Women have made marked gains in higher educational attainment since the 1980s, now outpacing men in both college attendance and completion (Conger and Dickson 2017; DiPrete and Buchmann 2013). However, women's progress in higher education has not translated into gender parity across academic disciplines (England and Li 2006). Despite receiving higher grades in high school than their men counterparts (Conger and Long 2010), women remain less likely to major in science, technology, engineering, and mathematics (STEM) fields (National Center for Science and Engineering Statistics 2021). Although women have made progress in some STEM disciplines, notably biology, they remain underrepresented in disciplines such as engineering, physics, and computer science and continue to face gendered stereotypes and unwelcoming environments (Blackburn 2017; Ceci et al. 2014; Sax et al. 2016). In short, the stubborn gender segregation of academic fields has important implications for broader patterns and manifestations of gender inequality in the United States.

Scholars and practitioners alike have sought to intervene on this issue. For example, since 2001, the National Science Foundation has invested nearly $300 million to support projects that encourage the “representation and advancement of women in academic science and engineering careers” (National Science Foundation 2022). Previous scholarly inquiry has suggested that high school academic achievement (Crisp, Nora, and Taggart 2009; Tyson et al. 2007), self-assessed math ability (Green and Sanderson 2018), family-work orientation (Sassler et al. 2017), and occupational planning (Ganley et al. 2018) are key predictors of STEM persistence in college. Once in college, factors positively

### Abstract

How does gender inform initial academic commitments and narrative self-presentation in science, technology, engineering, and mathematics (STEM) fields during the college application process? Analyzing 60,000 undergraduate applications to the University of California, the authors surface two key findings. First, extant gender segregation of academic disciplines also manifests in intended major choice. Additionally, gender and SAT Math scores together strongly predict intent to major in biology and engineering, the most popular and gender-segregated majors. Second, using natural language processing, the investigators find that author gender is more predictive of essay topics written by prospective engineers than prospective biologists. Specifically, women intending to major in engineering write about essay topics that signal their gender identity to a greater degree than women intending to major in biology, perhaps to mitigate gender-transgressive academic commitments. The authors subsequently argue that prescriptive and proscriptive ideas about men and women’s academic choices remain highly salient in a moment of imagining future academic and professional selves.

### Keywords

gender, STEM, college applications, computational methods, higher education
affecting STEM persistence include performance in early math and science courses, research experiences, and institutional supports such as tutoring and training programs (Chang et al. 2008; Estrada et al. 2016; Hurtado et al. 2007; Vieyra, Gilmore, and Timmerman 2011), particularly for women and underrepresented students of color (Espinosa 2011).

Missing from this literature, to our knowledge, is any exploration of the consequential process that links these pre- and postcollegiate experiences: the college application process. The admissions process carries clear implications for who enters STEM pathways—by design, it selects and rejects applicants—but it also obliges applicants to make preliminary academic declarations that carry important implications for major selection and degree completion (Wolniak 2016). Furthermore, the college application calls upon applicants to postulate about their academic and occupational futures a process which activates gendered perceptions of ability (Correll 2004) and cultural beliefs about gender and academic fit (Charles and Bradley 2009). Recent findings by Weeden, Gelbgiser, and Morgan (2020) revealed that gender differences in occupational planning have a strong association with gender gaps in STEM. Therefore, examining the initial academic commitments college applicants make, as well as how they describe their academic interests and ambitions, is important to further our understanding of persistent gender segregation in academic disciplines and related occupations, before students embark on what can often be inflexible academic pathways through college (Armstrong and Hamilton 2013; Chambliss and Takacs 2014).

We have a rare opportunity to address this gap in the literature. Leveraging a data set of 60,000 applications to the University of California system in the 2016–2017 academic year, we investigate how gender informs applicant behavior during the college admissions process. The University of California system is the largest public university system in the United States, attracting nearly 250,000 undergraduate applications in the 2021–2022 admissions cycle (University of California Office of the President 2021) and enrolling 280,000 undergraduates across all of its nine undergraduate campuses (University of California 2021). Our data provide uncommon insight into applicant choices and behavior at the juncture between secondary and tertiary education, capturing student self-presentation as they reflect on their previous academic experiences and narrate, in their own words, who they intend to be in college. How might gender influence this process of self-making?

We posit that gender will inform applicant decision making and self-presentation at two important points within the application: (1) the choice of intended academic major and (2) the crafting of personal essays describing one’s academic interests. The former necessitates self-evaluation of one’s academic achievement and potential in certain fields, a decision-making process that operates in markedly different ways for men and women (Correll 2004). The latter presents an opportunity for applicants to provide context, justification, and richness to their academic ambitions, providing fertile ground for investigating how applicants simultaneously make sense of their gender and academic identities in narrative form. We speculate that gender will emerge to varying degrees in personal essays, either signaled, because of gender’s function as a primary frame (Ridgeway 2011), or suppressed, as a safeguard against potential tokenization in academic disciplines where women are numerically rare (Kanter 1977).

We have two empirical aims. First, we use descriptive analysis to surface gender patterning in applicant intended major choice, examining the degree of gender segregation in academic disciplines at a novel site along the educational trajectory. Perceptions of ability are linked to academic and occupational aspirations in gendered ways, especially in math contexts (Correll 2001). We therefore complement this analysis with an examination of how the relationship between SAT Math scores and intended major varies for men and women in fields with differing gender compositions and cultural associations. Second, we use natural language processing (NLP) to investigate how gender informs the construction of personal essays narrating academic interest in engineering and biology. We focus on engineering and biology because they are two STEM subfields with starkly different gender compositions, with women making up the majority of biology majors nationally but an acute minority of engineering majors (Alon and DiPrete 2015). Focusing on these two disciplines allows us to investigate women’s gendered narrative self-presentation in gender-typical and gender-atypical STEM contexts.

Collectively, our study offers two key empirical contributions. First, we offer novel insight into how prevailing beliefs about gender and academic disciplines manifest in intended major choice, demonstrating that extant gender segregation of academic disciplines is also borne out in applicants’ preliminary academic commitments. Second, we demonstrate how gender informs women’s narrations of academic interests, especially for those crossing gendered boundaries. We find that women writing in engineering contexts write about topics that implicate their gender identity to a greater degree than women writing in biology contexts, perhaps in a bid to mitigate gender-transgressive disciplinary choices. These findings suggest that prevailing cultural norms around men and women’s academic choices remain highly salient for young people as they imagine their future academic and professional selves. Our study yields important insights for sociologists of gender, higher education, and stratification, as well as for educators and practitioners in secondary and post-secondary education already working toward gender inclusivity in STEM (Posselt 2020).
Gender informs the development of one’s academic identity across the educational life course. Much literature has documented how educational institutions are key sites at which signals about gender, gender performance, and academic aptitude are communicated and learned. Gendered associations about academic capacity and disciplinary fit are learned as early as in the first grade (Master, Meltzoff, and Cheryan 2021) and develop over the course of the educational life trajectory: middle school students are primed to perceive their White, male peers to be exceptional (Musto 2019), high school students develop perceptions of gender and math abilities that correlate with major intention (Riegle-Crumb and Peng 2021), the transition to college prompts students to consider their gender identity as they choose everything from initial courses to housing accommodations (Armstrong and Hamilton 2013), and selecting a college major can be a purposeful signal of femininity or masculinity with implications for persistence in academic disciplines (Mullen 2014; Riegle-Crumb, King, and Moore 2016).

These critical decisions during formal education extend into individuals’ occupational trajectories. Although major intention does not always translate to occupational ambition (Mullen 2014), prior literature suggests that occupational ambitions and major intentions are strongly correlated (Torre 2014; Weeden et al. 2020). Academic major is an important signal as college graduates enter the labor market, and gendered associations of academic fields extend into gendered associations of occupations (Cech 2015). Pointedly, Cech (2013) demonstrated how students of varying gender ideologies nearly uniformly adhere to gendered scripts when considering future careers. In short, prior literature has thoroughly documented how the cultivation of academic and gender identity is punctuated by key decision-making points: which courses to take, which major to declare, whether to persist in an academic field, and which kind of occupational field to pursue.

We consider a novel site of academic commitments along the educational trajectory: the choice of intended major during the college admissions process. Decisions around coursework and major intention made early in one’s collegiate career create path dependencies that set students on academic pathways that are difficult to deviate from (Chambliss and Takacs 2014; Wolniak 2016). Initial courses, particularly in STEM disciplines, can often stipulate whether a student continues in that major discipline (Thompson 2021), as well as how they view their potential to succeed within a course of study (Harrison, Hernandez, and Stevens 2022). Lang et al. (2021) documented how initial course selection is more than 30 times more predictive of major declaration than random guessing. If the college application is a site where students imagine commitments to certain academic disciplines, then intended major selection has important implications for the courses students ultimately take, the majors they declare, and in all likelihood, their ultimate career aspirations. Applicants to the University of California are asked to assert their academic interests in two forms: choosing an intended major and in an essay where applicants discuss an academic subject that inspires them. We attend to how gender informs each of these elements of the application in turn.

