2018). Furthermore, the NSCG item measured only one goal (i.e., benefitting society), but not other communal goals such as working with others or directly helping others through face-to-face interactions. A battery of communal goal items should therefore be used to study employment outcomes in future research.

However, in defense of my findings, items that measure different communal goals have correlated strongly in prior research (e.g., Diekman et al., 2010).

Third, another limitation was that the NSCG survey measured a narrow range of psychological attributes. For instance, one important construct to include in future employment research would be goal affordance stereotypes (i.e., perceptions that computing and engineering jobs offer communal opportunities). Based on prior research (e.g., Diekman et al., 2010), I assumed that workers without pSTEM degrees generally viewed computing and engineering jobs as lacking in ways to help others, but I was unable to directly test this assumption for this NSCG sample. Variation in affordance stereotypes could have helped better characterize the underlying psychological considerations leading to individuals' career decisions. Nevertheless, despite such limitations, the NSCG measured some other psychological attributes such as career goals for wanting intellectual challenge and opportunities for advancement. These other items helped me estimate the unique link between communal goal endorsement and employment outcomes, controlling for other attributes.

Implications for Employers

This study provides STEM employers with knowledge on recruiting workers with non-STEM degrees and leveraging those workers' educational training. This knowledge is particularly important for computer science employers (e.g., technology companies) given the openness of pathways for joining the computing workforce after college. For instance, non-pSTEM college graduates were roughly one in three (31%) college-educated computer scientists in 2015. In contrast, this percentage was much smaller for college-educated engineers (9%).

For-profit technology companies might see joiners' non-technical skills such as communication and organizational skills as an asset for enhancing business outcomes (e.g., profit). Consistent with this hypothesis, joiners spent more time than persisters on some non-STEM tasks such as management and finance. Nevertheless, questions remain about how joiners developed their technical skills. Employers may be hesitant to hire non-pSTEM graduates as computer scientists because on-the-job training programs for learning programming skills could be costly. However, joiners could have also learned computing skills by gradually taking on technical tasks (e.g., database administration) in non-STEM jobs, easing the burden on computer science employers. Future research should examine such possibilities to inform employers' decisions.

Computer science employers should also be aware that, based on my study, computing careers appear to be unattractive to workers who value benefitting society. This finding is concerning because a clear majority of non-pSTEM graduates rated contributing to society as somewhat important (36%) or very important (55%) when thinking about a job. If these communally oriented workers had joined computing as often as other workers, the college-educated computer science workforce would have been 29% larger in 2015, holding all other factors constant. My analyses did not address whether employers could have accommodated or needed so many more computer scientists, but these results nevertheless illustrate the potential importance of joining pathways for addressing workforce needs (National Science Board, 2018).

Employers therefore may want to consider how to highlight the communal aspects of computing careers to potential employees. For instance, experimental studies could test how variations in job advertisements (e.g., emphasizing the importance of interpersonal skills or not) could change communally oriented workers' interest in applying. This research would be valuable especially considering that computer scientists often rate interpersonal activities such as providing advice to others and communicating with peers as important to job performance (National Center for O*NET Development, 2018). In other words, computing jobs appear to offer ways to work with and help others, despite stereotypes suggesting otherwise.

Chapter 4: Conclusions and Implications for STEM Education and Workforce Needs

My dissertation found that the educational and career trajectories for becoming scientists and engineers in the United States are far more varied and complex than commonly assumed. For instance, nearly one fifth of STEM graduates started college as a non-STEM major, and one fifth of college-educated STEM workers had no postsecondary degree in any STEM field. Many college students and working adults therefore already have joined STEM fields from non-STEM backgrounds. Further widening these joining pathways could offer educators and employers new strategies for broadening the national supply of workers with STEM skills. These late entry points into STEM are particularly important for increasing gender diversity, given the large pool of women who start college as a non-STEM major and graduate with non-STEM degrees.

Despite illustrating these opportunities, my analyses also identified barriers that could limit access to these pathways for joining STEM. In Study 1, undergraduates were unlikely to later join STEM if they had not taken a high number of STEM courses early in college (e.g., less than two STEM courses in students' first semester in college). In Study 2, computing and engineering careers appeared to be unattractive to non-STEM graduates who valued benefitting society. Joining STEM during or after college is therefore not always a straightforward process given the structural barriers (e.g., course requirements) and psychological barriers (e.g., stereotypes about the nature of work in the field) that could impede late entry into STEM. Educators, employers, and policymakers therefore may want to consider how to overcome these obstacles through targeted interventions such as modifying course-taking policies for non-STEM majors and communicating how STEM careers offer ways to work with and help others.

My dissertation identified these opportunities and challenges by organizing analyses into three guiding research questions: (a) who are STEM joiners? (b) what predicts STEM joining?

(c) what is the potential national impact of widening joining pathways? By investigating these questions, I aimed to develop foundational knowledge that could inform future studies on policies and interventions for leveraging the diverse entry points into STEM. My research informs both educational and employment practices because I studied both the transitions from beginning of college to graduation and college graduation to the workforce.

Key insights from both studies suggested that (a) joiners' interdisciplinary training could be an asset for STEM careers that often require social and organizational skills in a business context; (b) taking STEM courses early in college and wanting to benefit society were robust predictors of STEM joining; and (c) even small changes in STEM joining rates could generate many more STEM graduates and workers. These results advance theoretical models of individuals' pursuit of STEM majors and careers as well as suggest new practical strategies for solving workforce needs.

Integrating Results Across Studies

Who Are STEM Joiners?

One overarching conclusion from both studies was that STEM joiners retained aspects of their non-STEM backgrounds and training in their ongoing education and work. Compared to STEM persisters, joiners took more non-STEM courses throughout college (Study 1) and worked more often on non-STEM job activities such as finance and management (Study 2). One hypothesis for these results is that joiners had a broader array of interests and skills than persisters (Valla & Ceci, 2014; Wang et al., 2013). Joiners had demonstrated a strong prior interest in at least one non-STEM field (e.g., by starting college as a non-STEM major) and therefore may have retained that interest even after formally transitioning into STEM majors and careers. In contrast, persisters may have spent their education and careers honing their skills more narrowly in a single technical discipline like mechanical engineering or biology.

Joiners' broad educational training could be an asset for STEM careers because those careers often require strong communication and organizational skills in a business context (National Center for O*NET Development, 2018). Consistent with this hypothesis, most non-STEM graduates who worked in computing (63%) or engineering (85%) said that their job was somewhat or closely related to their non-STEM educational training. In other words, based on workers' self-reports, joiners appeared to use their non-STEM training in their technical jobs.

Nevertheless, joiners' educational training might have also been a liability because joiners had fewer learning experiences than persisters to develop technical skills. Joiners' comparatively weaker STEM preparation therefore might have threatened their educational and career success (Kokkelenberg & Sinha, 2010). Results from Study 1, however, did not support this concern empirically for postsecondary education outcomes. Joiners and persisters did not significantly differ in terms of undergraduate STEM course performance and rates of graduating college on time (i.e., within four years), even though joiners had less high school STEM preparation than persisters. In other words, joiners achieved similar undergraduate success in STEM compared to persisters. Nevertheless, the career success of joiners is still largely unknown because Study 1 only examined educational outcomes and Study 2 did not have measures of job performance. Future research should therefore extend my dissertation by studying employers' evaluations of joiners' job performance and the pathways for non-STEM graduates to learn technical skills.

What Predicts STEM Joining?

My dissertation identified potential points of intervention by examining structural and psychological factors that predicted STEM joining. These analyses did not intend to definitively test the causal impact of specific policies and programs but instead aimed to rigorously identify promising directions for future intervention research. Based on my results, two variables emerged as consistently predictive of later STEM joining: (a) taking STEM courses early in college and (b) wanting to benefit society. The first positively predicted later earning of STEM bachelor's degrees among non-STEM majors (Study 1) and the second negatively related to working in computing and engineering among non-STEM graduates (Study 2). These results remained after controlling for several theoretically relevant covariates such as high school STEM preparation and other career goals such as wanting opportunities for advancement.

Postsecondary educators and policymakers therefore should study how early college STEM course-taking could create opportunities for non-STEM majors to later join STEM. For instance, future research could examine how to widen STEM joining pathways by redesigning institution-wide graduation requirements as early-college requirements (e.g., requiring all undergraduates to take a mathematics course in their first college semester). Increasing the quantity and quality of STEM courses that non-STEM majors take early in college could create opportunities for those students to develop interests in STEM and satisfy the course requirements for becoming STEM majors (Eccles, 2011; Merolla et al., 2012). Furthermore, employers should consider how to highlight the communal opportunities available in computing and engineering careers to potential job applicants, building on prior experimental studies with undergraduates (see Diekman et al., 2017 for a review). As an example, future experiments could study how emphasizing the importance of interpersonal skills in STEM job advertisements might encourage non-STEM graduates to apply.

