THE GENDERED SUBSTRUCTURE OF STEM: A QUANTITATIVE ANALYSIS

OF ORGANIZATIONAL CULTURE, ORGANIZING PROCESSES,

AND THE PROPORTION OF FEMALE GRADUATES

IN SIX DISCIPLINES

By

SARAH ANN MORTON

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
Department of Sociology

MAY 2019

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of SARAH ANN MORTON find it satisfactory and recommend that it be accepted.

____

Julie A. Kmec, Ph.D., Chair

____

Thomas Rotolo, Ph.D.

____

Amy S. Wharton, Ph.D.

ACKNOWLEDGMENT

There are several people I would like to thank for making this dissertation possible. First, I would like to thank my dissertation committee members Julie Kmec (chair), Tom Rotolo, and Amy Wharton for their continued support, guidance, and mentorship throughout graduate school. I thank Julie Kmec for serving as my advisor, including me on opportunities to practice scientific research, and for her endless support as I took on the challenge of completing the Sociology PhD and Master of Science in Statistics over the last five years. I also thank Tom Rotolo and my Statistics MS committee members Jan Dasgupta, Jave Pascual, and Abhishek Kaul for providing me with the statistical, methodological, and programming knowledge I needed to take on the challenging analytical techniques I use in this dissertation. I thank Stephanie Kane, Coleen McCracken, and Ali O. Ilhan for providing helpful advice on using IPEDS data (and Erik Johnson for introducing me to Ali). I thank Karen Monique Gregg, my undergraduate advisor in sociology, for her encouragement and mentorship that got me into sociology in the first place. Without her guidance, I would not have known that a graduate education was accessible to me. Lastly, I thank my family and friends, especially my mother Cheryl Stidam, father LeRoy Morton Jr., and sister Leah Morton, for supporting me throughout my education and in life.

THE GENDERED SUBSTRUCTURE OF STEM: A QUANTITATIVE ANALYSIS

OF ORGANIZATIONAL CULTURE, ORGANIZING PROCESSES,

AND THE PROPORTION OF FEMALE GRADUATES

IN SIX DISCIPLINES

Abstract

by Sarah Ann Morton, Ph.D.
Washington State University
May 2019

Chair: Julie A. Kmec

Women remain underrepresented in Science, Technology, Engineering, and Math (STEM) disciplines. A lot of research has examined individual-level impacts of their absence, such as women's lack of interest in STEM, their lack of sense of belonging, and low math confidence. In this dissertation, I use Acker's theory of gendered organizations as a theoretical foundation to conceptualize STEM disciplines as multi-layered gendered organizations. In three quantitative studies, I test hypotheses based on components of Acker's theory, specifically, organizational culture and organizing processes. I use Acker's idea of gendered subtext to justify using text data to measure organizational culture in STEM programs.

The text, program, and institutional data come from 1,758 STEM programs in six STEM disciplines: biology, chemistry, computer science, mathematics, physics, and psychology. The data sources are texts from each STEM program's "about us" or other introductory webpage, the Integrated Postsecondary Data System (IPEDS), and the IPEDS Completions Survey. In Study

1, I examine the extent to which the cultures of STEM disciplines are gendered using social network analyses of mental models. In Study 2, I explore how organizational cultures in programs differ across STEM fields using machine learning, multinomial logistic regression, and multivariate regression. In Study 3, I examine how organizational culture and organizing processes are related to the proportion of female bachelor's graduates in each discipline.

I find the largest difference in organizational culture between the most female-dominated STEM fields (biology and psychology) and the most male-dominated field in the sample (computer science). I also find that feminine organizational culture does not distinguish the cultural differences between male-dominated versus female-dominated STEM disciplines, but that the presence of masculine organizational concepts (such as expectations of brilliance and competition) distinguish the cultures in these disciplines. While I am not able to link organizational culture and the proportion of female graduates, at least in the way I hypothesize, I do find that programs housed in certain interdisciplinary departments in computer science, chemistry, and psychology have higher proportions of female graduates than programs in single-disciplinary departments. I discuss this dissertation's contributions, limitations, and implications, as well as avenues for future research.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Women remain underrepresented in several academic Science, Technology, Engineering, and Math (STEM) fields in the United States, and this underrepresentation has been a social problem of interest to scientists in many disciplines. Statistical evidence shows women's underrepresentation and inequality in academic STEM; currently, women earn about 37 percent of undergraduate STEM degrees in the United States (National Science Foundation 2014a). Disparity in STEM increases the higher one climbs up the academic ladder; while women make up about 42 percent of assistant professors and 40 percent of associate professors in science, engineering, and health fields, women only make up about 24 percent of full professors in these disciplines (National Science Foundation 2015). Women also face salary inequity in the STEM labor market outside of academia. In 2015, women earned about 82 percent the median salary of what men earned for full-time occupations in Science and Engineering overall (National Science Foundation 2017). By discipline, women earned less than men for physical scientists (25 percent less), social scientists (24 percent less), computer and information scientists (17 percent less), and earned 10 percent or less than men for biological/life scientists, psychologists, and engineers (National Science Foundation 2017). In fact, the only discipline where women earned more than men was the mathematical sciences, where women earned about 12 percent more than men in 2015 (National Science Foundation 2017).

The current literature on gender and STEM is interdisciplinary, spanning the social, life, and physical sciences, as well as engineering and education. Much of this literature focuses on gender differences in individual aspirations or interest in STEM. Studies looking at STEM aspirations have found that women are less likely to aspire to STEM careers as men (Riegle-Crumb et al. 2011; Sadler et al. 2012; Cech 2013), take STEM-related courses (Riegle-Crumb and Moore 2014), or persist in STEM majors (Sadler et al. 2012), but that factors such as extra-curricular activities can increase women's interest, confidence, and self-identification with STEM (Heaverlo et al. 2013; Stout et al. 2013).

Research also focuses on women's perceptions of and sense of belonging in STEM. Women have been found to have lower confidence in math and science than men (Ganley and Lubienski 2016; Ellis et al. 2016), as well as lower self-perceptions towards STEM subjects, sense of belonging, or professional role confidence in STEM (Hazari et al. 2013b; Cech et al. 2011; Ong 2005). Despite these gender disparities in interest, women's lack of confidence, and their lower sense of belonging, studies drawing on data from over 60 countries have found that women's actual performance in math and science is equal to or better than men's, suggesting that these disparities are not entirely due to gender differences in academic achievement (Riegle-Crumb et al. 2012; Else-Quest et al. 2010; Stoet and Geary 2018).

Although a large body of work focuses on individual-level decisions and perceptions that support women's inequality in STEM, cultural or structural factors have also been used to explain women's experiences, choices, and underrepresentation in STEM. Cultural factors such as "chilly" university climates (Hirshfield 2010; Riffle et al. 2013; Smith-Doerr et al. 2016), masculine work cultures in academia and STEM more generally (Richman et al. 2011; Gupta 2007; Uriarte et al. 2007; Burger 2009; Smith-Doerr et al. 2016), and negative stereotypes about women and STEM (Cheryan et al. 2013; Cheryan et al. 2017; Kmec 2013) have been identified as reasons for women's underrepresentation and inequality in STEM.