Gender, Achievement, and Intended Major

Selecting an intended major is not only a preliminary statement of academic identity but also a self-assessment of one’s “fit” with an academic discipline, including one’s potential to be academically successful (Porter and Umbach 2006). These moments of self-assessment and identity declaration implicate one’s gender identity, as well as prevailing beliefs about men and women. As Ridgeway (2011) argued, gender is a primary category by which we unconsciously understand ourselves, others, and how to structure social interactions. By extension, cultural beliefs about gender are ubiquitous and powerful, prescribing and proscribing norms, scripts, and status beliefs about men and women, including which majors are better suited to men and women.

Academic fields vary both in their gender demography and in the cultural meanings associated with them. Women have long been overrepresented in the humanities and social sciences and underrepresented in STEM fields; the humanities are understood as feminine fields by virtue of being more interpretative and subjective, whereas STEM fields are understood in positivist terms, reflecting a long-held association between positivism and masculinity (Leslie et al. 2015; Sprague 2016). These cultural associations provide encouragement or discouragement to students as they consider which academic disciplines would be appropriate for them to pursue (Beasley and Fischer 2012; Beutel, Burge, and Borden 2017). In an influential analysis, Charles and Bradley (2009) found that gender segregation of fields of study is even stronger in industrialized societies where individuals have ostensibly more choice; they argue this is evidence of individuals “indulging” their gendered selves and making academic choices that comport with prevailing gender beliefs.

Cultural beliefs about what men and women are better suited to study not only prescribe what individuals should pursue; they also proscribe what they should not. Prior literature has documented how men resist fields typed as feminine (Mullen 2014), a finding consistent with England and Li’s (2006) analysis of gender segregation in college majors: whereas women began to move into men-dominated disciplines such as business, men did not contribute to gender integration by moving into women-dominated disciplines. In short, the decision to enter a certain academic field is not only an expression of one’s academic identity, but also of one’s gender identity.
The choice of an intended major is also a form of academic self-evaluation. Applicants assess their academic accomplishments and aptitudes, and translate them into intended majors in which they believe they will be successful. Much literature has documented how men and women differentially react to signals of academic ability. Owen (2008) demonstrated how women are more likely to perceive “A” grades on final exams as signals of encouragement to persist in the field, whereas men require no such signal of encouragement. Similarly, Goldin (2013) documented how men are likely to persist in a chosen field of study, despite poor grades, whereas women take poor grades as cues to exit. Signals of ability are especially important in areas where beliefs about gender status are salient. Correll’s (2004) study demonstrated how women underrated their abilities and men overrated their abilities when primed to believe a task was more suited to men; these ratings of ability also correlated to aspirations in related careers, with men having higher ambitions. Legewie and DiPrete (2014) argued that context strongly influences the salience of gender in STEM career aspirations and subsequent decision making. The college application process, a context which emphasizes prior academic performance—grade point average (GPA), curricular rigor, and (prior to the test-optional movement) SAT and ACT scores—may therefore solidify extant impressions of academic abilities in gendered ways. In our study, we offer novel insight into how prevailing gendered cultural associations with academic fields and abilities manifest in intended major choice during the college application process.

Gender and Narrative Self-Presentation

The college application process is also a site of narrative performance where gender is likely to frame personal presentation. In addition to quantitative measures of achievement, many American colleges and universities, including the University of California, require personal statements to better assess applicants as whole persons (Bastedo et al. 2018). Students are therefore encouraged to draw upon their experiences, contexts, and identities when they present themselves in narrative form. Given the aforementioned salience of gender in academic contexts, and gender’s primary role in structuring social interactions, gender is likely to be one of the identities leveraged or performed in the crafting of personal essays (West and Zimmerman 1987). Furthermore, because personal essays resist quantification and commensuration, there is no standard by which to assess whether one’s essay will be more highly valued than another’s (Espeland and Stevens 1998). In other words, the subjective nature of personal essay writing means that applicants do not know which narratives might be deemed meritorious (Gebre-Medhin et al. 2022), creating conditions with high evaluative uncertainty. These conditions are further heightened by the fact that applicants do not know who will be evaluating their application (Aukerman and Beach 2018). Uncertainty in evaluative contexts is likely to activate gender as a primary sensemaking lens, with widely shared cultural beliefs about gender providing guardrails for individuals as they navigate uncertain contexts (Correll et al. 2020; Ridgeway 2011).

Given these characteristics of the college admissions process, we stipulate contrasting hypotheses about how gender informs women’s narrations of academic interests. To what extent do women intending to major in gender-atypical disciplines implicate their gender when writing about their academic interests, relative to women making more gender-typical decisions? We draw upon two differing theories to hypothesize potential answers to this question. First, we consider gender as a background identity, inflecting applicant self-presentation in ways that make gender identity legible (Ridgeway 2011). Second, we draw upon Kanter’s (1977) token theory, which stipulates that tokens, those who are numerically rare, minimize their minoritized identity to safeguard against social and material risks. We translate these theories to consider whether gender is signaled or suppressed by women in the production of personal narratives.

When gender operates in the background, it provides gendered inflections to the behaviors and actions associated with a given institutional role so that gender remains implicitly discernible to others, thereby facilitating interactions with expected gendered behavior. For example, a teacher has certain institutional roles to fulfill as the authority figure and knowledge bearer in a classroom, but a woman teacher may soften her presentation with more feminine expressions of warmth and niceness (Ridgeway 2011). Carli’s (2001) study of gender and influence found that when highly competent women temper their competence with displays of communalty and warmth, they mitigate the resistance they are likely to face from men peers. Furthermore, when women fail to behave in traditionally feminine ways, they can be punished. For example, women in leadership positions who showcase authority experience a backlash effect for not displaying traditionally feminine attributes, such as being nurturing and collaborative (Rudman et al. 2012). In effect, when gender operates as a background identity, it pressures women to subtly signal their gender identity so that both gender identity and institutional role are legible simultaneously.

The pressure to fulfill both sets of these expectations is heightened in contexts in which women occupy a gender-transgressive position. In our case, because “engineer” is not synonymous with “woman,” we would expect that women pursuing an engineering major would signal their gender identity in essays to counteract having selected a gender-atypical major. Differently put, women in engineering face dual pressures: (1) present as a legitimate engineer, perhaps by virtue of their academic achievement, and (2) present as a woman engineer, perhaps by virtue of how they craft their essays. Because women entering fields in which women dominate do not face the same burden to render their gender identity meaningfully, we posit that gender signaling will be stronger in academic contexts in which women are in the minority:
Hypothesis 1: Women pursuing a gender-atypical major will write about topics that implicate their gender identity, either implicitly or explicitly, to a greater degree than women pursuing a gender-typical major.

Alternatively, women who intend to major in a gender-atypical field of study may see themselves as potential tokens who are at social and material risk for crossing gendered boundaries. In her seminal work on tokenism in organizations, Kanter (1977) theorized that those who occupy a token position are perceived by numerically dominant members as highly visible, distinctively different, and stereotypical representatives of their group. Under such conditions, tokenized women are perceived and treated as less able anomalies, especially when their gender status deviates from societal norms, such as in engineering (Yoder 1994). Several studies have documented the “chilly climate” women encounter in STEM fields, including microaggressions or overt harassment (Bilimoria and Liang 2014; Hall and Sandler 1982). To protect themselves from social and material consequences, women may downplay their gender minority status, or present in a way that mirrors dominant gender expectations. For example, in Turco’s (2010) study of the leverage buyout industry, some women, tokenized by virtue of their gender, opted to learn about sports in order to gain access to important professional networks.

In the context of our study, women intending to major in gender-atypical fields may choose to write in ways that make their gender identity less salient to minimize tokenized scrutiny. Therefore, women intending to major in engineering may choose to write about topics that suppress their gender identity, or even write about the same topics as their men peers.

Hypothesis 2: Women pursuing a gender-atypical major will write about topics that implicate their gender identity, either implicitly or explicitly, to a lesser degree than women pursuing a gender-typical major.