In addition to identifying these practical implications, my dissertation demonstrated how studying STEM joining can modify and enrich theories about what psychological factors are most important to pursuing STEM (e.g., Diekman et al., 2017; Eccles, 2011). Results suggested asymmetries in the psychological processes that lead to STEM joining versus persistence. In Study 1, grades earned in introductory undergraduate STEM courses weakly predicted later

STEM joining, despite strongly predicting persistence among initial STEM majors. In Study 2, the career goal of benefiting society strongly and negatively related to computing and engineering employment among non-STEM graduates, despite relating more weakly to persistence (e.g., computer science graduates working in computing).

Both of these asymmetries in predicting joining versus persistence have implications for psychological theory. The results regarding grades are particularly interesting given the harsh grading standards and demanding work in introductory undergraduate STEM courses that could create self-doubt, driving students away from STEM (see Ceci et al., 2014 for a review). Consistent with this hypothesis, introductory STEM course grades have been a robust predictor of STEM persistence among students starting college as a STEM major (e.g., Chen & Soldner, 2013). But students who started college as a non-STEM major appeared to be much more resistant to the self-doubt that might come from mediocre grades in these so-called "weed out" classes. Only particularly poor STEM grades (GPA < 2.00) predicted a decline in rates of later joining STEM. In fact, STEM joiners earned an average of 0.26 grade points lower in their first-year STEM courses than non-STEM courses, but still decided to major in STEM anyway. One hypothesis is that below average STEM course performance may be less threatening to those that do not already identify with the domain, consistent with broader literature on social identity threat (e.g., Murphy et al., 2007).

In addition, the results regarding communal goals (i.e., benefitting society) suggest distinct goal pursuit processes in the joining versus persistence pathways that lead to computer science and engineering employment. For joining pathways, non-STEM graduates likely base career decisions on cultural stereotypes that portray computing and engineering careers as not offering communal opportunities (see Diekman et al., 2017 for a review). Such workers have limited direct exposure to the career opportunities available in those jobs. In contrast, for persistence pathways, computer science and engineering graduates may realize that employment in those fields offers varied ways to work with and help others (e.g., collaborating with colleagues, using data science to advance organizational goals). Communally oriented computer science and engineering workers might also align their jobs with their goals over time through role reconstruction (e.g., negotiate different working conditions) or role reconstrual (e.g., mentally reframe the nature of their work), as Diekman et al.'s (2017) goal congruity perspective would also predict. In other words, communal goal endorsement may have predicted joining more strongly than persistence because of differences in the underlying psychological processes in how workers pursued their career goals.

These results therefore both advance theoretical models of individuals' pursuit of STEM pathways as well as suggest new practical strategies for meeting workforce needs. Future research can build on my dissertation to identify other structural and psychological factors that may be important to entry into STEM during college and beyond. For instance, during college, students' mentoring relationships with STEM professors and informal social relationships with peers in STEM courses may be critical for non-STEM majors to develop a sense of belonging and identity in STEM (Wu & Uttal, 2017). After college, the high-paying job opportunities in STEM careers may be attractive to non-STEM graduates who might struggle finding such employment in other fields (National Science Board, 2018). My dissertation therefore demonstrates promise for studying STEM joining as a programmatic line of future, generative research.

What is the Impact of Widening Joining Pathways?

Impact analyses also demonstrated promise for how STEM joining pathways could broaden participation in postsecondary education and careers in STEM. One key insight from both studies was that even small changes in STEM joining rates could have a large national impact. For instance, Study 2 demonstrated the impact of making computer science more attractive to non-pSTEM graduates who valued benefitting society. If communally oriented non-pSTEM graduates had joined computing as often as other non-pSTEM graduates (2.2% vs. 4.6%), an additional 820,000 workers would have been computer scientists in 2015, which would have increased the size of the computing workforce by 29%. Even though the change was small in terms of joining rates (2.4 percentage points), the impact on the number of workers was considerable given the large pool of potential joiners. Among those who valued benefitting society, non-pSTEM graduates (i.e., potential joiners) outnumbered computer science graduates (i.e., potential persisters) by a ratio of 21 to 1.

These considerations are particularly important for increasing gender diversity in STEM. Few women start college and earn degrees in male-dominated STEM fields such as computer science and engineering. Standard interventions focusing on increasing women's persistence in pSTEM during college and beyond therefore may have limited impact on gender diversity. In contrast, the pool of women who start college and earn degrees in non-STEM fields is much larger. For instance, in Study 1, women who started college as a non-STEM major outnumbered women who started college as a pSTEM major by a ratio of 14 to 1. Impact analyses found that, compared to "plugging the leaky pipeline," closing the gender gap in undergraduate joining would more potently increase women's representation in pSTEM.

Furthermore, Study 2's impact analyses found that closing the gap between communally oriented versus other women would have mainly generated more female computer scientists through joining rather than persistence pathways. More than a half million more women (588,000) would have been computer scientists if communally oriented women had joined computer science as often as other women without pSTEM degrees. This increase would have nearly doubled the number of female computer scientists in 2015. In contrast, the potential increase in gender diversity through persistence pathways was much smaller. Only 60,000 more women would have computer scientists if communally oriented women had persisted as often as other women with computer science degrees. In other words, sex differences in communal goal pursuit processes appeared to matter more for joining than persistence in computer science.

These findings indicate that educators, employers, and policymakers should not dismiss the importance of STEM joining pathways because the act of joining STEM may seem infrequent (e.g., only 1 in 41 workers without pSTEM degrees were computer scientists). These stakeholders must consider the large pool of potential joiners when forming policy decisions and designing interventions to address workforce needs and increase gender diversity in STEM. As the President's Council of Advisors on Science and Technology argued, much attention has been given to "off-ramps" and attrition patterns in STEM, but "equal attention should be given to on-ramps, multiple routes to enter or re-enter STEM" (PCAST, 2012, p. 31).

Rethinking Common Assumptions About STEM Education and Employment

My dissertation's results collectively show that educators, employers, and policymakers should rethink several common assumptions about STEM education and careers in the United States. The processes for obtaining STEM degrees have often been viewed as a "pipeline" that requires completing key educational milestones such as taking calculus in high school, starting college as a STEM major, and persisting until graduation. Furthermore, STEM jobs "are generally assumed to require at least a bachelor's degree of education in [a science or engineering] field," as the National Science Board (2018) noted (p. 3-12). However, as my dissertation shows, both of these assumptions are often inaccurate and may direct attention and resources away from more comprehensive strategies for broadening participation in STEM.

Evaluating the Role of Postsecondary STEM Education

The results from Study 2 also raise broader questions about the role of postsecondary STEM education in meeting workforce needs. For instance, more than one million STEM workers were college graduates without any STEM degree at the bachelor's level or higher. Given this fluidity for joining the STEM workforce, are more STEM college graduates even needed to fulfill workforce demands? Or can workforce demands instead be met by more non-STEM graduates developing technical skills after college? Furthermore, many STEM graduates do not work in STEM careers after college. For instance, in Study 2, 40% of working computer science graduates and 62% of working engineering graduates were not employed in their educational field of study. Given this rate of attrition after college, are efforts to increase the number of STEM graduates justified?

One answer to these questions is that the need for STEM skills and knowledge in the modern U.S. economy extends well beyond jobs formally classified as STEM. For instance, marketing analysts are typically categorized as non-STEM business jobs, but their work often requires statistical knowledge and data analysis skills (Bidwell, 2014). Physicians and nurses are also usually categorized as non-STEM healthcare jobs, but their work heavily relies on life science knowledge. Most STEM graduates report using their educational training in their work, despite often working in jobs not formally categorized as STEM (National Science Board, 2018). For instance, among computer science graduates not formally working in computer science in 2015, 82% said their job was somewhat or highly related to their highest degree (National Survey of College Graduates, 2015). As Noonan (2017) argued in a report for the U.S. Department of Commerce, "increased technology in the workplace means that, to handle non-repetitive tasks, workers need the critical thinking and technical skills that come with STEM training" (p. 1).

In this regard, the U.S. economy still needs more STEM college graduates, even though many will work in jobs not traditionally categorized as STEM. Results from Study 1 show that one strategy for meeting this need is to create opportunities for non-STEM majors to join STEM during college. Joiners' broad training in both STEM and non-STEM subjects may be in particularly high demand because the fastest growing job areas in the United States have required both analytical and social skills (Deming, 2017). In contrast, the share of math-intensive jobs requiring only low levels of social skills has actually declined from 1980 to 2012 (see Figure IV in Deming, 2017). In other words, undergraduate joiners' interdisciplinary training may be uniquely suited to meet the needs of an economy that increasingly relies on both social skills and technological innovation.

Furthermore, postsecondary educators and policymakers should consider how course requirements for non-STEM majors might also give students the skills and knowledge to later pursue STEM careers. For instance, in its 2017 report on data science talent, the Business-Higher Education Forum (BHEF) recommended to expand the pathways leading to a diverse analytical workforce by "teaching foundational [data science and analytics] skills in a broad number of degrees" (BHEF, 2017, p. 21). Fitzgerald et al. (2016) recognized the national need for data scientists but argued, "there is an even greater need for data-enabled professionals who can marry a deep background in a particular field (e.g., engineering or economics) with a strong understanding of the application of data science tools" (p. 434).