Recently, Cheryan et al. (2017) reviewed hundreds of quantitative studies examining gender inequality in six STEM disciplines (computer science, engineering, physics, mathematics, biology, and chemistry), citing culture, or the dynamic system of behaviors and beliefs that influence and are influenced by ideas and values, interactions, and societal structures (Fiske and Marcus 2012), as among the factors with the most potential influence on women's underrepresentation. It is important to note that the masculine culture they attribute to male-dominated disciplines differs from traditional views of masculinity (e.g., physical and sexual prowess); instead, masculine culture in STEM is a social construction that gives men a greater ability to succeed and a better sense of belonging in STEM than women (Cheryan et al. 2017). Their review identifies three components of masculine culture that attribute to women's inequality in STEM: stereotypes of STEM fields that are often not compatible with the way women view themselves (e.g., computer scientists have to be obsessed with technology), negative stereotypes and perceived bias against women in STEM (e.g., stereotypes that women are bad at math), and a lack of role models for women in STEM (Cheryan et al. 2017).

A different body of research examines how organizational demography, such as the proportion of female teachers or faculty, impacts women's representation in STEM. Overall, the findings from these studies are mixed; some studies have found that a larger proportion of female faculty or teachers is related to higher proportions of female STEM students (Sharpe and Sonnert 1999; Sonnert et al. 2007) or women's decisions to major in STEM (Stearns et al. 2016), but others have found that female students are not more likely to persist in STEM or aspire to STEM careers when more of their STEM courses are taught by female instructors (Price 2010; Griffith 2010; Hazari et al. 2013a). Similarly, research in this area has found that women in classes with higher proportions of female classmates are also more likely to choose STEM subjects (Schneeweis and Zweimuller 2012), and that a higher proportion of female STEM graduate students positively impacts the persistence of female STEM students (Griffith 2010).

To summarize, statistical evidence and research has found gender inequality in both academic and non-academic STEM, especially regarding women's underrepresentation in these fields. The body of literature examining women's current underrepresentation and inequality in STEM is very broad and interdisciplinary. The primary focus of this research is on individual level perceptions of STEM, such as aspirations, performance, interest, and confidence in STEM. A smaller body of research considers how cultural factors such as "chilly" university climates, masculine work cultures, and the sex composition of STEM faculty play a role in women's experiences and underrepresentation in STEM. Although this research has been very informative, I explore the limitations of this research, specifically the way this research aggregates STEM disciplines, in the section below.

Importance of Disaggregating STEM

While previous research on gender and STEM has provided input about the mechanisms behind women's current inequality and underrepresentation in STEM, much of this research examines STEM as a whole, rather than comparing individual STEM disciplines. As Cheryan et al. (2017) emphasize in their review of research, it is important to not only consider STEM in the general sense, but to consider possible differences of women's representation and experiences *across* disciplines because doing so provides a way of evaluating the underlying mechanisms that are most likely to impact women's current underrepresentation in STEM as a whole. It is especially important to disaggregate STEM disciplines because the underrepresentation of women is not equal across fields, as demonstrated from the much higher proportions of women in fields such as biology, chemistry, and mathematics than physics, computer science, and engineering (National Science Foundation 2016). The differences in women's representation in math-intensive fields are especially telling; though women have reported lower confidence in mathematics than men (Ganley and Lubienski 2016; Ellis et al. 2016), women still make up over 40 percent of undergraduate degree holders in the math-intensive fields of chemistry and mathematics, although they make up less than 20 percent of graduates in physics and computer science (National Science Foundation 2016). In their comprehensive review of the gender and STEM literature, Cheryan et al. (2017) cite cultural differences in different STEM disciplines as having potential influence on women's unequal representation across different STEM fields. However, no prior research to date has compared the culture of multiple STEM disciplines while also attempting to find a statistical relationship between culture and the proportion of women, leaving a large gap in the gender and STEM literature as a whole. Ultimately, disaggregating STEM fields by discipline will help find potential disciplinary cultural differences that have not been previously considered, as well as discover the potential mechanisms behind women's current STEM underrepresentation in certain disciplines compared to others.

Introduction to Dissertation Framework and Methodological Design

In this dissertation, I account for these gaps in prior research by considering the structure, culture, and sex composition of six STEM disciplines: biology, chemistry, computer science, mathematics, physics, and psychology. I explore the following research questions:

1) To what extent are the cultures of STEM disciplines gendered? (study 1)
2) To what extent do the organizational cultures of programs differ across STEM fields? (study 2)
3) How is the organizational culture of STEM programs related to the proportion of female bachelor's graduates? (study 3)
4) How are the departmental and institutional structures of STEM programs related to the proportion of female bachelor's graduates? (study 3)

This dissertation consists of three quantitative studies. I use a variety of data sources, including text data from STEM program "about us" or equivalent webpages, departmental and institutional-level data from university websites, and the Integrated Postsecondary Data System (IPEDS). Study 1 uses social network analyses of mental models in the six STEM disciplines to address research question 1, Study 2 uses machine learning, multinomial logistic regression, and multivariate logistic regression to address research question 2, and Study 3 uses logistic regression to explore research questions 3 and 4.

To fill another gap in existing literature regarding how structure and culture might be related to women's underrepresentation in STEM, I use Acker's theory of gendered organizations as the underlying theoretical framework. Specifically, I use Acker's idea of the gendered substructure to conceptualize STEM disciplines as multi-layered gendered organizations (Acker 1990; Acker 2012). I take an organizational level approach by focusing on the concepts of organizational culture and organizing processes from Acker's theory. I also use Acker's idea of gendered subtext to justify the use of text data to examine the organizational culture in STEM disciplines (Acker 2012). From my application of Acker's theory to gender and STEM, I identify four additional concepts that I use to measure organizational culture, which are two feminine and two masculine concepts: *collaboration across disciplines* (feminine), *socially-connected science* (feminine), *socially-disconnected science* (masculine), and *expectations of brilliance and competition* (masculine).

Overview of Dissertation Chapters

I structure this dissertation as follows. In chapter 2, I provide a description of Acker's theory of gendered organizations, which is the theory at the core of the three studies. I specifically apply Acker's idea of the gendered substructure and its four main concepts (gendered identities, interactions in organizations, organizational culture, and organizing processes) to the body of literature on gender and STEM to argue that STEM disciplines act as multi-layered gendered organizations. In the three studies, I specifically focus on organizational culture and organizing processes. After that, I turn to Acker's idea of gendered subtext, which is related to the gendered substructure, in order to justify the use of text data to examine the organizational culture of STEM disciplines. In this section, I also explore the literature that examines organizational culture, gender, and STEM using text data. I end chapter 2 by discussing and justifying the hypotheses for all three studies.