Together, our two hypotheses represent two potential modes of gendered self-presentation in response to gender-typical and gender-atypical academic choices. Although admission readers know the gender of each application upon evaluation—the vast majority of applicants provide gender identity—we are not concerned with evaluators’ ability to identify applicant gender. Rather, we are interested in the degree to which prevailing gender norms and expectations shape author narrations of academic interest. Gender-inflected or gender-deflected narrations may be intentional or subconscious on the part of the author; whether intentional or not, these personal essays provide an uncommon opportunity to observe powerful cultural mores around gender and academic choice. In the case of Hypothesis 1, we posit that women mitigate gender-atypical choices with gender-inflected narratives that signal their identity as women alongside their intended academic identity of engineer. In the case of Hypothesis 2, we posit that women may be reacting to awareness of tokenism and discrimination of women in STEM (perhaps even personal experience) and therefore suppressing their gender identity when narrating their academic interests.

Data and Methods

To test these hypotheses we leverage a racially representative random sample of 60,000 applications to the University of California system in the 2016–2017 academic year. These data include important demographic information, such as self-reported gender identity, high school type, parental education, and reported household income, as well as two achievement metrics, high school GPA and SAT scores. The data also include students’ intended majors, which the University of California collapses into seven categories: humanities, social sciences, physical sciences, biology, engineering, business, and other. Finally, the application required students to submit four personal essays, resulting in 240,000 essays in our data set.

To refine our data set, we exclude a total of 9,202 applications. First, we omit 872 applications with missing gender data. Second, given the particular role of gender in orienting action in industrialized and affluent countries where choice is assumed (Charles 2017; Charles and Bradley 2009), we also exclude applications from students whose reported status was not “U.S. citizen” or “permanent resident” (n = 7,334). Finally, we removed applications with intended majors denoted as “other” given its undefined nature and relative infrequency in the data (n = 1,304). Applying these filters, with some students falling into more than one of these categories, yielded a final sample of 50,798 applications.

Our goal in analyzing this sample unfolds across a series of descriptive and computational analyses. First, we use logistic regressions to surface how author demography informs major intention. We complement these analyses with visualizations of predicted probabilities of choosing a given intended major as a function of SAT Math percentile, split by

---

1Although the applications are representative of the racial composition of the 2016–2017 University of California applicant pool, our data do not include students’ racial or ethnic identities. Also unavailable are evaluative data such as applicant file scores and admissions decisions.

2Applicants wrote essays in response to four prompts, choosing them among eight options. See Appendix A for a list of the essay prompts (Table A1), and see University of California (2016) for more background information.

3Applicants were given the choice of male or female when asked to provide their gender identities. Throughout, we refer to men and women not as an endorsement of the gender binary but rather to reflect the language used by the University of California system.
gender. In our second set of analyses, we employ structural topic modeling (STM) to explore the extent to which gender is implicated in the topic of personal essays written by women intending to major in biology and engineering. We elaborate on the methodology of each of these analyses below.

The Gender Demography of Major Intention

We begin with an analysis of gender patterning across six possible intended major choices. Of particular relevance for our study is the fact that the national candidate pool for selective college admissions is increasingly made up of women (DiPrete and Buchmann 2013). In addition, medium-term enrollment trends have favored STEM disciplines such as biology and engineering at the expense of the humanities in particular and, to a lesser extent, the social sciences (Brint et al. 2012; Brint 2002). A descriptive analysis of intended major counts within our sample (Figure 1) mirror these two broad patterns: (1) overall, 58 percent of the sample identify as women, and (2) more than half (54 percent) of our sample expressed intent to major in the two most common areas: biology ($n = 13,811$) and engineering ($n = 13,668$).

The distribution of men and women across these six possible intended majors exhibits strong gendered patterns. Although physical science and business each generally mirror the proportion of women in the candidate in the overall sample (59 percent and 54 percent respectively), the remaining four majors exhibit significant deviations from the sample as a whole. Biology, the social sciences, and the humanities each substantially overrepresent women relative to the sample as a whole, while women intend to major in engineering at a far lower rate than the proportion of women in the overall sample. To further our analysis of gendered self-presentation in different contexts, we focus on biology and engineering. We do so because of the popularity of these two majors and the substantially different patterns of gender representation in each: a total of 9,761 men intend to major in engineering, compared to 3,907 women (29 percent women); a total of 10,097 women chose biology as their intended major, compared to 3,714 men (73 percent women).

The empirical pairing of biology and engineering allows us

Figure 1. Count of intended majors for six academic categories by gender.
Table 1. Logistic Regression Models: Odds Ratios for Major Intent (n = 49,206).

|                          | Engineering | Biology  |
|--------------------------|-------------|----------|
| Gender (women)           | .12***      | 1.77***  |
| Income (log)             | .89***      | 1.04**   |
| Not first generation     | .96         | .99      |
| School type (public)     | 1.20***     | 1.06     |
| GPA percentile           | 1.37***     | 1.48***  |
| SAT Math percentile      | 7.94***     | .37***   |
| SAT EBRW percentile      | .34***      | 1.06     |
| Gender (women) × SAT Math percentile | 2.78*** | 1.64*** |
| Intercept                | 1.20        | .17***   |

Note: Logistic regression models with engineering (left) and biology (right) as binarized outcomes. EBRW = Evidence-Based Reading and Writing; GPA = grade point average.

Evidence-Based Reading and Writing (EBRW) score (each of which is expressed as a percentile relative to the overall sample). For this analysis, we remove applicants reporting household incomes less than $10,000 because of probable misreporting (n = 1,592).

These two models demonstrate stark differences in the social and academic profile of applicants who intend to major in engineering and biology. In particular, the odds ratios for gender indicate large and significant differences. After accounting for social class⁴ and academic performance, women still express intent to major in engineering at nearly one eighth the rate of men. Alternatively, models incorporating the same controls show that women intend to major in biology at nearly twice the rate of men.

Although both models show significant positive relationships with high school grades after accounting for all other variables, the relationship between SAT scores (percentile rank) and major intent is less uniform. High SAT Math scores are strongly predictive of intending to major in engineering, whereas SAT EBRW scores have a strong negative relationship. Intending to major in biology has no significant relationship with SAT EBRW score but a strong and significant negative relationship with SAT Math score.

Our models suggest that SAT Math score is a strong predictor of whether a student selects engineering or biology as their intended major, with the former requiring higher scores than the latter. But how might men and women differ in how they interpret their math scores as they consider entry into highly gendered disciplines? We attend to SAT Math scores as a potential signal of encouragement or discouragement to enter these fields because math is a ubiquitous gateway or prerequisite course for many STEM majors (Sanabria and Penner 2017; Thompson 2021) and because prior literature has amply documented gendered associations with math ability (Cvencek, Meltzoff, and Greenwald 2011; Grunspan et al. 2016; Musto 2019). If both men and women applicants internalize prevailing beliefs about men’s predisposition for math and science and interpret standard metrics of achievement in gendered ways, we would expect that women will require a higher SAT Math score than men to attempt entry into male-dominated fields. Conversely, if men equate women-dominated fields with lower math abilities, we would expect that men with higher SAT Math scores will opt against majoring in fields in which women are overrepresented.

To test this prediction, we interact gender and SAT Math and find the interaction effect to be both large and significant: in engineering contexts, the math signal is almost three times stronger for women than for men; in biology, the math

⁴There is some debate about whether biological sciences are to be included in the STEM field (Xie, Fang, and Shauman 2015). We include biology in our conception of STEM, following the theoretical grounds laid out by Weeden et al. (2020); biology degrees are science degrees and are likely to be perceived as such by high school students as they apply for BS and BA degrees.

⁵Table A2 provides the distributions of our independent variables. Social class measures are less discerning. In the biology model, there is no significant relationship between social class and major intention. In the engineering model, however, choosing engineering is negatively related to family income and positively related to attending a private school.
signal is 1.64 times stronger for women than for men. We then take the insights from the logistic regression models and generate predicted probabilities for whether a student intends to major in engineering (Figure 2) and biology (Figure 3) given their gender and academic achievement. We then model the linear relationships between these predicted probabilities and SAT Math percentiles and stratify by major and gender. To generate the predicted probabilities, we use logistic regression models with intended major as the dependent variable (with biology and engineering treated separately) and SAT Math, SAT EBRW, high school GPA percentiles, gender, and the interaction of gender and SAT Math as predictors. We then fit the predicted probabilities to SAT Math percentiles using locally estimated regression, a more flexible approach than traditional linear methods (Hout and Fischer 2014).

For the engineering model, men were almost always predicted to be more likely than women to select engineering as their intended major, regardless of their SAT Math score. For women below the 25th percentile of the SAT Math distribution, the predicted probabilities vary little, suggesting that other achievement metrics are unlikely to encourage women to select an intended engineering major. For men and to a lesser extent high-scoring women, there are wider ranges of predicted probabilities, which suggest that other factors inform whether they intend to major in engineering. The slopes were positive for both men and women and show how higher math scores help funnel applicants toward an engineering major intent. However, there is a slight curvature to the women’s plot, with an inflection point around the 50th percentile of SAT Math scores. This suggests that upon scoring in the upper half of the math distribution, women become
more likely to choose engineering as their intended major. However, overall and across the SAT Math distribution, women were generally much less likely than men to select engineering as their intended major regardless of their academic performance.