Some colleges and universities have already begun exploring how to equip liberal arts and business majors with data analytic skills (BHEF, 2016). As an example, Case Western Reserve University created an applied data science minor in fall 2014 for majors in a broad variety of fields including business, engineering, and health (Fitzgerald et al., 2016). In other words, formal postsecondary education will likely still play a pivotal role in preparing the next generation of U.S.

workers with technical skills, but colleges and universities must also adapt to the needs of a rapidly changing economy. My dissertation suggests that fostering pathways for non-STEM majors to join STEM during college may be one viable way to help address these evolving workforce needs.

Preparing Workers for Lifelong STEM Learning

Although STEM degrees remain important, the opportunities to develop technical skills extend well beyond formal postsecondary education (Risen, 2016). Workers without STEM degrees could learn technical skills both through informal experiences (e.g., gradually taking on technical job tasks) and formal experiences (e.g., formal on-the-job training, certification programs). Online training courses increasingly offer inexpensive ways to develop a wide array of technological skills including computer programming, database administration, machine learning, and web development (e.g., see https://www.datacamp.com). Consistent with these varied learning opportunities, in Study 2, nearly one third (32%) of college-educated computer science workers in 2015 had no pSTEM degree at the bachelor's level or higher. Presumably, these computer science joiners had honed their computing skills, in part, after college. Policymakers should therefore view STEM education as a lifelong process of developing skills and knowledge to thrive in an increasingly technological world; STEM education is not limited to formal training in schools and universities (Dierking, Falk, Rennie, Anderson, & Ellenbogen, 2003).

Fostering lifelong STEM learning is critical especially given the rapid pace at which technological advances create changing demands for new types of technical skills and knowledge (BHEF, 2017). In other words, the specific computer algorithms and scientific knowledge that undergraduates learn today may likely become outdated in one to two decades from now. Emerging technologies such as artificial intelligence and robotics can even eliminate the need for certain types of jobs, especially those involving performing routine tasks (Autor, 2015). Displaced workers must then adapt to an evolving economy by developing skills that extend beyond their formal K-12 and college education. Reflecting these challenges, one of NSF's current top priorities for future investment (i.e., one of its "10 Big Ideas") is to advance understanding of the "future of work at the human-technology frontier" (NSF, 2018, p. 15). This strategic initiative, in part, "responds to the pressing societal need to educate and re-educate learners of all ages (students, teachers and workers) in [STEM] content areas to ultimately function in highly technological environments, including in collaboration with intelligent systems" (NSF, 2017, p. 1).

These processes of lifelong learning in a changing economy may partly explain the results I found regarding the varied pathways into the computing workforce. Many college graduates without pSTEM degrees may have pursued computer science employment because they seized opportunities for learning about new and emerging technologies in their current job or through other means (e.g., online courses). Coupled with this technical training, joiners' non-STEM educational training could have been as asset for launching their new careers, especially given the need for communications and interpersonal skills in computing careers (National Center for O*NET Development, 2018).

These considerations sharply contrast with widespread narratives found in many policy discussions (e.g., about the "STEM pipeline") that assume STEM degrees are almost always needed for STEM jobs (National Science Board, 2018, p. 3-12). This assumption overlooks over one million STEM workers without STEM degrees and directs attention away from considering the educational opportunities after formal postsecondary education. For instance, even though STEM graduates obviously remain a vital source of STEM talent, they must still update their knowledge over time to learn new advances in technology and scientific knowledge (NSF, 2017). Hence, postsecondary STEM education can be viewed as *preparation for future learning*, which is a term that Bransford and Schwartz (1999) first popularized to refer to, "people's abilities to learn new information and relate their learning to previous experiences" (p. 69). In this respect, joiners and persisters are alike because they both continue to learn new skills and knowledge after formal education.

Conclusions and Implications

Students often change majors during college (e.g., Chen & Soldner, 2013), and most workers change jobs throughout their careers (e.g., Lyons, Schweitzer, & Ng, 2015). Yet the diverse ways for entering STEM majors and careers are often overlooked. My dissertation addressed this critical oversight by showing how STEM joining pathways offer novel opportunities for meeting workforce needs and broadening participation in STEM.

Based on Study 1, postsecondary educators and policymakers should evaluate how to facilitate STEM joining pathways by increasing the quality and quantity of STEM courses that non-STEM majors take early in college. Based on Study 2, employers should consider how to communicate to potential applicants, especially non-STEM graduates, that STEM careers offer ways to help others. These late entry points also provide new strategies to recruit women into STEM fields, especially given that many women start college as non-STEM majors and earn non-STEM degrees. For instance, in Study 2, the number of female computer scientists in 2015 would have nearly doubled if communally oriented women had joined computing as often as other women without pSTEM degrees.

These implications are merely examples of the many ways that the metaphor of STEM pathways can broaden thinking and research on developing STEM talent. In contrast, the leaky pipeline metaphor has directed attention on plugging "leaks" in the pipeline but away from more comprehensive strategies for addressing workforce needs. The pipeline metaphor especially constrains gender diversity initiatives because few women start college in male-dominated STEM fields such as computer science or engineering. For instance, in Study 1, compared to "plugging the leaky pipeline" for female pSTEM majors, closing gender gaps in pSTEM joining would have much more potently increased women's representation among pSTEM graduates.

Joiners' broad training in both STEM and non-STEM fields could be an asset in an economy that increasingly relies on both social skills and technological innovation (Deming, 2017). STEM careers often require strong communications and organizational skills in a business context (National Center for O*NET Development, 2018). Consistent with these considerations, in Study 2, most joiners said their technical job was somewhat or very related to their non-STEM educational training. Hence, in addition to helping address needs for more STEM workers, joiners could enrich STEM fields with interdisciplinary perspectives and novel insights gained from training in other fields. Further studying the pathways for joining STEM during college and beyond offers promise for leveraging this diverse talent that has often been overlooked.

References

Alper, J. (1993). The pipeline is leaking women all the way along. *Science*, *260*, 409-411. doi:10.1126/science.260.5106.409

American Council of Trustees and Alumni (2014). *Education or reputation? A look at America's top-ranked liberal arts colleges*. Washington, DC: American Council of Trustees and Alumni.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*, 469-480. doi:10.1037//0003-066X.55.5.469

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, *29*, 3-30. doi:10.1257/jep.29.3.3

Bidwell, A. (2014, February 4). Report: STEM job market much larger than previously reported. *U.S. News & World Report*. Retrieved from http://www.usnews.com/news/stem-solutions/articles/2014/02/05/report-stem-job-market-much-larger-than-previously-reported

Bennett, C. (2011). Beyond the leaky pipeline: Consolidating understanding and incorporating new research about women's science careers in the UK. *Business and Economic Review*, *54*, 149 –176.

Berryman, S. E. (1983). *Who Will Do Science?: Minority and Female Attainment of Science and Mathematics Degrees: Trends and Causes*. New York, NY: Rockefeller Foundation.

Boucher, K. L., Fuesting, M. A., Diekman, A. B., & Murphy, M. C. (2017). Can I work with and help others in this field? How communal goals influence interest and participation in STEM fields. *Frontiers in Psychology*, *8*, 901. doi:10.3389/fpsyg.2017.00901

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61-100. doi:10.3102/0091732X024001061

Business-Higher Education Forum (2016). Equipping liberal arts students with skills in data analytics. Washington, D.C.: Business-Higher Education Forum. Retrieved from http://www.bhef.com/sites/default/files/BHEF_2016_DSA_Liberal_Arts.pdf

Business-Higher Education Forum (2017). Investing in America's data science and analytics talent: The case for action. Washington, D.C.: Business-Higher Education Forum. Retrieved from http://www.bhef.com/sites/default/files/bhef_2017_investing_in_dsa.pdf

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*, 31-72. doi:10.1111/j.1467-6419.2007.00527.x

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

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, *15*, 75-141. doi:10.1177/1529100614541236

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregration by field of study in 44 countries. *American Journal of Sociology*, *114*, 924-976. doi:10.1086/595942

Chen, X., & Soldner, M. (2013). *STEM attrition: College students' paths into and out of STEM fields*: Washington, DC: National Center for Educational Statistics.

Charles, C. Z., Fischer, M. J., Mooney, M., & Massey, D. S. (2009). *Taming the river:*

Negotiating the academic, financial, and social currents in selective colleges and universities. Princeton, NJ: Princeton University Press.