In Chapter 3, I explain the data collection and sampling procedure I use for the three studies. Chapters 4, 5, and 6 cover the methods and results for each study. In chapter 4, I discuss the methods and results for Study 1, which examines the extent to which the cultures of STEM disciplines are gendered using social network analyses of mental models. In chapter 5, I go over the methods and results for Study 2, which examines if there are cultural differences between STEM disciplines using machine learning, multinomial logistic regression, and multivariate regression. In chapter 6, I present the methods and results for Study 3, which uses binomial logistic regression to explore the associations between organizational culture and organizing processes with the proportion of female STEM bachelor's graduates. When I interpret each study's results, I emphasize the differences between male-dominated STEM disciplines (e.g., computer science) and STEM disciplines with at least gender parity (e.g., biology). In chapter 7, I discuss the results from the three studies, their connections to the broader gender and STEM literature, and possible explanations for each study's unusual findings. I also explain the limitations of each study, and conclude this dissertation with directions for future research and practical implications.

CHAPTER 2: APPLYING ACKER'S THEORY OF GENDERED ORGANIZATIONS TO THE GENDER AND STEM LITERATURE

This chapter provides a literature review that synthesizes two literatures: scholarship on Acker's theory of gendered organizations and empirical studies on gender and STEM. I begin this chapter by explaining the context behind Acker's theory of gendered organizations and emerging research using this theory to build the theoretical framework for this dissertation. I then explore Acker's idea of the gendered substructure to in order to conceptualize STEM disciplines as multi-layered gendered organizations. Specifically, I relate the gender and STEM literature to the four concepts of the gendered substructure: gendered identities, organizational interactions, organizational culture, and organizing processes. After that, I explain Acker's idea of gendered subtext to justify the use of text data for examining organizational culture in STEM programs, and explore the current studies examining texts (e.g., course syllabi) and organizational culture in STEM. I end this chapter by justifying and stating the hypotheses that I will test in chapters 5 and 6.

Gendered Organizations

Neoclassical organizational studies (e.g., Mayo 1933; Barnard 1968; Simon 1945) failed to consider gender, operating under the assumption that organizational structures are gender neutral, meaning that organizational structures or processes were not considered to be impacted by gender (Acker 1990). Using traditional organizational ideas, scholarship built upon this work by examining women and organizations, arguing overall that gender is a crucial part of organizational analysis and that organizational structure produces gender inequality in organizations, such as women having less opportunity and power than men (e.g., Kanter 1977; Mills 1988). While those studies took a much needed look at women's positions and experiences in organizations, these studies also assumed organizational structures were gender neutral. Gender was considered to be outside of ongoing structural processes in work organizations, rather than an integral component of organizational structure. Contrary to prior feminist and non-feminist organizational research, Acker theorized that organizations are not gender neutral; rather, assumptions about gender are present in documents used to structure organizations and in the organization of work itself (Acker 1990; Acker 2012).

Acker noted two reasons why organizational structures were previously conceptualized as gender neutral. The first reason was that it is difficult to "see" gender when only the masculine is present. In other words, men in organizations assume their behavior represents the "normal" human experience in work organizations (Acker 1990). The second reason was that when gender differences in organizations are acknowledged, it is often argued that gendered attitudes and individual behavior "contaminate" gender neutral organizations, rather than that the organizational structures themselves are gendered. This viewpoint ultimately separates organizations from the individuals in them, ignoring the interconnections between the two (Acker 1990).

Due to prior research and organizational theory's neglect of gender, as well as prior studies that explained sex differences in organizations in terms of differences in biology, socialization, and attitudes (e.g., Furstenberg 1968; Blauner 1964), Acker called for researchers examining gender inequality to consider factors that are related to the sex structuring of organizations, such as the differential recruitment of women into jobs requiring dependence (Acker and Van Houten 1974). In her theory of gendered organizations, Acker defined an organization as "gendered" if dichotomies such as advantage and disadvantage, exploitation and control, action and emotion, and meaning and identity have patterns of masculine and feminine distinctions (Acker 1990). That is, gender is not simply an addition to ongoing gender neutral organizational processes, such as gendered attitudes and behavior "contaminating" gender-neutral structures, but is an integral part of organizational processes.

Much empirical research has used Acker's theory of gendered organizations as a theoretical framework. In their review of this research, Britton and Logan (2008) outline three emerging areas: intersectionality, mechanisms for organizational change, and organizational context. First, researchers have used the theory of gendered organizations as a framework for intersectionality studies. Researchers in this area have examined topics such as gendered racism in the workplace and inequities by gender and sexuality in religious organizations. Wingfield looked at gendered racism in the workplace, finding that black men in the female-dominated field of nursing did not experience the same upward mobility as white men; in other words, the "glass escalator" effect that has been said to give men upward mobility in female-dominated fields is impacted by race (Wingfield 2009). Whitehead (2013) examined inequality by gender and sexuality in religious congregations, finding that a congregation's stance on allowing women to serve as members of the lead clergy is significantly associated with the acceptance of gays and lesbians as members or leaders in the congregation.

A second emerging area in research on gendered organizations has looked at mechanisms for organizational change. Research in this area has looked at topics such as gender inequality in leadership, promoting gender equity in academic STEM, and workplace gender desegregation. Rindfleish and Sheridan (2003) found that women in senior leadership positions identified structural barriers to more women entering leadership (e.g., the "old boys network"), but when asked to offer solutions for the "old boys network," they chose for change to come about naturally over time, rather than a structural change to this structural problem. This solution was different from their ideas for organizational change regarding the mother/executive dual role as a barrier to women in leadership, to which these senior women suggested the structural solution of increasing childcare (Rindfleish and Sheridan 2003). Latimer et al. (2014) assessed an intervention designed to promote gender equity and organizational change within STEM departments at two colleges, finding that levels of collective efficacy toward gender equality significantly increased, while levels of conflict significantly decreased after the implementation of this program. Huffman et al. (2010) looked at the dynamics behind workplace gender desegregation, finding that the presence of women in managerial positions positively impacts gender integration, and that trends toward gender integration are more due to change within workplaces instead of new integrated workplaces entering the population over time.

Lastly, studies using Acker's theory of gendered organizations have examined organizational contexts that gender workplace outcomes. Research in this area includes the topics of sexual harassment, workplace sex discrimination, and work arrangements. Chamberlain et al. (2008) explored how different organizational attributes are related to sexual harassment, finding that organizational attributes such as workplace culture, the gender composition of work groups, and worker power are related to the severity and form of sexual harassment in the workplace. On the other hand, Stainback et al. (2011) found that the experience of sex discrimination is lower for both men and women when they comprise the numerical majority of their working group. In their study of gender disparity in patenting, Whittington and Smith-Doerr (2008) found that women scientists are more likely to become patent-holding inventors in industry settings with more flexible organizational structures than in hierarchical settings in industry or academia.