We observe the opposite pattern in biology: regardless of SAT Math performance, women were always more likely to select biology as their intended major than men. The predicted probabilities vary more for men and women, regardless of math score. This suggests that for many University of California applicants, the decision to choose biology as their intended major is likely mediated by other achievement metrics; our logistic regressions suggest that GPA is likely to have a stronger effect in intending to major in biology. Although the slope for women is only moderately negative, the slope for men is steeper, meaning that the relationship between SAT Math scores and the probability of intending to major in biology is stronger for men than for women; as SAT Math scores increase, the probability of men intending to major in biology decreases at a greater rate than for women. Among students at the lowest end of the SAT Math score distribution, women, on average, were about 13 percent more likely to select biology as their intended major; among students at the highest end of the SAT Math score distribution, women were about 20 percent more likely than men to choose biology as their intended major, on average. These results provide a descriptive suggestion that applicants translate SAT Math scores into major intentions in gendered ways: high SAT Math scores may actually discourage men from biology, perhaps because men interpret strong math scores as a signal to pursue fields typed as masculine, where their math acuity is better used, rather than pursue less quantitative and therefore more feminine-typed fields. This pair of results suggests that although low SAT Math scores may
dissuade women from pursuing men-dominated disciplines, high SAT Math scores may dissuade men from pursuing women-dominated disciplines.

These analyses offer an initial look into gender patterning in major intent at the largest public university system in the United States. The college application process, at least by virtue of major intent, is a site where gender seems to operate as a primary frame, with applicants reproducing rather than challenging gendered associations of academic disciplines. Academic achievement signals also appear to moderate major intention choices in gendered ways that are likely to reinforce asymmetrical gender integration patterns across academic fields. We next turn to surfacing the extent to which gender surfaces in essays written by women who make typical major intention choices (biology), and those who cross gender boundaries and intend to major in engineering.

Observing Gender Inflection in Personal Essays

In this analysis, we use computational methods to model how applicant gender predicts essay topics in engineering and biology. What do women who aspire to engineering and biology—who do and do not cross gendered academic boundaries—write about when asked to narrate their academic interests? And to what extent is their gender identity implicated in the topics they choose to narrate? We operate under two contrasting hypotheses. Considering gender as a signal, we expect that women in engineering will write about topics that distinguish them as women engineers to mitigate gender-transgressive choices. In this case, gender will be more predictive of essay topics in engineering contexts, relative to biology. Figure 4 summarizes our two hypotheses.

To test these hypotheses, we analyze 32,408 essays written in response to the following prompt: “Think about an academic subject that inspires you. Describe how you have furthered this interest inside and/or outside of the classroom.” We draw upon these essays to analyze the degree to which gender inflects essay topics.

We use STM to analyze essay content at scale (Alvero et al. 2021). Given our analyses of major intent above, we factor in the effects of gender and achievement for essays among biology- and engineering-intending majors. Topic modeling, including STM, is a branch of machine learning methods for textual data that model word co-occurrence patterns to generate insight into the thematic content of a corpus (for a broad overview of topic models, see Mohr and Bogdanov 2013). Topic modeling identifies groups of words that tend to co-occur (“topics”) in documents and creates numeric representations of essays as mixtures of topics (topic scores). Unlike other topic modeling approaches, STM allows researchers to use metadata about the text or author (in this case, gender and academic achievement markers) as part of the topic generation process. This allows us to see the content of the essays at scale (i.e., the “topics”) and to identify topics for which applicant gender has a significant effect in the topic generation process (i.e., which topics are more strongly associated with women applicants compared with men). We use STM to generate topics from two different subsets of the filtered data: all biology applicants and all engineering applicants. To determine the number of topics to generate, we follow Mimno and Lee (2014; see Roberts, Stewart, and Tingley 2019 for additional details). This approach allows for a simpler, standardized process to

---

**Hypothesis 1:**

**Signaled Gender**

Gender Process: gender inflects behavior associated with an institutional role so that gender identity is signaled alongside institutional role

Expectation: gender will be more predictive of essay topic in gender atypical majors, because women will write about different topics than men to a greater degree

**Hypothesis 2:**

**Suppressed Gender**

Gender Process: gender minorities suppress gender identity to avoid material and social risks, may behave more like dominant group

Expectation: gender will be less predictive of essay topic in gender atypical majors, because women will write about different topics than men to a lesser degree

---

7We use the term topic scores for simplicity, but scholars also frequently use the terms beta and gamma.

8This approach finds the exact convex hull of the word co-occurrence matrix using low-dimensional embeddings.
determine an optimal number of topics rather than testing different values or sets of values. Furthermore, as the essay data used for each model are of different sizes and character, using Mimno and Lee’s method helps us avoid over- or underfitting one batch of essays on the basis of the hyperparameter tuning of another set of essays.

We report our STM results in two ways: the proportion of topics for which gender had a significant effect and the sizes of the effects on the topics. The effects we describe below are, to use probabilistic notation, the differences between a word belonging to a topic and a word belonging to a topic given that it came from a woman applicant: $P(\text{word})$ versus $P(\text{word} | \text{gender})$. To see if either of our hypotheses holds, we examine whether there is a greater proportion of topics significantly associated with gender among engineering essays or biology essays. We find that the effect of gender on academic inspiration essay topics is more prevalent among prospective engineering applicants compared with their peers in biology. Among engineering essays, 37 percent of essay topics were significantly associated with gender (40 of 107 topics). Among the biology essays, 26 percent of the topics exhibited significant gender effects (27 of 102 topics). In other words, gender-differentiated topics were detected in engineering essays at a higher rate than in biology by 11 percentage points.

To ensure that these differing proportions were not due only to the fact that gender is differently distributed in each major, we estimated a structural topic model that replaced the gender covariate with a random binary variable of similar proportions as the respective distributions of women for the two focal majors (27 percent for engineering, 72 percent for biology). This novel approach could be useful for robustness checks of STM results and helps ensure that a significant relationship between the text and covariate is not spurious or because of data imbalance. In Figure 5, we display the proportions of essay topics significantly associated with author gender in biology and engineering and juxtapose those with the proportion of essays significantly associated by our random binary variable.

The structural topic model with the random binary variable covariate yielded zero essay topics in which it was a significant predictor. This was true whether the binary variable matched the gender proportions in biology or engineering. These results strongly suggest that our findings are not simply the product of gender imbalances across the two majors, but rather due to men and women applicants writing about distinctive topics to different degrees as aspiring biologists and engineers. Together, these findings support our first hypothesis that gender is signaled to a greater extent in essays written in engineering contexts, relative to biology. We next consider the degree to which gender is collectively predictive of essay topic in each discipline, and surface what applicants write about in academic inspiration essays.

Figure 6 shows the topics that are significantly associated with gender, ordered by their strength of association. The
Figure 6. Topics with significant associations with gender ($p < .05$) for biology (top) and engineering (bottom) applicants. Topics on the right side of the respective dashed vertical lines were associated with women applicants; topics on the left were associated with men applicants.
vertical dashed line is a visual aid to indicate which topics were more strongly associated with women writers (right side of the figure) and men writers (left side). For each topic, the dots represent the average increase or decrease of the respective topic score by applicant gender. As topic scores sum to unity for any given essay, they represent the proportion of an essay that contains a particular topic. Note that the figures were generated with respect to essays written by women; hence, negative values represent topics that, on average, are less prevalent for women than for men.

Although the topics of “academic inspiration” generally fall among those we might expect for STEM majors generally (e.g., science, research, scientific careers), there is variation in the degree to which each of the topics is gendered. For example, women in both intended majors were more likely to reference academic subjects outside of STEM that are more readily associated with traditional ideas of femininity (e.g., foreign languages, books and English, the creative arts). In contrast, essays written by men in both intended majors largely remain within the bounds of STEM subjects (e.g., physics laws, mathematics, and majoring in science), with history a notable outlier in both engineering and biology. Conspicuously, the essay topic most likely to be written by women was “women in science.” Specifically, the topic “women in science” was more prevalent in essays written by women intending to major in engineering by approximately 1.2 percent. Aggregating the absolute values of all topic scores in Figure 6 for each applicant pool, we find that gendered topics account for a 16.4 percent difference in the topical composition of the average engineering essay and a 10.6 percent difference in the total composition of the average biology essay. In sum, we find that gender is not only significantly predictive of a greater proportion of topics in engineering than in biology, but also that gender is predictive of essay topic to a greater degree in engineering than in biology. Across all of our analyses, we find that women signal their gender identity to a greater degree in gender-transgressive fields, evincing the persistent role gender plays in self-presentation, including during consequential sites of academic performance and selection.