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

Cortes, P., & Pan, J. (2017). *Occupation and gender*. Institute for Labor Economics (IZA) Discussion Papers, No. 10672. Retrieved from https://hdl.handle.net/10419/161295

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*, 924-942. doi:10.3102/0002831209349460

Croft, A., Schmader, T. & Block, K. (2015). An under-examined inequality: Cultural and psychological barriers to men's engagement with communal roles. *Personality and Social Psychology Review*, *19*, 343-370. doi:10.1177/1088868314564789

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*, 1593-1640. doi:10.1093/qje/qjx022

Diekman, A. B., Brown, E. R., Johnston, A. M., & 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. *Psychological Science*, *21*, 1051-1057. doi:10.1177/0956797610377342

Diekman, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to STEM careers:

Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology*, *101*, 902-918. doi:10.1037/a0025199

Diekman, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., & Clark, E. K. (2017). A goal congruity model of role entry, engagement, and exit: Understanding communal goal processes in STEM gender gaps. *Personality and Social Psychology Review*, *21*, 142-175. doi:10.1177/1088868316642141

Dierking, L. D., Falk, J. H., Rennie, L., Anderson, D., & Ellenbogen, K. (2003). Policy statement of the “informal science education” ad hoc committee. *Journal of Research in Science Teaching*, *40*, 108-111. doi:10.1002/tea.10066

Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*, 195-201. doi:10.1177/0165025411398185

Fiegener, M. K., & Proudfoot, S. K. (2013). *Baccalaureate origins of U.S.-trained S&E doctorate recipients*. Washington, DC: National Science Foundation.

Fitzgerald, B. K., Barkanic, S., Cárdenas-Navia, I., Chen, J., Elzey, K., Hughes, D., & Troyan, D. (2016). The BHEF National Higher Education and Workforce Initiative: A model for pathways to baccalaureate attainment and high-skill careers in emerging fields, Part 3. *Industry and Higher Education*, *30*, 433-439. doi:10.1177/0950422216674839

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A. B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science*, *341*, 1455-1456. doi:10.1126/science.1240487

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*, 911-922. doi:10.1016/j.econedurev.2010.06.010

Highsmith, J., & Cockburn, A. (2001). Agile software development: the business of innovation. *Computer*, *34*, 120-127. doi:10.1109/2.947100

Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, *23*, 305-327. doi:10.1002/jae.998

Kokkelenberg, E. C., & Sinha, E. (2010). Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students. *Economics of Education Review*, *29*, 935-946. doi:10.1016/j.econedurev.2010.06.016

Konrad, A. M., Ritchie, J. E. J., Lieb, P., & Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: A meta-analysis. *Psychological Bulletin*, *126*, 593-641. doi:10.1177/1059601100252002

Legewie, J., & DiPrete, T. A. (2014). Pathways to science and engineering bachelor's degrees for men and women. *Sociological Science*, *1*, 41-48. doi:10.15195/v1.a4

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*, 262-265. doi:10.1126/science.1261375

Lyons, S. T., Schweitzer, L., & Ng, E. S. (2015). How have careers changed? An investigation of changing career patterns across four generations. *Journal of Managerial Psychology*, *30*, 8-21. doi:10.1108/JMP-07-2014-0210

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

Matskewich, H. E., & Cheryan, S. (2016). *Comparing stereotypes and attitudes across STEM fields* [data files and codebooks]. Retrieved from osf.io/6gmn5

Merolla, D. M., Serpe, R. T., Stryker, S., & Schultz, P. W. (2012). Structural precursors to identity processes: The role of proximate social structures. *Social Psychology Quarterly*, *75*, 149-172. doi:10.1007/s11218-013-9233-7

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

Miller, D. I., & Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends in Cognitive Sciences*, *18*, 37-45. doi:10.1016/j.tics.2013.10.011

Miller, D. I., & Wai, J. (2015). The bachelor's to PhD STEM pipeline no longer leaks more women than men: A 30-year analysis. *Frontiers in Psychology*, *6*, 37. doi:10.3389/fpsyg.2015.00037

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, *18*, 879-885. doi:10.1111/j.1467-9280.2007.01995.x

National Center for O*NET Development (2018). *O*NET 22.2 Database* [Data files and data dictionary]. Retrieved from https://www.onetcenter.org/db_releases.html

National Science Board (2018). *Science & engineering indicators 2018* (NSB Publication No. 2018-1). Arlington, VA: National Science Foundation. Retrieved from https://www.nsf.gov/statistics/2018/nsb20181/assets/nsb20181.pdf

National Science Foundation (2017). *Cyberlearning for work at the human-technology frontier* (NSF Publication No. 17-598). Alexandria, VA: National Science Foundation. Retrieved from https://www.nsf.gov/pubs/2017/nsf17598/nsf17598.pdf

National Science Foundation (2018). *Building the future: Investing in discovery and innovation – NSF strategic plan for fiscal years (FY) 2018–2022* (NSF Publication No. 18-45). Retrieved from https://www.nsf.gov/pubs/2018/nsf18045/nsf18045.pdf

National Wildlife Federation (2013). Green STEM: An educational collision of epic proportion. Retrieved from http://blog.nwf.org/2013/04/green-stem-an-educational-collision-of-epic-proportion/

New York Hall of Science (2012). Design-make-play: Growing the next generation of science innovators. Retrieved from http://www.yingtrsef.org/wp-content/uploads/DesignMakePlay_Report-2012.pdf

Nix, S., Perez-Felkner, L., & Thomas, K. (2015). Perceived mathematical ability under challenge: A longitudinal perspective on sex segregation among STEM degree fields. *Frontiers in Psychology*, *6*, 530. doi:10.3389/fpsyg.2015.00530

Noonan, R. (2017). *STEM jobs: 2017 update*. *ESA issue brief #02–17*. Washington, DC: U.S. Department of Commerce, Economics and Statistics Administration, Office of the Chief Economist.

North Shore Technology Council (2013). Local STEM education initiatives highlighted at Sept. 25 business breakfast. Retrieved from https://www.nstc.org/previous-events/local-stem-education-initiatives-highlighted-at-sept-25-business-breakfast/

President's Council of Advisors on Science and Technology (2012). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology,* *Engineering, and Mathematics*. Washington, DC: President's Council of Advisors on Science and Technology.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Risen, T. (2016, May 20). Job training a boon for the shifting tech industry. *U.S. News and World Reports*. Retrieved from https://www.usnews.com/news/articles/2016-05-20/job-training-a-boon-for-the-shifting-tech-industry

Saujani, R., & Sweet, J. (2016, December 27). The gender gap in computer science is hurting U.S. businesses. *The Washington Post*. Retrieved from https://www.washingtonpost.com/opinions/the-gender-gap-in-computer-science-is-hurting-us-businesses/2016/12/27/fa4f6680-c247-11e6-9578-0054287507db_story.html

Steinberg, M., & Diekman, A. B. (2018). Considering "why" to engage in STEM activities elevates communal content of STEM affordances. *Journal of Experimental Social Psychology*, *75*, 107-114. doi:10.1016/j.jesp.2017.10.010

Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*, *15*, 250-267. doi:10.1037/a0018719

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

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*, 859-884. doi:10.1037/a0017364

University of Notre Dame (2012). *The first year of studies* (Retrieved March 1, 2014 from http://fys.nd.edu/incoming-students/first-year-requirements/)

Valla, J. M., & 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). *Perspectives on Psychological Science*, *9*, 219-224. doi:10.1177/1745691614522067

Wai, J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2010). Accomplishment in science, technology, engineering, and mathematics (STEM) and its relation to STEM educational dose: A 25-year longitudinal study. *Journal of Educational Psychology*, *102*, 860-871. doi:10.1037/a0019454

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*, 73-82. doi:10.1177/0963721413512856

Wang, X. (2013). Modeling entrance into STEM fields of study among students beginning at community colleges and four-year institutions. *Research in Higher Education*, *54*, 664-692. doi:10.1007/s11162-013-9291-x

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

Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching*, *42*, 36-41.

Whitten, B. L., Dorato, S. R., Duncombe, M. L., Allen, P. E., Blaha, C. A., Butler, H. Z., … , Williams, B. A. (2007). What works for women in undergraduate physics and what we can learn from women's colleges. *Journal of Women and Minorities in Science and Engineering*, *13*, 37-76. doi:10.1615/JWomenMinorScienEng.v13.i1.30

Wine, J., Janson, N., & Wheeless, S. (2011). *2004/09 Beginning postsecondary students longitudinal study (BPS:04/09) full-scale methodology report*. Washington, DC: National Center for Educational Statistics.

Wise, L. L., McLaughlin, D. H., & Steel, L. (1979). *The Project TALENT data bank*. Palo Alto, CA: American Institutes for Research.

Wu, J., & Uttal, D. H. (2017). *Why did students join STEM? A qualitative examination of students' learning experiences*. Unpublished manuscript, School of Education and Social Policy, Northwestern University, Evanston, IL.

Xie, Y., & Killewald, A. A. (2012). *Is American science in decline?* Cambridge, MA: Harvard University Press.

Xie, Y., & Shauman, K. (2003). *Women in science: Career processes and outcomes*. Cambridge, MA: Harvard University Press.

Appendix: Supplemental Methods and Results for Study 1

These supplemental materials provide additional information about the statistical specifications about analyses involving multilevel modeling and propensity score analysis in Study 1.