In light of its influence and potential in organizational research, the role of Acker's theory of gendered organizations in empirical studies is debatable; is this theory testable, or does it act as a framework for seeing inequality? Following Britton and Logan (2008), I consider the theory of gendered organizations as a framework for seeing inequality in the context of STEM academic fields. Specifically, I use the theory of gendered organizations to conceptualize academic STEM disciplines as multi-layered gendered organizations. Unlike most prior research, I also test aspects of this theory. Specifically, I consider the gendered substructure and gendered subtext of programs in six academic STEM disciplines to test hypotheses related to Acker's ideas of organizational culture and organizing processes. I describe these concepts and their relationship to STEM disciplines below.

Gendered Substructure of Modern Organizations

Acker (2012) expanded her theory of gendered organizations by conceptualizing the gendered substructure. This gendered substructure provides the foundation for my conceptualization of academic STEM disciplines as multi-layered gendered organizations. The gendered substructure consists of hidden processes in organizations where gendered assumptions about men and women are embedded and reproduced, and gender inequities perpetuated (Acker 2012). One example of the gendered substructure is the "ideal worker" norm, or the idea of the unencumbered worker who has no obligations outside of the workplace (Acker 2006; Williams et al. 2013). The gendered substructure helps answer a persistent question: why do gender inequities continue to exist despite women's movements, equal opportunity laws, women's high representation in the paid labor force, and gender parity in college graduation rates (Acker 2012)?

According to Acker (2012), the gendered substructure consists of four components: organizing processes, organizational culture, interactions in organizations, and gendered identities. In the sections below, I summarize each of these components and apply them to current gender and STEM research to conceptualize academic STEM disciplines as gendered organizations. Figure 1 maps themes from the gender and STEM literature, which are covered in the sections below, onto the four components of the gendered substructure.

Figure 1: Proposed Theoretical Model Connecting the Gendered Substructure of Organizations to Current Research on Gender and STEM

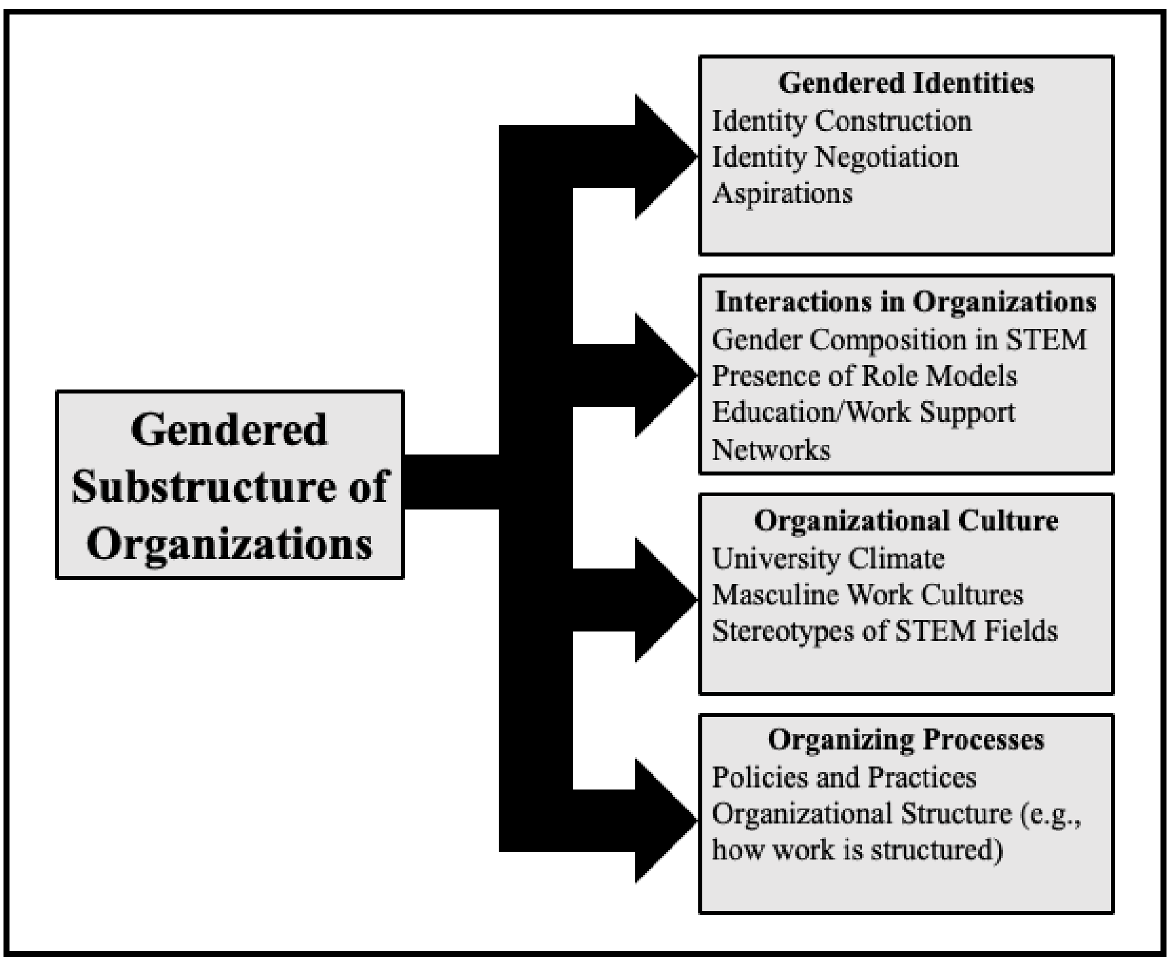


Figure 1. The concept of focus from Acker's theory of gendered organizations is the gendered substructure of modern organizations. The arrows point to each component of Acker's theory, and the themes below each component emerge from the connections between Acker's concepts

and the current literature on gender and STEM. Together, these ideas form a theoretical model that I use to conceptualize STEM disciplines as multi-layered gendered organizations.

Concept 1: Gendered Identities

According to Acker (2012), gendered identities consist of understandings of what it means to be a man or woman in a work organization, or expectations for how men and women are to behave in the workplace. These individual identities are brought into work organizations, but are constantly formed and changed as men and women participate in work processes. Acker notes that these gendered identities vary, and people modify these gendered identities as they gain work experience. According to Acker, one example of gender identities in the workplace are pressures for women in managerial positions to manage "like men," even if they may face social backlash for not conforming to a more stereotypically feminine management style (Rudman and Glick 2001).

Gendered Identities and STEM

First, STEM disciplines can be conceptualized as multi-layered gendered organizations because of gendered identities present at the individual level. These gendered identities are constantly formed and changed in STEM organizations and modified with experience (Acker 2012). In STEM disciplines, three themes arise from gendered identities: identity construction, identity negotiation, and aspirations.