Discussion and Conclusion

By taking advantage of rare undergraduate application data, we are able to (1) surface gender patterning across intended major fields and (2) provide a window into how students, situated at the fulcrum between secondary and tertiary education, inflect gender into narrations of academic ambitions. Our results collectively demonstrate the persistence of traditional gendered associations and stereotypes in informing academic commitments and self-presentation during a consequential site of academic selection, the college admissions process. More than 15 years after England and Li’s (2006) study regarding asymmetrical gender integration patterns in major choice, we find that intended major choices reflect persistent gender segregation of academic disciplines. Moreover, we find that gendered self-presentation surfaces to a higher degree among women pursuing gender-atypical disciplines, suggesting that prescriptive and proscriptive ideas about men and women’s academic choices remain highly salient in a moment of imagining future academic and professional selves.

Our descriptive analyses of applicant major intention illustrates that extant gender segregation of academic disciplines also manifests during the college admissions process. In particular, engineering and biology, the two most popular intended majors, display nearly mirrored gender imbalances, with women outnumbering men nearly 3 to 1 in biology and men accounting for more than 70 percent of all intended engineering majors. We find descriptive evidence that SAT Math scores have gendered relationships with intention to major in these two STEM subfields; the relationship between SAT Math scores and intention to major in engineering is more strongly positively associated for women than for men. This difference is also apparent in biology, though reflective of an opposite gender-academic identity proscription; SAT Math scores are more strongly negatively associated with intention to major in biology for men than for women. Our results are consistent with Correll’s (2004) assertion: individuals do not form aspirations in a vacuum, but rather draw on “perceptions of their own competence at career-relevant tasks,” which are “differentially biased by cultural beliefs about gender” (p. 111). In short, the declaration of intended major betray gendered understandings of academic abilities.

Collectively, our descriptive results portend the reinforcement of gender segregation of academic fields upon college matriculation. Initial commitments applicants make during the college admissions process may trigger course sequences and prerequisites that are difficult to disrupt, and therefore funnel students toward specific disciplines (Chambliss and Takacs 2014; Chaturapruek et al. 2021; Lang et al. 2021). This is particularly concerning in engineering, which has the lowest rate of field switching across all majors (Shaw and Barbuto 2010). Furthermore, our investigation into the relationship between SAT Math scores and major intention reveal that any modest gender integration may reinforce gendered associations between academic discipline, gender, and rigor: the small subset of women most likely to cross gendered boundaries into engineering are those with high SAT Math scores; in contrast, whereas men are nearly uniformly less like to intend to major in biology than women, the men most likely to cross gendered lines are those with lower SAT Math scores. If similar patterns hold for other disciplines, women-typed fields are likely to attract students with lower
standard metrics of achievements, while men-typed fields are likely to attract high-scoring students. This could reinforce stereotypes of feminine-typed fields as less rigorous.

Second, we hypothesized that for women entering gender-atypical fields, gender is either signaled or suppressed in the construction of essays describing academic interest. The former predicted that essay topic choice would implicate author gender to a greater degree in gender-atypical contexts, whereas the latter predicted topic choice would implicate gender to a lesser degree. We develop a novel analytical approach to a popular computational method (STM) and find that gender is predictive of more essay topics, and to a greater degree, in contexts in which women are underrepresented. Specifically, women writing in engineering contexts write about topics that distinguish them as women engineers in ways that women in biology do not. We take this as evidence of implicit or unconscious awareness of gender-transgressiveness in academic commitments by women intending to major in engineering; their academic narrations allow them to cue their gender identity, perhaps as a way to redress their gender-atypical major intention. If women are responding to even implicit pressure to soften gender-transgressive choices with narrations that signal traditional gender norms, then the cultural stereotypes around which academic disciplines men and women should and should not pursue remain salient. In other words, women who cross gendered academic boundaries seem to do so at least implicitly aware of the cultural stereotypes they are defying.

Conversely, we do not find support for token theory, which would predict gender suppression in narrative self-presentation in response to the extreme underrepresentation of women in engineering. The admissions context is likely notable. A selective college application is both an occasion for imagining a future self and demonstrating one’s unique qualifications for selection over others. As tokens in engineering, women resist being seen as different from men and therefore less acceptable as an engineer (Faulkner 2009; Yoder 1994). The high-stakes admissions context of our study, on the other hand, ostensibly rewards “standing out,” especially when applicants advance narratives of hardship, challenge, or experiences as marginalized or minoritized individuals (Gebre-Medhin et al. 2022; Silva 2013; Waller-Bey 2020). Women seeking limited spots to enter a gender-transgressive environment are thus faced with a conundrum of tokenism: how do they demonstrate worthiness as prospective engineers, irrespective of gender, while simultaneously showcasing their distinguishing experiences as women breaking gendered boundaries? Seron et al. (2018) suggested that women engineers resolve this quandary by embracing the dominant meritocratic values of the field, which allows the expression of femininity without penalty to their status as legitimate engineers. Operating under this framework, the women in our study find legitimacy meritocratically through their academic record and appear, most often rather subtly, to distinguish themselves from their men counterparts in their narrative self-presentations. If this interpretation is correct, it is additional evidence of how the meritocratic ideology of engineering contexts reinscribes dominant masculine values despite the boundary-crossing efforts of women (Seron et al. 2018).

However, we do see some evidence of women’s writing about gender stereotypes or being minoritized in STEM fields. In engineering contexts, the essay most likely to be written by women was “women in science.” Although we do not see this phenomenon at scale—“women in science” was the only topic produced by our STM that explicitly noted gender minoritization—we take this topic as evidence that at least a small but vocal minority of women engineers feel comfortable calling out their experiences as gender minorities. American colleges and universities are among many organizations that have at least publicly committed themselves to diversity and inclusion initiatives (Berrey 2015). Increasing women’s representation in STEM fields has long been an institutional priority for many American higher education institutions (National Science Foundation 2022). Awareness of these gender-egalitarian norms may encourage women pursuing STEM fields to eschew the risks of visibilizing their minoritized gender and instead foreground their gender to stand out among a male-dominated applicant pool (Holgersson and Romani 2020). Women intending to major in engineering may be responding, to a modest degree, to these egalitarian efforts. The emergence of this essay topic is perhaps a more optimistic finding among results that largely suggest the maintenance of extant academic gender segregation and cultural stasis regarding the disciplines men and women should and should not pursue.

Data constraints prevent us from taking a fully intersectional perspective (Choo and Ferree 2010; Ma and Xiao 2021), as we do not have access to race data. Further analyses could also focus more centrally on the role of class in encouraging students to pursue gender-atypical studies or deploy robust qualitative analyses to capture finer grained narrative details that may have escaped computational modeling. Because women’s underrepresentation in STEM fields has implications for broader patterns of gender inequality in society, we have chosen to focus on women’s narrative self-presentation. However, gender proscriptions also powerfully inform the decisions and behavior of men (Pascoe 2007). Future analyses may attend to the narrative self-presentation of men making gender-atypical decisions. We furthermore do not have access to decision data, so are unable to ascertain if students were admitted to their intended fields of study. As alluring as admissions decisions may be, we take these findings as consequential in their own right: the college application process is a site in which the dual performance of academic and gender identity appears to be reproduced, rather than disrupted.

Our study yields important insights for scholars and practitioners alike. We suspect that our findings about gendered performances during the college application process are
likely to be reproduced in other evaluative narrative settings, including graduate school applications, fellowship personal statements, and cover letters for jobs, especially in contexts with stark gender disparities. Competitive settings that elicit narrative presentations of self are likely to agitate gendered understandings of identity and subsequently influence gendered performances; therefore, organizations, ranging from colleges and universities to fellowship organizations and places of employment, play a key role in the solicitation and possible maintenance of gender performances and stereotypes (Mullen and Baker 2018). To the extent that these performances serve processes that reinforce gender segregation and marginalization, further attention is warranted.

Many organizations, including colleges and universities, are invested in making STEM a more welcoming space for women and may even act affirmatively on their behalf during the selection process. However, as the majority of women in our study intend to major in women-dominated disciplines, affirmative action for women intending to major in STEM disciplines reaches only a small cohort of college-going women, rendering these interventions relatively moot. However, given the gendered relationships between SAT Math scores and intended major, we suspect that widespread adoption of test-optional policies (but see the Massachusetts Institute of Technology’s recent announcement [Schmill 2022]) may remove some gendered barriers to academic disciplines, especially for women considering engineering. This prediction would be analogous to the marked increase in applications submitted to highly selective institutions following test-optional policy adoption (Rickard 2021).