Longitudinal response rates. Students were excluded if their six-year graduation status and/or later major field of study was unknown. This exclusion resulted in longitudinal response rates of 82% for the BPS sample, 91% for the NLSF sample, and 69% for the Project TALENT samples. The longitudinal response rate was therefore lowest for Project TALENT, but that study's original researchers took steps to address non-response bias. Those researchers estimated characteristics of non-respondents by aggressively tracking down randomly selected samples of participants who did not respond to the standard mail-in surveys. Researchers used phone interviews and credit agencies to locate these initial non-respondents. Final response rates for these initial non-respondents were high (roughly 70-90%). For more details, see pages 17-19 of Wise, McLaughlin, & Steel (1979). STEM joining rates were equal ($P = .75$) among these special respondents ($n = 446$) and normal mail-in survey respondents ($n = 19{,}316$). The relationship between early-college STEM course taking and STEM joining was also the same across respondent type ($P = .73$). Non-respondent bias therefore likely had a minimal influence on our central results for Project TALENT.

Regression models for inferential analyses. We used multilevel logistic regression to model the dichotomous outcome of earning a STEM or non-STEM bachelor's degree (Raudenbush & Bryk, 2002). Students were modeled as nested within institutions (intraclass correlation = .09 for NLSF, assuming a level-1 variance of $\pi^2/3$; see chapter 14 of Snijders & Bosker, 1999). Mixed-effects models assumed that students' log odds of earning a STEM bachelor's degree, relative to earning a non-STEM bachelor's degree, were combinations of fixed effects of predictor variables (e.g., STEM course-taking) and random effects of between-institution heterogeneity. The regression models assumed that the log odds of STEM joining were normally distributed across institutions. Based on a multilevel model with no predictor variables, the estimated log odds at an average institution in NLSF was -2.40 and the estimated between-institution variance was 0.24. We used the *xtmelogit* and *gllamm* commands in Stata to obtain model estimates and used adaptive quadrature as the estimation algorithm (Rabe-Hesketh & Skrondal, 2008).

Within- and between-institution relationships for inferential analyses. Predictor variables varied both within and between institutions. For instance, in the NLSF sample, 11% of the variance in SAT Mathematics scores was between institutions (see intraclass correlations in Table S3). We therefore group-mean centered all student-level predictors to control for this between-institution variance (Raudenbush & Bryk, 2002). Conceptually, multilevel models with group-mean centered predictors will compare students within the same institution and aggregate these comparisons across multiple institutions. This approach aligns with recommendations that local control participants should be used when estimating causal effects (Shadish, 2011). Contrasting the magnitude of within-institution and between-institution relationships can provide additional insight. Although NLSF's small number of institutions (N = 28 institutions) limited the precision of between-institution estimates, Project TALENT was ideal because it included 1,190 four-year institutions (including 762 institutions with five or more participants, 274 institutions with twenty or more participants, and 86 institutions with fifty or more participants).

Survey weights for inferential analyses. We used probability survey weights in some regression models to estimate coefficients that were approximately representative of the sampled undergraduate population (Rabe-Hesketh & Skrondal, 2006). Such weighted regression models explicitly addressed the issue of non-equal sampling probabilities (e.g., racial/ethnic minorities were oversampled in NLSF such that Asian, Black, Hispanic, and White students were each roughly one quarter of the sample). However, using survey weights presented unique challenges for the NLSF sample because the differences in survey weights were large and sample sizes were modest. For instance, survey weights for White students were 8.4 times as large as for Black students on average. Among STEM joiners, White students were 23% of observed counts but 74% of the sum total of survey weights. Hence, weighted models primarily reflected results for the minority of White students in the observed sample. This feature meant that weighted models provided less precise regression estimates than unweighted models. These concerns were heightened because weighting can perform poorly for multilevel logistic regression (Rabe-Hesketh & Skrondal, 2006) and sample sizes were modest in some cases (e.g., n = 98 STEM joiners in NLSF, yielding an effective sample size of n = 31 when using weights). These sampling design issues were less extreme for the BPS and Project TALENT samples.

Comparing results with and without using survey weights. For the reasons described above, we generally placed greater emphasis on unweighted than weighted models when analyzing the NSLF sample. Unweighted models can be seen as reflecting the experiences of a racially diverse, rather than primarily White, college population. Most importantly, we compared the results of unweighted and weighted models to ensure our results were robust. Compared to unweighted models (e.g., Model 4 in Table S5), models that weighted by survey weights generally showed similar results (e.g., Model 5). Early-college STEM course taking continued to predict STEM joining using weights or not. However, for the NLSF sample, the predicted effect of first-semester STEM course taking was twice as strong in the weighted than unweighted model (*b*s = 1.00 vs. 0.50, respectively). The difference between models likely reflected the moderating influence of race. To test this possibility, we added interaction terms (e.g., *ug_STEM* × *Black*) to Model 4 to test for a moderating effect of race/ethnicity. We found that first-semester

STEM course-taking significantly predicted increases in STEM joining for White ($b = 1.15$, $p =$ .0004) and Asian ($b = 0.59$, $p = .008$) students, but not for Black ($b = 0.26$, $p = .23$) or Hispanic ($b = 0.23$, $p = .31$) students. The course-taking effect was significantly greater for Asian/White students than Black/Hispanic students ($\Delta b = 0.62$, $p = .011$).

Hence, even though Black and Hispanic students were as likely as White students to join STEM, first-semester STEM course-taking was a less relevant factor for explaining individual differences in STEM joining among Black and Hispanic students in NLSF. We did not find this same interaction in the BPS sample. In the BPS sample, early-college STEM course taking (the percent of first-year credits earned in STEM department) predicted STEM joining among Asian/White students ($b = 0.054$, $SE = 0.006$, $p < .0001$) and Black/Hispanic ($b = 0.071$, $SE =$ .03, $p = .033$) students. The racial/ethnic difference in regression slopes was not significant ($p =$ 0.62) and was in the opposite direction compared to that for NLSF. The interaction between race/ethnicity and course taking might therefore be confined to selective institutions, but sample sizes for racial/ethnic minorities were too small in the BPS sample (e.g., $n \sim 100$ Black/Hispanic students starting at selective institutions) to rigorously test this possibility using another dataset. The generalizability of the interaction between race/ethnicity and STEM course taking is therefore unclear, but this interaction was important when interpreting and comparing the unweighted and weighted models for specifically the NLSF sample.

Conclusion regarding NLSF's survey weights. The most important result regarding weights was that early-college STEM course taking predicted STEM joined with or without them (see Table S4). Hence, results for early-college STEM course taking were robust. Unweighted models generally yielded smaller estimates of our central treatment effect (i.e., effect of early-college STEM course-taking on STEM joining). Unweighted estimates were therefore likely conservative, and focusing on them therefore reduced the possibility of overstating the policy implications of our results. In the main text, we reported unweighted estimates for the NLSF sample.

References

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*, 399-424. doi:10.1080/00273171.2011.568786

Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, 23, 305-327. doi:10.1002/jae.998

Rabe-Hesketh, S., & Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society: Series A*, *169*, 805-827. doi:10.1111/j.1467-985X.2006.00426.x

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. College Station, TX: Stata.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.

Shadish, W. R. (2011). Randomized controlled studies and alternative designs in outcome studies challenges and opportunities. *Research on Social Work Practice*, *21*, 636-643. doi:10.1177/1049731511403324

Snijders, T., & Bosker, R. (1999). *Multilevel analysis*. London, United Kingdom: Sage Publications.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101, 817. doi:10.1037/a0016127

Table S1

Descriptions of variables included in inferential analyses of the NLSF sample (n = 1,108).

Variable name	Description
Demographic variables	
male	Dummy code for being male. (M = 0.36, SD = 0.48, nonresponse rate = 0%, intraclass correlation = .04, correlation with first-semester STEM course taking = .01).
Asian	Dummy code for being Asian. (M = 0.22, SD = 0.42, nonresponse rate = 0%, intraclass correlation = .01, correlation with first-semester STEM course taking = .11).
Black	Dummy code for being Black. (M = 0.29, SD = 0.45, nonresponse rate = 0%, intraclass correlation = .11, correlation with first-semester STEM course taking = -.04).
Hispanic	Dummy code for being Hispanic. (M = 0.25, SD = 0.44, nonresponse rate = 0%, intraclass correlation = .05, correlation with first-semester STEM course taking = -.03).
parent_STEM	Dummy code for having at least one parent employed in a STEM field one year prior to students' freshman year. (M = 0.11, SD = 0.31, nonresponse rate = 0.8%, intraclass correlation = .00, correlation with first-semester STEM course taking = .04).
SES	Composite index of socioeconomic status. Average of z scores for parents' highest level of education (0 = high school degree or lower, 1 = some college, 2 = college degree, 3 = graduate degree) and household income during students' senior year of high school. (M = 0.00, SD = 0.86, nonresponse rate = 0.1%, intraclass correlation = .05, correlation with first-semester STEM course taking = .04).
High school variables	
ap_STEM	Number of AP STEM courses taken. Possible courses: biology, calculus, chemistry, computer science, environmental science, and/or physics. (M = 1.04, SD = 1.06, nonresponse rate = 0.1%, intraclass correlation = .08, correlation with first-semester STEM course taking = .07).
ap_nonSTEM	Number of AP non-STEM courses. Possible courses: art, economics, English, foreign language, government, music, psychology, U.S. history, and/or world history. (M = 1.96, SD = 1.47, nonresponse rate = 0.1%, intraclass correlation = .15, correlation with first-semester STEM course taking = -.12).
hs_gpa_STEM	Grade Point Average in high school STEM courses (averaged across mathematics and natural science courses). (M = 3.54, SD = 0.52, nonresponse rate = 0.2%, intraclass correlation = .09, correlation with first-semester STEM course taking = .09).
hs_gpa_nonSTEM	Grade Point Average in high school non-STEM courses (averaged across English, foreign language, history, and social studies courses). (M = 3.75, SD = 0.34, nonresponse rate = 0.2%, intraclass correlation = .09, correlation with first-semester STEM course taking = .00).