Identity Construction

First, gendered identities are constructed through experiences and confidence in STEM subjects, especially mathematics. Numerous studies have shown men have higher confidence in their math skills than women, even when they obtain the same math scores (Correll 2001; Ceci et al. 2014; Bench et al. 2015). Although this lack of confidence in math ability may deter women from choosing a STEM major (Ceci et al. 2014), this finding should be taken with a grain of salt. In the most math intensive field, mathematics, women have attained gender parity in representation at the undergraduate level, and cultural factors in other math-intensive fields (e.g., physics and engineering) may be more salient than math confidence and ability in explaining women's underrepresentation (Cheryan et al. 2017). In addition, women's math confidence can be changed; Heaverlo et al. (2013) found that involvement in extracurricular STEM activities and math teacher influence positively impacted 6^{th}-12^{th} grade girls' interest and confidence in math. Regardless, women are more hesitant to enter STEM if they perceive they are unprepared, and women who express interest in STEM fields are more likely to be academically well-prepared. In other words, the discrepancy between those who are interested in STEM but who are academically under-prepared to enter STEM is greater for men than for women (Iskander et al. 2013). In addition to math confidence, Cech et al. (2011) acknowledged that math self-assessment does not impact persistence in STEM. Instead, women's lack of professional role confidence, which consists of an individual's confidence in their ability to successfully fulfill the roles and competencies of an occupation, lowers their STEM retention.

Identity Negotiation

Another way gendered identities are present in STEM organizations is through the process of identity negotiation. Identity negotiation occurs when minority groups attempt to divert attention away from their minority status and towards their identity as a full group member (Hatmaker 2013). When applied to gender and STEM, women use identity negotiation to cast attention away from their minority status as women in male-dominated STEM fields and towards their identity as STEM professionals. It has been shown that women attempt to downplay their femininity so that they will be recognized as professional engineers and scientists instead of women (Dryburgh 1999; Ong 2005; Rhoton 2011; Barnard 2012; Hatmaker 2013). Women often go to great lengths to accomplish this task, including changing their speech and the way they dress (Ong 2005; Rhoton 2011). Women of color face additional challenges, perceiving that their peers think they are less competent because they do not "fit" the typical appearance of a scientist. As a result, these women practice identity negotiation by trying to disprove negative stereotypes attributed to women and racial minorities (Ong 2005). Other STEM women negotiate their identities by distancing themselves from other women, especially those who have more stereotypically feminine behavior; some women go as far as denying gender inequality based on their own perceptions of STEM as a meritocracy (Rhoton 2011; Cech and Blair-Loy 2010).

Aspirations

Aspirations and self-conceptions are linked to gendered identities in STEM because these aspirations and self-conceptions impact interests in STEM and choices to enter STEM majors. Self-conceptions can have more influence on career choice than gender cultural beliefs; Cech (2013) found that students with emotional, unsystematic, or people-oriented self-conceptions were significantly likely to go into fields that were more stereotypically female, net of their gender cultural beliefs. These findings ultimately suggest that cultural ideals of self-expression reinforce sex segregation in STEM by framing gender-stereotypical self-conceptions and aspirations as "self-expressive" career decisions (Cech 2013). Aspirations and values also play a role in gendered identities; it is possible that cultural constructs of gender may impact men and women by altering their self-conceptions and aspirations to fit gendered roles (Cech 2015). For example, STEM women have been found to value feminine traits such as social consciousness, public welfare, and social responsibility more so than men, who typically have more masculine values such as technological leadership (Cech 2015; Cech 2014; Canney and Bielefeldt 2015). Taken together, women may develop different professional identities in STEM (Cech 2015), and the traits they value may be considered incompatible with working in STEM disciplines.

Concept 2: Organizational Interactions

Interactions between colleagues with different power levels also produce and reproduce the gendered substructure. These interactions may be formal or informal and can occur between individuals or in group settings. These interactions are often where inequality is reinforced, such as joking or harassment. Most inequalities are subtle, such as opposition to work proposals favored by women or criticisms of women's abilities to do a job that are seemingly objective (Acker 2012).

Work organizations especially reinforce gender inequality when people within organizations exclude or belittle women in their interactions. Negative interactions directed toward women in the workplace, especially in male-dominated work organizations, are related to Kanter's work on tokenism, which shows that women's underrepresentation in traditionally male-dominated spheres can be disadvantageous for women (Kanter 1977). Tokenism plays a role in workplace interactions because women's underrepresentation in work organizations results in negative experiences with workplace interactions, such as increased visibility and social isolation (Kanter 1977). It is only when women are no longer tokens, or when women's representation is over 15 percent of the specific organizational setting, when women begin to have power in organizations and be less socially isolated (Kanter 1977). However, tokenism theory has not entirely held under empirical testing; while this tipping point has been found in some empirical studies (e.g., Allmendinger and Hackman 1995; Sharpe and Sonnert 1999;

Stainback et al. 2016), other studies do not support Kanter's theory (e.g., Budig 2002; Glass et al. 2013).

Organizational Interactions and STEM

Much research has examined how the representation of female faculty and interactions with mentors and other role models impact women's experiences in STEM fields. This research has primarily focused on the topics of gender composition in STEM, the presence of role models, and educational and work support networks. Most of this research has examined how these interactions and higher representation of women in faculty positions have improved the experiences of women in STEM.

Composition of Women in STEM

First, several studies have specifically looked at how the gender composition of female faculty in STEM impacts the gender composition of female students, as well as their decisions to major in STEM. This work builds off of Tidball's (1986) study, which found a positive relationship between the number of female faculty and number of female doctoral students in the natural sciences. More recent studies have found that institutions with higher shares of female STEM faculty have larger proportions of female STEM students (Sharpe and Sonnert 1999; Sonnert et al. 2007). Higher proportions of female graduate students in a university or discipline at large also positively impact women's representation as STEM undergraduates (Griffith 2010). A positive relationship has been found between the presence of female faculty and women's decisions to major in STEM (Bettinger and Long 2005; Schneeweis and Zweimuller 2012; Stearns et al. 2016), take STEM courses (Bettinger and Long 2005; Riegle-Crumb and Moore 2014), or persist in STEM programs (Robst et al. 1998; Price et al. 2010; Griffith 2010; Marra et al. 2012), suggesting that a greater representation of female faculty in STEM increases women's sense of belonging. However, other studies have found no relationship between the proportion of female faculty and women's interest in science in secondary education settings (Gilmartin et al. 2007; Hazari et al. 2013).

Role Models

The impact of role models and mentors for STEM female faculty and students on various outcomes (e.g., recruitment and retention) has also been researched. First, the presence of role models prior to entering college has been found to encourage female students to enter STEM fields (Chanderbhan-Forde et al. 2012). Female role models have also been shown to positively impact the retention and performance of female students in STEM courses (Herrmann et al. 2016; Stout et al. 2011). Mentoring has been an effective solution for increasing the presence, retention, and advancement of female faculty and students in STEM (Chesler and Chesler 2002; Gardiner et al. 2007; Cheryan et al. 2015), and poor advising can impact student's decisions to leave STEM fields (Marra et al. 2012). Even having female advisors can create positive experiences for female STEM students, such as a heightened sense of belonging (De Welde and Laursen 2008). Although female advisors have been beneficial to women's confidence in academia, at least one study has suggested that the most successful students are those with male mentors, possibly because there are fewer female or minority faculty members in high status positions (Spalter-Roth et al. 2011).