Finally, much prior scholarship has encouraged women to enter STEM fields and move away from the humanities and social sciences. Although such a conclusion makes sense to achieve better gender parity and to increase women’s earnings, it also has the effect of reinforcing the idea that the humanities and social sciences are inferior to STEM, an idea rooted in gendered valuations of femininity and masculinity (Sprague 2016). We do not take our results as a wholesale endorsement of compelling women to pursue STEM (and men to pursue the humanities) but rather as a call for scholars and practitioners alike to consider how prevailing ideas about gender are iteratively communicated to, and potentially challenged by, students as they progress to and through American higher education.

## Appendix A

### Table A1. University of California Application Essay Prompts.

| Summary                              | Verbatim Prompt Text                                                                 |
|--------------------------------------|--------------------------------------------------------------------------------------|
| 1. Leadership experience             | Describe an example of your leadership experience in which you have positively influenced others, helped resolve disputes, or contributed to group efforts over time. |
| 2. Creative side                     | Every person has a creative side, and it can be expressed in many ways: problem solving, original and innovative thinking, and artistically, to name a few. Describe how you express your creative side. |
| 3. Talent or skill                   | What would you say is your greatest talent or skill? How have you developed and demonstrated that talent over time? |
| 4. Educational barrier or opportunity| Describe how you have taken advantage of a significant educational opportunity or worked to overcome an educational barrier you have faced. |
| 5. Significant challenge             | Describe the most significant challenge you have faced and the steps you have taken to overcome this challenge. How has this challenge affected your academic achievement? |
| 6. Academic inspiration              | Think about an academic subject that inspires you. Describe how you have furthered this interest inside and/or outside of the classroom. |
| 7. School and community service      | What have you done to make your school or your community a better place? |
| 8. Why I stand out                   | Beyond what has already been shared in your application, what do you believe makes you stand out as a strong candidate for admissions to the University of California? |

### Table A2. Distribution of Independent Variables.

| Variable                        | Proportion | Mean       | Median     | 25th and 75th Percentiles |
|---------------------------------|------------|------------|------------|---------------------------|
| First-generation status (not)   | 48.4% (n = 24,593) |           |            |                           |
| School type (public)            | 90.0% (n = 45,684)  |           |            |                           |
| Income                         | $105,897.20 | $69,000    | [$34,000, $140,000] |
| GPA                             | 3.75       | 3.82       | [3.47, 4.10]  |
| SAT Math                       | 600.04     | 600        | [530, 670]   |
| SAT EBRW                       | 597.96     | 600        | [540, 660]   |

Note: EBRW = Evidence-Based Reading and Writing; GPA = grade point average.
### Table A3. Verbatim Excerpts from Exemplar Essays.

| Topic               | Highest Probability | Frequent Exclusive | Excerpt                                                                                                                                                                                                 |
|---------------------|---------------------|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **Engineering**     |                     |                   |                                                                                                                                                                                                          |
| Women in science    | want, scienc, career, interest, girl, women, field | women, girl, femal, forens, stem, male, career | She nominated me to an all-girl math and science program called Tech Trek. I met girls throughout Northern California who shared mutual interests and with whom I collaborated. We discovered how to make home-made ice cream with liquid nitrogen, extracted the DNA from a strawberry and learned how genetic mutations occur. My interest in science grew exponentially. |
| Microbiology        | biolog, bodi, human, learn, cell, life, function | biolog, bodi, cell, anatomi, physiolog, genet, cellular | Although this topic was briefly discussed in Biology, it inspired me to pursue a career in the medical field. At the university, I hope to learn more about the human body. One day, I hope to be an expert and help people. I admire how incredibly well the human body works. |
| Medical practice    | medic, doctor, medicin, biomed, field, career, want | medic, doctor, hospit, medicin, biomed, surgeon, nurs | The ultimate goal of mine is to help others while practicing medicine. In addition to being a part of the Medical Pathway Program I was given the opportunity to take the course, CTE Medical Assistant Clinical-Administrative which included a five week internship at Scripps Mercy Hospital Chula Vista. Being given the opportunity to participate in this internship reaffirmed my goals to become a physician. |
| **Biology**         |                     |                   |                                                                                                                                                                                                          |
| Writing and reading | write, english, read, stori, express, essay, alway | write, essay, poem, express, wrote, poetri, stori | Ever since i was eight years old i would always keep a journal and in that journal i would always write what i did that day, my feelings and thoughts. I still write in that journal till this day it’s a place where i can express and say whatever i want. Whenever i felt upset,happy or i just needed to let my thoughts out i would just write and it made me realize that it was an amazing way for me to be able to express myself where none else would know it was private. |
| Time words          | day, one, first, time, week, go, felt | minut, week, morn, month, told, day, let | Two days prior, my calculus teacher, Mr. [Harrison], challenged my class to learn how to solve the puzzle using calculus algorithms. The stakes were high: if I wasn’t able to solve the Rubik’s Cube during a class period of 45 minutes, I would receive an F towards my semester calculus grade. If I solved the Rubik’s Cube in time, I would receive an A. It was all or nothing. When the test started, I fumbled with the Rubik’s Cube nervously. Gradually, I gained focus and confidence, forcing myself to move quickly. |
| Mental health       | psycholog, mental, peopl, learn, behavior, mind, ill | mental, psycholog, disord, psychologist, ill, psychiatrist, behavior | Mental illnesses are something I knew nothing about. All I knew was that depression causes a person to live in extreme sorrow and without a will to live. I didn’t know that there were an abundant number of mental illnesses. Since I wasn’t educated on mental illnesses I thought that they didn’t hold much importance, but now I know that they do. |

Note: Excerpts from essays written by women that were most representative of the topics most associated with women applying to engineering (top three) and biology (bottom three).

a. “Highest probability” are the words with the highest word-to-topic distribution parameters.

b. “Frequent exclusive”, often called “FREX,” is the weighted harmonic mean of a given word in terms of its overall frequency in the essay corpus and its exclusivity to a given topic (Airoldi and Bischof 2016). This metric was designed to balance the frequency of a given word and its exclusivity to a given topic, relative to other topics.
Acknowleggment

We are grateful to the University of California for sharing the relevant data and for the camaraderie of the Student Narrative Lab. We would also like to thank Shelley Correll and the members of Stanford’s Advanced Topics in the Sociology of Gender seminar for their helpful feedback. Sonia and AJ are additionally grateful for the support of the Stanford Interdisciplinary Graduate Fellowship and the Diversifying Academia, Recruiting Excellence fellowship.

ORCID iDs

Sonia Giebel https://orcid.org/0000-0002-4730-8980
AJ Alvero https://orcid.org/0000-0003-4642-9788

References

Airoldi, Edoard M., and Jonathan M. Biischof. 2016. “Improving and Evaluating Topic Models and Other Models of Text.” Journal of the American Statistical Association 111(516):1381–1403.

Alon, Sigal, and Thomas A. DiPrete. 2015. “Gender Differences in the Formation of a Field of Study Choice Set.” Sociological Science 2:50–81.

Altonji, Joseph G., Erica Blom, and Costas Meghir. 2012. “Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers.” Annual Review of Economics 4:85–223.

Alvero, A. J., Giebel, S., Gebre-Medhin, B., Antonio, A. L., Stevens, M. L., & Domingue, B. W. (2021). Essay content and style are strongly related to household income and SAT scores: Evidence from 60,000 undergraduate applications. Science Advancer, 7(42).

Armstrong, Elizabeth A., and Laura T. Hamilton. 2013. Paying for the Party: How College Maintains Inequality. Cambridge, MA: Harvard University Press.

Aukenman, Maren, and Richard Beach. 2018. “Student Conceptualizations of Task, Audience, and Self in Writing College Admissions Essays.” Journal of Adolescent & Adult Literacy 62(3):319–27.

Bastedo, Michael N., Nicholas A. Bowman, Kristen M. Glasener, and Jandi L. Kelly. 2018. “What Are We Talking about When We Talk about Holistic Review? Selective College Admissions and Its Effects on Low-SES Students.” Journal of Higher Education 89(5):782–805.

Beasley, Maya A., and Mary J. Fischer. 2012. “Why They Leave: The Impact of Stereotype Threat on the Attrition of Women and Minorities from Science, Math and Engineering Majors.” Social Psychology of Education 15(4):427–48.

Berrey, Ellen. 2015. The Enigma of Diversity. Chicago: Chicago University Press.

Beutel, Ann M., Stephanie Woodham Burge, and B. Ann Borden. 2017. “Femininity and Choice of College Major.” Gender Issues 35(2):113–36.

Bilimoria, Diana, and X. Liang. 2014. Effective Practices to Increase Women’s Participation, Advancement and Leadership in US Academic STEM. In Women in STEM careers: International perspectives on increasing workforce participation, advancement and leadership, edited by Lord, Linley, Bilimoria, Diana. London, UK: Edward Elgar.