diff_STEM	Students' ratings of the degree of difficulty of high school STEM courses (averaged across mathematics and natural science courses). Rating scale ranged 0 to 10. (M = 5.48, SD = 2.13, nonresponse rate = 0.1%, intraclass correlation = .03, correlation with first-semester STEM course taking = -.14).
diff_nonSTEM	Students' ratings of the degree of difficulty of high school non-STEM courses (averaged across English, foreign language, history, and social studies courses). Rating scale ranged 0 to 10. (M = 4.08, SD = 1.86, nonresponse rate = 0.1%, intraclass correlation = .00, correlation with first-semester STEM course taking = .04).
SAT_M	SAT – Mathematics score. (M = 643, SD = 89, nonresponse rate = 35.4%, intraclass correlation = .11, correlation with first-semester STEM course taking = -.07).
SAT_V	SAT – Verbal score. (M = 650, SD = 84, nonresponse rate = 35.2%, intraclass correlation = .19, correlation with first-semester STEM course taking = -.10).
Fall freshman variables	
ug_STEM	Number of STEM courses completed during the first semester or quarter of college. (M = 1.20, SD = 1.05, nonresponse rate = 2.6%, intraclass correlation = .17).
init_premed	Dummy code for having declared a pre-medicine major during freshman year. (M = 0.08, SD = 0.27, nonresponse rate = 0%, intraclass correlation = .04, correlation with first-semester STEM course taking = .40).
Institution-level variables	
ug_STEM_mn	Mean number of STEM courses taken by schools' non-STEM intenders during the first semester or quarter of college.
premed_intensive	Dummy code for attending a school with a high proportion (above a median split) of premed intenders among non-STEM intenders.
STEM_intensive	Dummy code for attending a STEM-intensive school. Defined by a high proportion (above a median split) of freshmen who intended a STEM major (excluding undecided students).
public_uni	Dummy code for attending a public university.
liberal_arts	Dummy code for attending a liberal arts college.
most_selective	Dummy code for attending one of the more selective NLSF institutions. Defined by a high mean total SAT score (above a median split among NLSF institutions).
Outcome variable	
final_STEM	Dummy code for later joining STEM. (M = 0.09, SD = 0.28, intraclass correlation = .09, correlation with first-semester STEM course taking = .28).

Table S2

Descriptions of variables included in inferential analyses of the Project TALENT sample (*n* = 19,762).

Variable name	Description
Demographic variables	
male	Dummy code for being male. (*M* = 0.40, *SD* = 0.49, nonresponse rate = 0%, intraclass correlation = .43, correlation with early-college STEM course taking = .03).
Asian	Dummy code for being Asian. (*M* = 0.01, *SD* = 0.10, nonresponse rate = 0.5%, intraclass correlation = .59, correlation with early-college STEM course taking = .03).
racial_minor	Dummy code for being a non-Asian racial/ethnic minority. These racial/ethnic minorities were collapsed into this category because of small sample sizes. (*M* = 0.02, *SD* = 0.15, nonresponse rate = 0.5%, intraclass correlation = .86, correlation with early-college STEM course taking = .05).
parent_STEM	Dummy code for having at least one parent employed in a STEM field when students were first sampled in high school. (*M* = 0.06, *SD* = 0.23, nonresponse rate = 4.0%, intraclass correlation = .06, correlation with early-college STEM course taking = .00).
SES	Composite index of socioeconomic status, created by the original Project TALENT researchers. Based on parents' highest level of education, household income, and presence of certain household items (e.g., television set). See Wise et al. (1979) for details. We computed *z* scores for this variable using the mean and standard deviation of the entire Project TALENT sample (e.g., based on the statistics below, the socioeconomic status of our analyzed subsample was 0.77 SDs above the 1960 high school population). (*M* = 0.77, *SD* = 0.85, nonresponse rate = 3.2%, intraclass correlation = .29, correlation with early-college STEM course taking = -.04).
High school variables (analogous to those in NLSF)	
hs_STEM_crs	Percent of high school courses taken in math and science. (*M* = 39.5, *SD* = 9.6, nonresponse rate = 7.2%, intraclass correlation = .14, correlation with early-college STEM course taking = .11).
hs_gpa_STEM	Grade Point Average in high school STEM courses (averaged across mathematics and science courses). (*M* = 3.24, *SD* = 0.66, nonresponse rate = 4.4%, intraclass correlation = .13, correlation with early-college STEM course taking = .07).
hs_gpa_nonSTEM	Grade Point Average in high school non-STEM courses (averaged across English, foreign language, history, and social studies courses). (*M* = 3.38, *SD* = 0.57, nonresponse rate = 4.3%, intraclass correlation = .12, correlation with early-college STEM course taking = .01).
verbal_conf	Composite measure of students' verbal confidence (average on 5-point Likert scale). Based on seven items (α = 0.76). Example item =

	"I have a difficult time expressing myself in written reports, examinations, and assignments" (reverse-coded). No comparable measure was available for mathematics confidence. (M = 3.90, SD = 0.66, nonresponse rate = 3.2%, intraclass correlation = .05, correlation with early-college STEM course taking = -.03).
math_test	Composite measure of standardized mathematics performance. Based on four subtests. See Wai et al. (2009) for details. We computed z scores for this variable using the mean and standard deviation of the entire Project TALENT sample. (M = 0.95, SD = 0.84, nonresponse rate = 0%, intraclass correlation = .32, correlation with early-college STEM course taking = -.04).
verbal_test	Composite measure of standardized verbal performance. Based on three subtests. See Wai et al. (2009) for details. We computed z scores for this variable using the mean and standard deviation of the entire Project TALENT sample. (M = 0.96, SD = 0.62, nonresponse rate = 0%, intraclass correlation = .39, correlation with early-college STEM course taking = -.02).

Early-college STEM course-taking[*]

ug_STEM	The number of STEM departments that students had taken courses in by the time of first longitudinal follow-up (~1.5 years after high school graduation). Range = 0 to 4 (possible departments = biological science, engineering, mathematics, physical science). Information on the number of courses taken was not available. However, in the NLSF sample, we found that the number of first-year STEM courses correlated highly with the number of first-year STEM departments (r = .83, where the categories of "departments" matched the categories available for Project TALENT: biological science, mathematics, physical science, and other). (M = 1.49, SD = 0.85, nonresponse rate = 1.6%, intraclass correlation = .26).

High school variables (not in NLSF)

hs_STEM_degree	Dummy code for intending a STEM college major when students were in high school. (M = 0.18, SD = 0.39, nonresponse rate = 4.9%, intraclass correlation = .07, correlation with early-college STEM course taking = .12).
hs_STEM_both	Dummy code for both intending a STEM college major and future STEM employment when students were in high school. (M = 0.12, SD = 0.32, nonresponse rate = 4.1%, intraclass correlation = .10, correlation with early-college STEM course taking = .10).
STEM_interest	Composite measure of students' interests in STEM activities (10 items) and STEM occupations (14 items). The items about activities

[*]Note that there was no variable for intending a pre-medicine major because the survey form for the first follow-up did not provide a response category for intending a health/medicine major. The closest response option was "biological science," which would have excluded students from the present analysis if they had selected it.

	asked students to indicate "how much you like or would like each of the following" and listed activities such as "study physics" or "learn about diseases." Students completed similar items about STEM occupations such as "chemist" and "research scientist." When answering items about occupations, students were instructed to, "assume that you would have any necessary training or education that would be required. Disregard salary, social standing, permanence, etc., in fact anything except how well you would like to do the work." Project TALENT researchers created subscales for interests in physical science and biological science/medicine (see Wise et al., 1979 for more details), and we averaged *z*-scores for those subscales to form the *STEM_interest* variable. (*M* = 0.22, *SD* = 0.84, nonresponse rate = 0.4%, intraclass correlation = .05, correlation with early-college STEM course taking = .10).
nonSTEM_interest	Composite measure of students' interests in non-STEM activities (22 items) and occupations (32 items). Based on averaging *z*-scores to Project TALENT-created subscales for artistic, business, literary-linguistic, musical, and social service interests. Example activities = help the poor, manage a large store, play an instrument, teach children. Example occupations = banker, lawyer, musician, office manager. (*M* = 0.34, *SD* = 0.70, nonresponse rate = 0.3%, intraclass correlation = .05, correlation with early-college STEM course taking = -.08).
STEM_info	Composite measure of students' knowledge of STEM-related information. Based on averaging *z*-scores to Project TALENT-created subscales for aeronautics and space, biological science, engineering, electricity and magnetism, and physical science domains (65 total items). We interpret this variable as a behavioral indicator of students' interests in STEM because many of the items focused on information that students would have learned outside of formal instruction (see Wise et al., 1979 for further description). High scores on this variable would therefore in part reflect students' behaviors to learn STEM-related information outside of formal instruction (e.g., following news coverage about Sputnik's launch in 1957). **Example items:** The minimum speed a rocket needs in order to get beyond the range where gravity will pull it back to earth is called (A) thrust, (B) anti-gravity pull, (C) gravitational acceleration, (D) escape velocity, or (E) gravity cancellation speed. On some electrical equipment, the plug has a third prong. This provides (A) a boost in current, (B) an increase in voltage, (C) a ground, (D) transfer from AC to DC, or (E) transfer from DC to AC. (*M* = 0.44, *SD* = 0.69, nonresponse rate = 0%, intraclass correlation = .21, correlation with early-college STEM course taking = .04).