Education/Work Support Networks and Social Capital

Finally, research has considered interactions in STEM via education and work support networks and social capital resources for female faculty, female scientists in industry, and female STEM students. Network analyses of faculty have found that female scientists are much more likely to network with other female scientists than with male scientists (Kegen 2013) and that

STEM faculty in male-dominated fields such as physics and engineering report the lowest number of female faculty in their personal support networks (Feeney and Bernal 2010). Network composition can play a role in career outcomes; larger, denser collaboration networks impact having a STEM leadership position, but having more women in one's collaboration network reduces the likelihood of having STEM discipline or research center leadership positions (Parker and Welch 2013). However, more connected networks of relationships in both courses and team-based competitions can improve the performance of STEM students (Taxler 2015; Yang et al. 2014; Dou and Brewe 2014), and women's presence on teams has performance benefits (Bear and Woolley 2011). Other network resources, such as participation in student chapters of professional STEM organizations, can help women and minority STEM students have a greater sense of belonging (Dalrymple and Evangelou 2006; Daily et al. 2007), as well as help students develop soft skills like leadership and service that are often not addressed in the classroom (Fisher et al. 2014).

Concept 3: Organizational Culture

Organizational culture, in a broad sense, consists of beliefs, attitudes, images, values, and behaviors present in organizations (Acker 2012). Culture defines "acceptable" and "unacceptable" behaviors for men and women and images of masculinity and femininity (Acker 1992; Acker 2012). It is important to note that organizational culture does not exist in isolation; organizational cultures are located in the larger cultural landscape of the surrounding society (Acker 2012). Societies themselves often have multiple perspectives on race, gender, and class relations, further shaping organizational culture. Similarly, subunits within organizations (e.g., STEM programs) may have differing cultures, some of which may perpetuate inequality (Acker 2012). This idea of a separate organizational culture of subunits is especially relevant to my conceptualization of STEM disciplines located within academic institutions as gendered organizations. Ultimately, organizational culture is important because it can support the continuation of structures and processes that produce inequality, stalling organizational change (Acker 2012).

Organizational Culture and STEM

Much research has examined how organizational culture impacts factors such as sense of belonging, perceptions of disciplines, and gender stereotypes in STEM fields at the disciplinary, institutional, and department levels. Some of this research especially focuses on how these cultures are gendered. Broader societal stereotypes influence organizational cultures present in STEM disciplines (Cheryan 2012; Cheryan et al. 2016), and both institutions and disciplines shape STEM academic culture (Lee 2007). Studies tend to explore university climate, masculine work cultures, or stereotypes of STEM fields.

University Climate

Several studies look at university climate and women's experiences in STEM, often finding the presence of a negative and isolating "chilly" university climate that excludes and marginalizes women (Hirshfield 2010; Riffle et al. 2013; Smith-Doerr et al. 2016). Workplace outcomes are impacted by department climate for both men and women faculty, and women have been found to perceive a more negative workplace climate in academia than men (Riffle et al. 2013). The climate in certain institutional contexts has also been considered in this line of research. For example, women at private institutions of higher education are more likely to persist and graduate in STEM than women at non-private academic institutions, possibly because private institutions are a more resource-rich environment with more robust communities than other institutional environments (Espinosa 2011; Zweben and Bizot 2016). Researchers have also compared university climates across different STEM fields; STEM fields which have better incorporated women into them value informal relationships more so than ones where women remain underrepresented (Cain and Leahey 2014). Although the "chilly" (or exclusive) university climate is pervasive in STEM, there have been solutions to "warming" this climate, including increasing the representation of women (Uriarte et al. 2007; Sanders et al. 2009; Maranto and Griffin 2010) and promoting a supportive, inclusive, and collaborative culture (Fox 2000; Rhoten and Pfirman 2007; Kongar et al. 2008; Kasarda et al. 2010).

Masculine Work Cultures

Researchers have also examined masculine work cultures as they relate to STEM disciplines. These masculine cultures often emphasize several traits, including individualism, competition (Gupta 2007; Uriarte et al. 2007; Burger 2009; Smith-Doerr et al. 2016), and socially-disconnected science (Uriarte et al. 2007; Schiebinger and Schraudner 2011; Smith-Doerr et al. 2016). Competition and individualism have been found to be highly discouraging to women in STEM (Uriarte et al. 2007; Smith-Doerr et al. 2016; Burger 2009), and women perceive competitive environments more negatively than men (Gupta 2007). Burger (2009) found that one of the biggest factors discouraging female students in engineering was competition in the classroom. Even opinions regarding explanations of women's underrepresentation in STEM are individualistic; many Americans attribute women's underrepresentation in STEM to their lack of human capital instead of structural explanations (Cech and Blair-Loy 2010), and men often see women's underrepresentation as an individual issue stemming from gender stereotypes (e.g., women having natural nurturing capacities that turned them away from engineering; Gill et al. 2008). Socially disconnected science, or a scientific paradigm that disconnects scientific research from practical or societal issues and is very limited in scope (particularly to theoretical work in a single discipline), has also been disadvantageous for women in STEM (Uriarte et al. 2007). In several studies, women have expressed interest in doing research that impacts society (Smith-Doerr et al. 2016; Kongar et al. 2008; Espinosa 2011), and women leave STEM fields in part because of the inability of professors to make STEM education accessible or align with their goals to contribute to society (Espinosa 2011). Women in STEM engage in social conscious research that crosses disciplines and serves multiple stakeholders and missions outside of academia (Rhoten and Pfirman 2007), and women tend to value collaborative cultures and working in teams while conducting their research (Kongar et al. 2008; Fox 2000; Rhoten and Pfirman 2007; Smith-Doerr 2005).

As stated earlier, organizational culture does not exist in isolation from the rest of society, but is instead shaped by culture that exists in the surrounding society. Likewise, organizational culture in STEM is not immune to this outside societal influence. This influence primarily consists of stereotypes about gender and STEM fields. Much research in STEM has focused on this topic, and current stereotypes of STEM fields relevant to the current study include ideas about people who work in STEM and the perceived status of STEM fields (Cheryan et al. 2017). People who work in STEM fields are stereotyped as being masculine (Cheryan et al. 2013). Stereotypes about women as nurturing and family oriented are often viewed as incompatible with STEM culture (Haines and Wallace 2003). These stereotypes vary across STEM disciplines and have produced different outcomes. For example, STEM fields with higher proportions of women tend to be low in status, and an influx of women into these fields lowers the status (Cain and Leahey 2014; Kessel 2014). The devaluation of STEM fields with higher proportions of women is further evidenced by women's smaller share of representation in STEM fields with higher salaries (Kessel 2014).

Concept 4: Organizing Processes

While organizing processes may appear gender neutral, gender inequality can be embedded into organizational structure as a result of structural factors such as job design, the distribution of decision-making, and supervisory power (Acker 2012). For example, Acker notes how job classifications and groupings of jobs are used to justify paying women lower wages than men (Acker 2012).