Blackburn, Heidi. 2017. “The Status of Women in STEM in Higher Education: A Review of the Literature 2007–2017.” Science & Technology Libraries 36(3):235–73

Brint, Steven, Kristopher Proctor, Scott Patrick Murphy, and Robert A. Hanneman. 2012. “The Market Model and the Growth and Decline of Academic Fields in U.S. Four-Year Colleges and Universities, 1980—2000.” Sociological Forum 27(2):275–99.

Brint, Steven G. 2002. “The Rise of the ‘Practical Arts.’” Pp. 231–59 in The Future of the City of Intellect: The Changing American University, edited by S. G. Brint. Stanford, CA: Stanford University Press.

Carli, Linda L. 2001. “Gender and Social Influence.” Journal of Social Issues 57(4):725–741.

Cech, Erin. 2013. “The Self-Expressive Edge of Occupational Sex Segregation.” American Journal of Sociology 119(3):747–89.

Cech, Erin. 2015. Engineers and Engineeresses? Self-Concepts and the Development of Gendered Professional Identities.” Sociological Perspectives 58(1):56–77.

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy W. Williams. 2014. “Women in Academic Science: Explaining the Gap.” Psychological Science in the Public Interest 15(3):75–141.

Chambless, Daniel F., and Christopher G. Takacs. 2014. How College Works. Cambridge, MA: Harvard University Press.

Chang, Mitchell J., Oscar Cerna, June Han, and Victor Sàenz. 2008. “The Contradictory Roles of Institutional Status in Retaining Underrepresented Minorities in Biomedical and Behavioral Science Majors.” Review of Higher Education, 31(4):433–64.

Charles, Maria. 2017. “Venus, Mars, and Math: Gender, Societal Affluence, and Eighth Graders’ Aspirations for STEM.” Socius 3. Retrieved September 13, 2022. https://journals.sagepub.com/doi/10.1177/2378023117697179.

Charles, Maria, and Karen Bradley. 2009. “Indulging Our Gendered Selves? Sex Segregation by Field of Study in 44 Countries.” American Journal of Sociology 114(4):924–76.

Chaturapruek, Sorathai, Tobias Dalberg, Marissa E. Thompson, Sonia Giebel, Monique H. Harrison, Ramesh Johari, and Mitchell L. Stevens, et al. 2021. “Studying Undergraduate Course Consideration at Scale.” AERA Open 7:2332858421991148.

Choo, Hae Yeon, and Myra Marx Ferree 2010. “Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions and Institutions in the Study of Inequalities.” Sociological Theory 28(2):129–49.

Conger, Dylan, and Lisa Dickson. 2017. “Gender Imbalance in Higher Education: Insights for College Administrators and Researchers.” Research in Higher Education 58(2):214–30.

Conger, Dylan, and Mark C. Long. 2010. “Why Are Men Falling Behind? Gender Gaps in College Performance and Persistence.” Annals of the American Academy of Political and Social Science 627(1):184–214.

Correll, Shelley J. 2001. “Gender and the Career Choice Process: The Role of Biased Self-Assessments.” American Journal of Sociology 106(6):1691–1730.

Correll, Shelley J. 2004. “Constraints into Preferences: Gender, Status, and Emerging Career Aspirations.” American Sociological Review 69(1):93–113.

Correll, Shelley J., Katherine R. Weisshaar, Alison T. Wynn, and JoAnne Delfino Wehner. 2020. “Inside the Black Box of Organizational Life: The Gendered Language of Performance Assessment.” American Sociological Review 85(6):1022–50.
Crisp, Gloria, Amaury Nora, and Amanda Taggart. 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(4):924–42.

Cvencek, Dario, Andrew N. Meltzoff, and Anthony G. Greenwald. 2011. “Math–Gender Stereotypes in Elementary School Children.” Child Development 82(3):766–79.

DiPrete, Thomas A., and Claudia Buchmann. 2013. The Rise of Women: The Growing Gender Gap in Education and What It Means for American Schools. New York: Russell Sage.

England, Paula, and Su Li. (2006). “Desegregation Stalled: The Experiences That Contribute to Persistence.” Harvard Educational Review 81(2):209–41.

Estrada, Mica, Myra Burnett, Andrew G. Campbell, Patricia B. Campbell, Wilfred F. Denetclaw, Carlos G. Gutiérrez, and Sylvia Hurtado, et al. 2016. “Improving Underrepresented Minority Student Persistence in STEM.” CBE—Life Sciences Education 15(3):es5.

Faulkner, Wendy. 2009. “Doing Gender in Engineering Workplace Cultures. II. Gender In/authenticity and the In/visibility Paradox.” Engineering Studies 1(3):169–89.

Ganley, Caroline, Casey E. George, Joseph B. Cimpian, and Martha B. Makowski. 2018. “Gender Equity in College Majors: Looking beyond the STEM/nonSTEM Dichotomy for Answers Regarding Female Participation.” American Educational Research Journal 55(3):453–87.

Gebre-Medhin, Ben, Sonia Giebel, AJ Alvero, Anthony Lising Nolan, L. Wiggins, Alison J. Crowe, and Steven M. Goodreau. 2016. “Males Under- Estimate Academic Performance of Their Female Peers in Undergraduate Biology Classrooms.” PLoS ONE 11(2):e0148405.

Goldin, Claudia. 2013. “Notes on Women and the Undergraduate Economics Major.” CSWEP Newsletter (Summer):4–6, 15.

Green, Alan, and Danielle Sanderson. 2018. “The Roots of STEM Achievement: An Analysis of Persistence and Attainment in STEM Majors.” American Economist 63(1):79–93.

Grunspan, Daniel Z., Sarah L. Eddy, Sara E. Brownell, Benjamin L. Wiggins, Alison J. Crowe, and Steven M. Goodreau. 2016. “Males Under-Estimate Academic Performance of Their Female Peers in Undergraduate Biology Classrooms.” PLoS ONE 11(2):e0148405.

Hall, Roberta M., and Bernice R. Sandler. 1982. “The Classroom Climate: A Chilly One for Women?” Washington, DC: Project on the Status and Education of Women.

Harrison, Monique H., Philip A. Hernandez, and Mitchell L. Stevens. 2022. “Should I Start at Math 101? Content Repetition as an Academic Strategy in Elective Curriculums.” Sociology of Education 95(2):133–52.

Holgersson, Charlotte, and Laurence Romani. 2020. “Tokenism Revisited: When Organizational Culture Challenges Masculine Norms, the Experience of Token Is Transformed.” European Management Review 17(3):649–61.

Hout, Michael, and Claude S. Fischer. 2014. “Explaining Why More Americans Have No Religious Preference: Political Backlash and Generational Succession, 1987–2012.” Sociological Science 1:423–47.

Hurtado, Sylvia, June C. Han, Victor B. Sáenz, Lorelle L. Espinosa, Nolan L. Cabrera, and Oscar S. Cerna. 2007. “Predicting Transition and Adjustment to College: Biomedical and Behavioral Science Aspirants’ and Minority Students’ First Year of College.” Research in Higher Education 48(7):841–87.

Kanter, Rosabeth Moss. 1977. “Some Effects of Proportions on Group Life: Skewed Sex Ratios and Responses to Token Women.” American Journal of Sociology 82(5):965–90.

Kirkeboen, Lars, Edwin Leuven, and Magne Mogstad. 2016. “Field of Study, Earnings, and Self-Selection.” Quarterly Journal of Economics 131:1057–1111.

Lang, David, Alex Wang, Nathan Dalal, Andreas Paepeke, and Mitchell Stevens. 2021. “Forecasting Undergraduate Majors Using Academic Transcript Data.” Pp. 243–46 in Proceedings of the Eighth ACM Conference on Learning@ Scale. New York: Association for Computing Machinery.

Llegewie, Joscha, and Thomas A. DiPrete. 2014. “The High School Environment and the Gender Gap in Science and Engineering.” Sociology of Education 87(4):259–80.

Leslie, Sarah-Jane, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. “Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines.” Science 347(6219):262–65.

Ma, Yingyi, and Shiyang Xiao. 2021. “Math and Science Identity Change and Paths into and out of STEM: Gender and Racial Disparities.” Socius 7, Retrieved September 13, 2022. https://journals.sagepub.com/doi/full/10.1177/23780231211001978.

Master, Allison, Andrew N. Meltzoff, and Sapna Cheryan. 2021. “Gender Stereotypes about Interests Start Early and Cause Gender Disparities in Computer Science and Engineering.” Proceedings of the National Academy of Sciences 118(48):e2100030118.