nonSTEM_info	Composite measure of students' knowledge of information in the non-STEM domains of art, business, law, literature, music, and social studies. Based on averaging *z*-scores to Project TALENT-created subscales (92 total items). (*M* = 0.76, *SD* = 0.63, nonresponse rate = 0%, intraclass correlation = .30, correlation with early-college STEM course taking = -.08).
spatial_test	Composite measure of standardized spatial performance. Based on four subtests. See Wai et al. (2009) for details. We computed *z* scores for this variable using the mean and standard deviation of the entire Project TALENT sample. (*M* = 0.48, *SD* = 0.82, nonresponse rate = 0%, intraclass correlation = .16, correlation with first-semester STEM course taking = .04).
Institution-level variables	
ug_STEM_mn	The means of the *ug_STEM* variable at students' institutions.
Other variables	The means of other variables at students' institutions (these variables were *male*, *parent_STEM*, *hs_STEM_crs*, *hs_gpa_STEM*, *math_test*, *hs_STEM_degree*, *STEM_interest*, *nonSTEM_interest*, *STEM_info*, and *spatial_test*). These particular variables were selected for the institution-level model because they predicted STEM joining at the student level.
Outcome variable	
final_STEM	Dummy code for later joining STEM. (*M* = 0.02, *SD* = 0.14, intraclass correlation = .16, correlation with first-semester STEM course taking = .08).

Table S3

Educational backgrounds and trajectories of STEM joiners, compared to students from other educational pathways (BPS sample). Values on the top are weighted mean estimates and values in parentheses are standard errors. $^{***}p < .001$. $^{**}p < .01$. $^{*}p < .05$. $^{\dagger}p < .10$.

		Comparison Groups		
	STEM joiners versus…	STEM persisters	Non-STEM persisters	STEM switch-outs
High school				
Took calculus	49% (5%)	59%† (3%)	25%*** (1%)	38%† (3%)
Took 4 years or more of science	57% (5%)	67%† (2%)	50% (1%)	65% (4%)
SAT Mathematics score	570 (10)	610*** (6)	529*** (3)	550 (8)
SAT Verbal score	536 (8)	570*** (5)	537 (3)	542 (7)
Composite GPA	3.51 (0.035)	3.57 (0.015)	3.43* (0.015)	3.48 (0.03)
First year in college				
Number of STEM credits earned	12.7 (0.8)	18.7*** (0.4)	6.1*** (0.1)	11.1† (0.6)
Number of total credits earned	30.2 (0.6)	31.9* (0.4)	29.4 (0.2)	27.8** (0.7)
% of credits that were earned in STEM	42% (2%)	58%*** (1%)	21%*** (0.5%)	40% (2%)
Earned credit in calculus/advanced math	41% (5%)	68%*** (3%)	14%*** (1%)	38% (4%)
Withdrawn or failed a STEM course	12% (3%)	9% (2%)	10% (1%)	15% (3%)
STEM GPA	3.01 (0.07)	3.13 (0.04)	2.96 (0.03)	2.71*** (0.06)
Non-STEM GPA	3.27 (0.05)	3.36 (0.03)	3.21 (0.02)	3.16† (0.05)

Cumulative outcomes in college

Graduate within 4 years	55% (6%)	58% (3%)	62% (1%)	48% (4%)
Graduate within 5 years	91% (2%)	93% (1%)	93% (1%)	90% (3%)
Cumulative STEM GPA	3.02 (0.05)	3.12† (0.03)	2.94 (0.02)	2.77*** (0.05)
Cumulative non-STEM GPA	3.32 (0.05)	3.39 (0.03)	3.28 (0.02)	3.23† (0.03)
% withdrawn/failed STEM courses out of attempted	2.5% (0.3%)	1.8%† (0.2%)	3.8%** (0.4%)	5.2%*** (0.6%)
Number of total STEM courses taken	26.4 (1.1)	33.7*** (0.8)	6.9*** (0.2)	14.3*** (0.8)
Number of total non-STEM courses taken	25.2 (1.2)	18.9*** (0.5)	42.0*** (0.4)	35.9*** (0.8)
Number of STEM credits earned	71.1 (2.9)	90.1*** (1.9)	19.4*** (0.4)	38.1*** (2.1)
Graduated with double major	15.2 (2.8)	8.1 (1.3)	13.4 (0.9)	8.3 (2.1)
Sample size	~200	~700	~2,700	~300

Table S4

Unstandardized coefficients for multilevel logistic regression models of STEM joining (NLSF sample, n = 1,108). The coefficients for the SAT variables have been multiplied by 100 to facilitate presentation of results. Models 1-4 were unweighted and Model 5 weighted students by probability survey weights. ***$p < .001$. **$p < .01$. *$p < .05$. $^{\dagger}p < .10$.

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5
Demographic variables					
male	0.51*	0.40	0.42	0.39	0.68
Asian	0.53†	0.36	0.04	0.03	-0.36
Black	-0.37	-0.26	-0.43	-0.45	-0.28
Hispanic	-0.07	-0.03	-0.20	-0.22	-0.88†
parent_STEM	0.46	0.45	0.36	0.40	-0.30
SES	-0.08	-0.13	-0.17	-0.18	-0.44
High school variables					
ap_STEM		0.33**	0.20	0.20	0.19
hs_gpa_STEM		0.19	0.16	0.19	0.40
diff_STEM		-0.16*	-0.07	-0.06	0.11
SAT_M		-0.19	0.11	0.17	0.22
ap_nonSTEM		-0.02	0.08	0.08	0.29
hs_gpa_nonSTEM		0.08	-0.34	-0.39	-0.35
diff_nonSTEM		0.13†	0.02	0.02	0.13
SAT_V		0.10	0.02	-0.01	0.20
Fall freshman variables					
ug_STEM			0.50***	0.50***	1.00***
init_premed			2.03***	1.97***	2.22***
Institution-level variables					
ug_STEM_mn				0.46	2.07**
premed_intensive				0.92*	1.46
STEM_intensive				0.75**	2.21**
public_uni				-0.62†	-0.83
liberal_arts				0.41	1.57†
most_selective				0.16	0.44

Table S5

Unstandardized coefficients for multilevel logistic regression models of STEM joining (Project TALENT sample, n = 19,762). Models 1-5 were unweighted and Model 6 weighted students by probability survey weights. The coefficients for the *hs_STEM_crs* variable has been multiplied by 10 to facilitate presentation of results. All models also included dummy codes for the high school cohort (e.g., originally tested in 9th grade vs. 10th grade in high school). We also estimated models separately for each high school cohort and found no differences in regression coefficients for *ug_STEM* across cohorts. ***p < .001. **p < .01. *p < .05. $^{\dagger}p$ < .10.

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Demographic variables						
male	1.15***	1.04***	0.97***	0.22	0.25	0.18
Asian	0.74	0.60	0.48	0.25	0.20	0.24
racial_minor	0.20	0.47	0.53	0.51	0.52	0.70
parent_STEM	0.43*	0.42*	0.42*	0.32	0.32	0.42†
SES	-0.06	-0.04	-0.05	-0.06	-0.05	-0.10
High school variables (analogous to those in NLSF)						
hs_STEM_crs		0.26***	0.26***	0.15*	0.15*	0.14*
hs_gpa_STEM		0.47***	0.42**	0.16	0.17	0.13
math_test		0.35***	0.30**	0.12	0.13	0.13
hs_gpa_nonSTEM		-0.14	-0.11	0.16	0.09	0.10
verbal_conf		-0.12	-0.10	-0.10	-0.11	-0.12
verbal_test		-0.14	-0.12	-0.17	-0.14	-0.08
Early college STEM course-taking						
ug_STEM			0.51***	0.38***	0.37***	0.37***
High school variables (not in NLSF)						
hs_STEM_degree				0.52**	0.49**	0.49**
hs_STEM_both				-0.14	-0.18	-0.11
STEM_interest				0.55***	0.56***	0.57***
nonSTEM_interest				-0.34***	-0.34***	-0.28**
STEM_info				0.31*	0.29*	0.28*
nonSTEM_info				-0.29†	-0.34*	-0.37*
spatial_test				0.34***	0.36***	0.38***

Institution-level variables		
ug_STEM_mn	0.34*	0.39**
male_mn	-0.29	-0.39
parent_STEM_mn	0.61	0.58
hs_STEM_crs_mn	0.21†	0.20*
hs_gpa_STEM_mn	-0.03	-0.11
math_test_mn	0.24	0.33
hs_STEM_degree_mn	1.29*	1.51**
STEM_interest_mn	0.54†	0.58*
nonSTEM_interest_mn	-0.67†	-0.81*
STEM_info_mn	-0.82*	-0.83*
spatial_test_mn	1.27***	1.17***

Table S6

Additional results about the effect of early-college STEM course taking on later STEM joining.