Organizing Processes and STEM

STEM disciplines can also be conceptualized as multi-layered gendered organizations because of organizing processes that perpetuate the gendered substructure. In STEM disciplines, organizing processes underlie the gendered substructure in two ways: through policy and practices and organizational structure.

Policy and Practices

Several policies and practices in STEM organizations are a part of the gendered substructure, and these processes can both perpetuate and reduce gender inequality in STEM. In academia, these processes play out in several ways. Practices such as NSF ADVANCE initiatives and improving the macro and micro climates of academia can improve women's retention in STEM (Bilimoria et al. 2008; Stepan-Norris and Kerrissey 2016). However, a lack of work-life balance policies and supportive institutional and departmental environments can negatively impact women faculty's retention (Gardner 2012). Next, job advertisements in venues targeting women have been found to increase the proportion of women applicants for academic STEM positions, and the greater the proportion of female applicants, the greater the chance that a woman will be hired (Glass and Minnotte 2010). At the promotion stage, the informal practices and lack of clear standards may prove difficult for women seeking promotion to full professor (Britton 2010). For students and faculty, external alliances with STEM faculty through externally funded projects and course offerings can positively shape environments for women in STEM (Fox et al. 2011). Ultimately, the idea of women's retention in STEM as an individualized "pipeline" problem has created a misalignment between structural programs and the individualistic activities that take place in STEM; alternatively, scholars have suggested that STEM departments look beyond individually oriented practices (e.g., peer mentoring and social events) and towards more structural solutions (e.g., initiatives to recruit women in STEM and programs providing women connections to other science and engineering faculty) for retaining women in STEM (Fox et al. 2011).

Organizational Structure

Several studies in both academic and non-academic settings have also examined how different organizational structures perpetuate and reduce gender inequality in STEM. The relationship between flexibility or hierarchy in STEM work settings and the outcomes of STEM women is up for debate. On one side, less hierarchical and bureaucratic structures have benefitted women in industry and academia-industry partnerships, making them more likely to achieve positive outcomes such as promotions and patents (Smith-Doerr 2004; Whittington and Smith-Doerr 2008). Both within and outside of STEM, universities divide faculty into tasks they perform and by discipline, which often means female and male academics are segregated into different departments and have varying representations at different levels. As a result, female faculty members perform a disproportionate share of teaching, care work, and emotional labor (Bird 2011). Similarly, bureaucratic structures have disadvantaged women in STEM because of their proclivity towards interdisciplinary research (Rhoten and Pfirman 2007; van Rijnsoever and Hessels 2011); as long as disciplines are divided, interdisciplinary research will not be as highly valued as traditional research paradigms (Rhoten and Pfirman 2007). In the classroom, hierarchy may be discouraging to women because of excessive and inflexible standards in the classroom, as well as classes structured as highly competitive "weed out" courses (Parson 2016; Bejerano and Bartosh 2015; Mervis 2011). However, more flexible organizational structures have also been shown to perpetuate inequality for women in STEM industry; Roth and Sonnert (2011) found that women in flexible STEM organizations perceived gender bias in workplace decisions. One reason behind these discrepancies in findings could be that in organizations with exclusionist cultural frames of masculinity, as well as organizations that rely on informal networks with powerful individuals to which women have less access (Roth and Sonnert 2011), nonhierarchical and flexible structures may not be beneficial to women (Ridgeway 2009).

Gendered Subtext

A concept related to the gendered substructure of gendered organizations is the gendered subtext (Acker 2012). The gendered subtext is distinct from the gendered substructure because it refers to written or common practice texts (e.g., policies, memos, handbooks, and guides) that shape gendered processes and structures in organizations (Acker 1990; Acker 2012). Put differently, the gendered subtext shapes organizational functions and can potentially contribute to the reproduction of gendered organizations. Acker gives the example of an Oregon state government job evaluation text where secretary jobs could not be compared with management jobs, even if secretaries performed management tasks (which often happens). Because of this job evaluation system, the pay gap between managers and secretaries could not be reduced, despite the extra management work secretaries completed. This policy appears to be gender neutral, but it is an example of gendered subtext because the women in this organization were secretaries and the men were the managers. Texts such as policies, guides, and memos may contribute to the reproduction of gendered workplaces (Acker 1990; Sargent 2009; Acker 2012), and this ultimately produces and reinforces gender inequities in these settings. The gendered subtext of organizations consists of two elements: organizational logic and ideal worker norms (Acker 2012). These ideas are explored in the following sections.

Organizational Logic and Ideal Worker Norms

Organizational logic refers to how organizations are put together through their texts, such as formal policy, memos, books, and mission statements. Typically, organizational logic consists of bureaucracy and hierarchy (Acker 2012), but other organizational logics exist, such as team-based versus hierarchical control. These different types of organizational logic are created and transmitted through texts, such as bureaucratic policies that differentiate management positions in ways that tend to preserve sex segregation. The gendered logic of organizations also includes expectations for work behavior. While these expectations appear to be gender neutral, they are based on an abstract worker that is not a universal worker, but an ideal worker (Acker 1990; Acker 2012; Williams et al. 2013). This ideal worker is unencumbered, having no obligations outside of the workplace (Williams et al. 2013). While the concept of the ideal worker appears to be gender neutral, it is part of the gendered subtext, differentiating women from men. Men are more likely to be seen as legitimate, ideal workers than women because women have traditionally performed the unpaid labor that has made it possible for men to play the role of the ideal worker (Acker 2012). This structural separation perpetuates images of femininity and masculinity, and these images impact organizational processes that perpetuate gender inequality and occupational sex segregation (e.g., Skuratowicz and Hunter 2004), especially in male-dominated disciplines where ideal worker norms are so prevalent (Acker 2012).

Gendered Subtext and STEM

Studies of gendered subtext and STEM have focused on three key areas: organizational culture in STEM texts (e.g., course syllabi, mission statements, and admissions materials), the organizational logic of STEM programs, and ideal worker norms for STEM students and employees. These studies of the gendered subtext in STEM are important because it has been found in a study of engineering that non-experts in engineering identify engineering students as dominant, forceful and masculine solely from institutional mission statements, showing the masculine culture of engineering (De Pillis and De Pillis 2008). This masculine culture is one of the problems in attracting a diverse group of students into STEM (Cheryan et al. 2017).

Organizational Culture in STEM Texts

First, several studies have analyzed organizational culture through STEM texts, such as departmental mission statements (De Pillis and De Pillis 2008), course syllabi (Parson 2016; Bejerano and Bartosh 2015), websites (Moreau et al. 2010; Moreau and Mendick 2012), admissions materials (Osei-Kofi and Torres 2015), and university internship etiquette manuals (Heflin 2015). Although these texts appear to be gender neutral (Bejerano and Bartosh 2015), their language has been found to depict both overt and subtle gendered subtext. Scholars have pointed out several instances of overt masculine culture in course syllabi, including language with an authoritarian and individualistic tone (De Pillis and De Pillis 2008; Parson 2016; Bejerano and Bartosh 2015), language that promotes a competitive and isolating "chilly" university climate (Parson 2016), and an overemphasis of "masculine" thinking with a focus on rational thought and knowledge as static and unchanging (Berjerano and Bartosh 2015; Parson 2016). These overt masculine cultures in STEM texts contribute to the gendered subtext in STEM that marginalizes women.