Mimno, David, and Moontae Lee. 2014. “Low-Dimensional Embeddings for Interpretable Anchor-Based Topic Inference.” Pp. 1319–28 in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar: Association for Computational Linguistics.

Mohan, John W., and Petko Bogdanov. 2013. “Introduction—Topic Models: What They Are and Why They Matter.” Poetics 41(6):545–69.

Mullen, Ann L. 2014. “Gender, Social Background, and the Choice of College Major in a Liberal Arts Context.” Gender & Society 28(2):289–312.

Mullen, Ann L., and Jayne Baker. 2018. “Gender Gaps in Undergraduate Fields of Study: Do College Characteristics Matter?” Socius 4. Retrieved September 13, 2022. https://journals.sagepub.com/doi/full/10.1177/2378023118789566.

Musto, Michela. 2019. “Brilliant or Bad: The Gendered Social Construction of Exceptionalism in Early Adolescence.” American Sociological Review 84(3):369–93.

National Center for Science and Engineering Statistics. 2021. “Women, Minorities, and Persons with Disabilities in Science and Engineering.” Retrieved September 13, 2022. https://ncses.nsf.gov/pubs/nsf21321/report.

National Science Foundation. 2022. “ADVANCE at a Glance.” Retrieved September 13, 2022. https://www.nsf.gov/crssprgm/advance.
Pascoe, C. J. 2007. *Dude, You’re a Fag: Masculinity and Sexuality in High School*. Berkeley: University of California Press.

Porter, Stephen R., and Paul D. Umbach. 2006. “College Major Choice: An Analysis of Person–Environment Fit.” *Research in Higher Education* 47(4):429–49.

Posselt, Julie. 2020. *Equity in Science: Representation, Culture, and the Dynamics of Change in Graduate Education*. Stanford, CA: Stanford University Press.

Owen, Ann L. 2008. “Grades, Gender, and Encouragement: A Regression Discontinuity Analysis.” Retrieved September 13, 2022. https://ssrn.com/abstract=1295583.

Rickard, Jenny. 2021. “Dear Colleagues.” Retrieved September 13, 2022. https://mailchi.mp/f3ddd6c16b4a/application-trends-through-march-1.

Ridgeway, Cecilia L. 2011. *Framed by Gender: How Gender Inequality Persists in the Modern World*. Oxford, UK: Oxford University Press.

Riegle-Crumb, Catherine, Barbara King, and Chelsea Moore. 2016. “Do They Stay or Do They Go? The Switching Decisions of Individuals Who Enter Gender Atypical College Majors.” *Sex Roles* 74(9–10):436–49.

Riegle-Crumb, Catherine, and Menglu Peng. 2021. “Examining High School Students’ Gendered Beliefs about Math: Predictors and Implications for Choice of STEM College Majors.” *Sociology of Education* 94(3):227–48.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2019. “stm: An R package for Structural Topic Models.” *Journal of Statistical Software* 91(2):1–40.

Rudman, Laurie A., Corinne A. Moss-Racusin, Julie E. Phelan, and Sanne Nauts. 2012. “Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice against Female Leaders.” *Journal of Experimental Social Psychology* 48(1):165–79.

Sanabria, Tanya, and Andrew Penner. 2017. “Weeded Out? Gendered Responses to Failing Calculus.” *Social Science* 6(2):47.

Sassler, Sharon, Jennifer Glass, Yael Levitte, and Katherine M. Michelmore. 2017. “The Missing Women in STEM? Analyzing Gender Differentials in the Factors Associated with Transition to First Jobs.” *Social Science Research* 63:192–208.

Sax, Linda J., M. Allison Kanny, Jerry A. Jacobs, Hannah Whang, Dayna S. Weintraub, and Amber Broch. 2016. “Understanding the Changing Dynamics of the Gender Gap in Undergraduate Engineering Majors: 1971–2011.” *Research in Higher Education* 57:570–600.

Schmill, Stu. 2022. “We Are Reinstating our SAT/ACT Requirement for Future Admissions Cycles.” Retrieved September 13, 2022. https://mitadmissions.org/blogsentry/we-are-reinstating-our-sat-act-requirement-for-future-admissions-cycles/.

Seron, Carroll, Susan Silbey, Erin Cech, and Brian Rubineau. 2018. “ ‘I Am Not a Feminist, but . . .’ : Hegemony of a Meritoricist Ideology and the Limits of Critique among Women in Engineering.” *Work and Occupations* 45(2):131–67.

Shaw, Emily J., and Sandra Barbuti. 2010. “Patterns of Persistence in Intended College Major with a Focus on STEM Majors.” *NACADA Journal* 30(2):19–34.

Silva, Jennifer M. 2013. *Coming Up Short: Working Class Adulthood in an Age of Uncertainty*. Oxford, UK: Oxford University Press.

Sprague, Joey. 2016. *Feminist Methodologies for Critical Researchers: Bridging Differences*. Walnut Creek, CA: Rowman & Littlefield.

Thompson, Marissa E. 2021. “Grade Expectations: The Role of First-Year Grades in Predicting the Pursuit of STEM Majors for First- and Continuing-Generation Students.” *Journal of Higher Education* 92(6):961–85.

Torre, Margarita. 2014. “The Scarring Effect of ‘Women’s Work’: The Determinants of Women’s Attrition from Male-Dominated Occupations.” *Social Forces* 93(1):1–29.

Turco, Catherine J. 2010. “Cultural Foundations of Tokenism: Evidence from the Leveraged Buyout Industry.” *American Sociological Review* 75(6):894–913.

Tyson, Will, Reginald Lee, Kathryn M. Borman, and Mary Ann Hanson. 2007. “Science, Technology, Engineering, and Mathematics (STEM) Pathways: High School Science and Math Coursework and Postsecondary Degree Attainment.” *Journal of Education for Students Placed at Risk* 12(3):243–70.

University of California. 2016. “Insights into the Personal Insight Questions.” Retrieved September 13, 2022. https://www.ucop.edu/cap-forum/_files/documents/insights-into-the-personal-insight-questions.pdf.

University of California. 2021. “Fall Enrollment at a Glance.” Retrieved September 13, 2022. https://www.universityofcalifornia.edu/about-us/information-center/fall-enrollment-glance.

University of California Office of the President. 2021. “All-Time Record-High Number of Applicants Apply to UC, with Chicano/Latino Students Comprising Largest Proportion.” Retrieved September 13, 2022. https://www.universityofcalifornia.edu/press-room/all-time-record-high-number-applicants-apply-uc-chicanolatino-students-comprising.

Vieyra, Michelle, Joanna Gilmore, and Briana Tzimerman. 2011. “Requiring Research May Improve Retention in STEM Fields for Underrepresented Women.” *Council of Undergraduate Research Quarterly* 32(1):13–19.

Waller-Bey, Aya. 2020. “Rediscovering Voices in Uncertain Times.” *Contexts* 19(4):84–85.

Weeden, Kim A., Dafna Gelbgiser, and Stephen L. Morgan. 2020. “ ‘Pipeline Dreams: Occupational Plans and Gender Differences in STEM Major Persistence and Completion.’” *Sociology of Education* 93(4):297–314.

West, Candace, and Don H. Zimmerman. 1987. “ ‘Doing Gender.’” *Gender & Society* 1(2):125–51.

Wolniak, Gregory C. 2016. “Examining STEM Bachelor’s Degree Completion for Students with Differing Propensities at College Entry.” *Journal of College Student Retention: Research, Theory & Practice* 18(3):287–309.

Xie, Yu, Michael Fang, and Kimberlee A. Shauman. 2015. “STEM Education.” *Annual Review of Sociology* 41:331–57.

Yoder, Janice D. 1994. “ ‘Looking beyond Numbers: The Effects of Gender Status, Job Prestige, and Occupational Gender-Typing on Tokenism Processes.’” *Social Psychology Quarterly* 57(2):150–59.

**Author Biographies**

**Sonia Giebel** is a PhD candidate in the sociology of education and higher education at Stanford University. She is interested in how
different dimensions of identity, including race and gender, are understood and performed during key academic transitions, especially the undergraduate admissions process. In January 2023, she will begin a postdoctoral fellowship at the WZB Berlin Social Science Center.

**AJ Alvero** is an incoming assistant professor of sociology and data science at the University of Florida. His current research uses computational methods and frameworks to study language, strategy, and culture in society.

**Ben Gebre-Medhin** is an assistant professor of sociology at Mount Holyoke College who focuses on the relationship between universities, academic identities, and social inequality. In other work he asks how a handful of wealthy, elite, private universities in the United States came to compete over free virtual courses during the massive open online course movement.

**Anthony Lising Antonio**’s research focuses on stratification and postsecondary access, racial diversity and its impact on students and institutions, student friendship networks, and student development.