Key result	Additional detail
1. Diverse pre-college variables were weak predictors of first-semester STEM course-taking among non-STEM intenders.	The self-selection hypothesis predicted that, among non-STEM intenders, endogenous factors such as prior STEM course-taking and STEM attitudes should have explained variance in first-semester STEM course-taking. However, among non-STEM intenders, the correlations between these pre-college factors and first-semester STEM course-taking were small (all observed *r*s < .15, Table S3). For instance, among non-STEM intenders in NLSF, only 4.7% of the observed variance in first-semester STEM course-taking was explained by gender, race/ethnicity, socioeconomic status, parents' employment in STEM fields, SAT scores, high school AP STEM and non-STEM course-taking, high school STEM and non-STEM grades, and perceived difficulty of STEM and non-STEM courses (see Table S3). Analogous variables in Project TALENT explained only 4.3% of the variance in early-college STEM course-taking; including other variables such as occupational plans and prior interests in STEM only explained another 1.2% of the total variance (see Table S4).
2. Pre-college variables were weaker predictors of first-semester STEM course-taking among non-STEM intenders than among undecided students.	Pre-college variables in NLSF explained 12.6% of first-semester STEM course-taking among undecided students. Hence, self-selection due to observed covariates among undecided students was nearly three times as strong as among non-STEM intenders. This finding supports the notion that self-selection was weaker among non-STEM intenders.
3. Compared to pre-college variables, institutional factors explained more variance in first-semester STEM course-taking among non-STEM intenders.	Multilevel models indicated that 17% of the observed variance in first-semester STEM course-taking (among non-STEM intenders) was explained by between-institution heterogeneity. This heterogeneity likely reflected, in part, external factors such as institutional requirements for first-year courses and graduation.
4. Causal estimates were robust to plausible	We used simulation-based methods, initially developed by Ichino, Mealli, and Nannicini (2008), to evaluate the potential consequences of unobserved confounders. These methods flexibly simulated how hypothetical confounders would have influenced selection into

confounders that were not measured.	treatment (STEM course-taking) and potential treatment outcomes (STEM joining). These simulations then used propensity score matching (Austin, 2011) to quantify how such confounders would have changed our causal estimates of STEM course taking. We used the *sensatt* command in the statistical software Stata to conduct these analyses (19). We describe the details below. • **Notation:** We used the same notation as Ichino et al. (2008). The variable U was an unobserved confounder, the *selection effect* Λ was the average odds ratio of U in logistic regression models predicting treatment status, and the *outcome effect* Γ was the average odds ratio of U in models predicting the treatment outcome. We defined the *confound effect* as $\ln(\Lambda) \times \ln(\Gamma)$, which would equal 0 if the variable U did not relate to selection ($\Lambda = 1$) or the outcome ($\Gamma = 1$). • **Model assumptions:** The unobserved confounder U was assumed to be binary (i.e., $U = 0$ or $U = 1$), moderately prevalent ($U = 1$ for 40% of cases, on average), and uncorrelated with other predictor variances. Using Monte Carlo simulations and empirical examples, Ichino et al. showed that their simulation method yielded similar results for alternate specifications (e.g., assuming a continuous confounder or different prevalences for $U = 1$). • **Simulation parameters:** We varied the selection effect Λ and treatment effect Γ by manipulating the probabilities that $U = 1$ in the four treatment-outcome cells (e.g., the probability that $U = 1$ among untreated STEM joiners). We used the same sets of parameter values that Ichino et al. used (e.g., the mean difference in U between treatment and untreated individuals varied from 0.1 to 0.7). These values covered a wide range of confound effects. We ran 1,000 simulations for 49 sets of parameter values (in total, 49,000 simulations). • **Results:** The confound effect, $\ln(\Lambda) \times \ln(\Gamma)$, linearly related to causal estimates: *ATT* = 5.1% – 0.9% × *confound effect*. Regression model statistics ($R^2 = .98$, RMSE = 0.4%) and visual inspection of scatterplots indicated that a linear model well summarized the data (e.g., it explained 98% of the variance in estimates). • Based on these simulations, a confound effect of 2.8 would have reduced our ATT causal estimate by 50% and a confound effect of 5.6 would have reduced the estimate to 0. Such confound effects are implausibly large. For instance, the confound effect for the perceived difficulty of STEM courses (variable name = *diff_STEM*) was 0.30. In Project TALENT, the confound effect for intending a college STEM major while in high school (variable name = *hs_STEM_degree*) was 0.57. Given that these observed

	covariates had some of the largest confound effects in NLSF and Project TALENT, an unobserved confound effect of 2.8 was unlikely. • A confound effect of 2.8 would have required that the unobserved confounder *U* quintupled the odds of selecting treatment (i.e., $\Lambda = 5$), assuming equal selection and outcome effects (i.e., $\Lambda = \Gamma$). As discussed earlier (Results 1–3), such large self-selection effects are implausible. Confounders with smaller selection effects (i.e., $\Lambda < 5$) could have had the same confound effect, but only if the outcome effect was larger (i.e., $\Gamma > 5$). • **Conclusion:** Our results were robust to plausible hidden biases; only unobserved variables with implausibly large confounding effects would have substantially reduced our causal estimates.

Table S7

Calculations about the hypothetical scenarios described in the main text.

Quote from the main text	Additional explanation
"For instance, we estimated that increasing the joining rate by 5 percentage points would generate between 26,000 to 63,000 more STEM graduates per year, depending on the definition of potential STEM joiners."	**Strictest definition of STEM joiners** = STEM bachelor's degree earners who initially intended a non-STEM major and started postsecondary education at a 4-year institution (this definition is the primary focus of this current manuscript and does not include beginning 2-year students, initially undecided students, or students whose highest degree was an associate's degree) **Broadest definition of STEM joiners** = STEM bachelor's degree earners or STEM associate's degree earners who were initially undecided or intended a non-STEM major and started postsecondary education at a 4-year or 2-year institution. For students who earned both a bachelor's and associate's degree, we used only the field of study for the bachelor's degree. Hence, students who earned a STEM associate's degree and non-STEM bachelor's degree would not be considered STEM joiners because their highest earned degree was not in STEM. This definition also meant that we did not double-count students who earned both a STEM associate's degree and STEM bachelor's degree. **Example calculation for strictest definition:** Pool of potential STEM joiners = bachelor's degree earners who had initially intended a non-STEM major and started at a 4-year university = 518,600 Effect of increasing joining rate by 5% = 518,600 × 5% = 26,000 more STEM bachelor's degrees per year These estimates are likely conservative because the total number of U.S. undergraduates has increased since the BPS study. For instance, based on data from the Integrated Postsecondary Education Data System, the number of first-time freshman enrolled at 4-year institutions increased by 11% from 2004 to 2012.
"Increasing women's rate to join pSTEM fields to match men's would generate 38%	Men's pSTEM joining rate = 5.0% Women's pSTEM joining rate = 1.8% Difference in rates = 3.2%

more female graduates in these male-dominated fields"	Number of female graduates who initially intended a non-STEM major = 337,700 Number of additional female pSTEM graduates generated by increasing women's joining rate to match men's = 337,700 × 3.2% = 10,806 Number of current female pSTEM graduates = 28,300 Percent increase in female pSTEM graduates = 10,806/28,300 = 38% We adjusted the numbers in Figure 7 by a multiplicative constant so that numbers representing "current supply" match the number of pSTEM bachelor's degrees awarded in 2012. This adjustment was based on population-level data from Integrated Postsecondary Education Data System.
"For instance, women currently earn 25% of the U.S.'s pSTEM bachelor's degrees, and 'plugging' the leaky pSTEM pipeline for female undergraduates would only increase this percentage to 27%"	Men's pSTEM switch-out rate = 30.3% Women's pSTEM switch-out rate = 41.1% Difference in rates = 10.8% Weighted sample size of female graduates who initially intended pSTEM major = 24,400 Number of additional female pSTEM graduates generated by increasing women's switch-out rate to match men's = 24,400 × 10.8% = 2,635 Number of current female pSTEM graduates = 28,300 Number of current male pSTEM graduates = 85,400 Current percentage of females among pSTEM graduates = 28,300/(28,300 + 85,400) = 25% [NOTE: this estimate of 25% aligns with current population-level data available through the Integrated Postsecondary Education Data System] Percent of females among pSTEM graduates if women's switch-out rate was changed to match men's = (28,300 + 2,635)/(28,300 + 2,635 + 85,400) = 27%