Other studies of gendered subtext in STEM have observed more subtle forms of gender and culture in STEM texts. First, depictions of men and women scientists have produced and reinforced gendered images. For example, websites and course syllabi often depict women in general or in STEM as being in supportive roles relative to men (e.g., photos of women providing assistance to men; see Osei-Kofi and Torres 2015; Bejerano and Bartosh 2015), featured as less established scientists (Moreau and Mendick 2012), and portrayed in domains of science that are culturally constructed as feminine (Moreau and Mendick 2012). Websites emphasized feminine traits of women scientists, such as their empathy towards living beings, communication skills, and ability to form relationships (Moreau and Mendick 2012). Texts from websites and university advertisements also emphasize women's relationships with men (especially more senior men scientists), details about their private lives, and appearance more than men's (Moreau and Mendick 2012; Osei-Kofi and Torres 2015; Heflin 2015). For example, a study of university internship manuals found that descriptions of appropriate dress codes for women tended to be longer and more detailed than men's (Heflin 2015). Even advertisements and stories intended to promote diversity promote gender stereotypes; in a study of college admissions material, women appeared in pictures in labs and other scientific spaces, but they were often not pictured as engaged and active in the scientific work (Osei-Kofi and Torres 2015). In articles about scientists, women were generally excluded from general articles about scientists and instead were relegated to a "Women in Science" webpage (Moreau and Mendick 2012). Ultimately, gendered subtexts in STEM support the gendered substructure while making STEM appear to have gender equality.

Organizational Logic and STEM

Various forms of organizational logics exist as part of the gendered subtext in STEM fields. Hierarchical structures and cultures have been prominent in STEM texts, and this hierarchy has been found to be discouraging to women (Gill et al. 2008). Hierarchy is especially present in the often inflexible standards of the classroom (Parson 2016; Bejerano and Bartosh 2015). Parts of this hierarchical classroom structure include strict and excessive rules and procedures, many prerequisites to take courses (especially in math-intensive fields), and the mentality of difficult "weed out" courses expressed in texts such as course syllabi (Parson 2016; Bejerano and Bartosh 2015; Mervis 2011). These texts ultimately show how masculine culture is present in STEM.

Ideal Worker Norms in Academic STEM

Finally, ideal worker norms are part of the organizational logic in STEM disciplines. Like other areas of the labor market, this ideal worker is conceptualized as a man without family responsibilities or other obligations that may interrupt time dedicated to work (Williams et al. 2013; Gappa et al. 2007; Hill et al. 2014). Ideal worker norms and expectations put a strain on women, especially in academia, a traditionally male-dominated discipline. Women have expressed that they often have competing expectations in academia (e.g., research and teaching), especially in "strident" institutions where prestige and status are the mission; however, women have also expressed that they do not have the institutional support to live up to these expectations (Gardner 2013). Having a family further complicates working in competitive institutions; women feel conflict between work and family and are even perceived negatively when they take parental leave (Gardner 2013). Ideal worker norms present in academia are internalized by STEM mothers; it has been found that STEM mothers perceive that they have to work harder in academia than STEM and non-STEM fathers and mothers who are not in STEM fields (Kmec

2013). In addition to ideal worker norms, expectations of "brilliance" infiltrate STEM disciplines, especially those which are more male-dominated. One study found that fields with the highest expectations of brilliance had the lowest proportions of women, most likely due to stereotypes that women do not have the innate ability to be brilliant in these fields (Leslie et al. 2015).

Alternative Explanations and Justification of Using Text Data

Alternatively to the research above, it is possible that it is difficult to detect variation in the organizational culture underlying STEM program texts across disciplines and academic institutions. This is because of symbolic isomorphism, or the similarity in symbolic attributes of organizations within the same institutional field (Glynn and Abzug 2002). In other words, it is possible that STEM programs create their texts based on what others in their field have done in the past, eliminating a large portion of potential variation across texts. Even course syllabi may be based on how previous professors taught the course, which may not accurately reflect the current culture of the department. Although the potential lack of variation in STEM program texts is a limitation to this series of studies, I address this by carefully measuring feminine and masculine culture as justified by the diverse STEM education and gender literature.

Nevertheless, the gendered substructure and subtext of organizations are important to this analysis for several reasons. The gendered subtext provides justification for using text data to examine the gendered substructure of STEM disciplines, specifically organizational culture. Since I have justified why STEM disciplines can be conceptualized as multi-layered gendered organizations, it follows that their gendered subtext is one way to empirically examine organizational culture. The gendered substructure also justifies looking at how organizing processes (e.g., interdisciplinary departments versus single disciplines in a department) are associated with women's representation in STEM disciplines. The ability to examine these two concepts (organizational culture and organizing processes) in a series of quantitative studies ultimately allows me to test hypotheses based on these concepts from Acker's theory of gendered organizations. The specific ways I form hypotheses to test Acker's theory are in the section below.

Summary of Studies and Hypotheses

Table 1 summarizes the research questions, potential data sources, and hypotheses. Although the literature on gender and STEM above supports the conceptualization of STEM disciplines as multi-layered gendered organizations, we know little about how culture is gendered in STEM disciplines at the macro level (e.g., in academic STEM programs). Similarly, we know little about cultural differences in programs across STEM fields (Cheryan et al. 2017), as well as if the organizational culture and organizing processes of STEM programs are associated with the proportion of female graduates in STEM. This dissertation provides several advances to sociology and other disciplines that examine gender and STEM. First, by testing if organizational culture or organizing processes are related to the proportion of female STEM graduates, we learn if these factors are a potential mechanism behind the unequal representation of women across different STEM fields. Beyond STEM disciplines, this dissertation's framework could help untangle the mechanisms behind the unequal progress of occupational sex desegregation in general, or why women have become more represented in some occupations but not others. Second, these analyses test if concepts from Acker's theory of gendered organizations (organizational culture and organizing processes) actually play a role in women's outcomes in STEM, specifically the proportion of female graduates. Even though this dissertation is limited to the realm of academic STEM, it provides an important case study for examining if organizational culture or organizing processes in gendered organizations are actually related to women's representation in an organization. Third, the quantitative methods I use in this study provide alternative ways for operationalizing organizational culture and organizing processes to the typical qualitative methods in this line of research, which future studies can use to test aspects from Acker's theory of gendered organizations in other contexts. Lastly, this dissertation helps us learn if different STEM disciplines actually have different organizational cultures. Learning if the organizational cultures of different STEM disciplines differ, especially between male and female-dominated STEM disciplines, provides insight into factors that may be behind women's general inequality in STEM.

My research aims to account for the literature gaps I mentioned above in the following ways. In Study 1, I do a social network analysis of program descriptions and other "about us" website pages from six academic STEM disciplines to examine the extent to which STEM programs are gendered. This study helps us learn the nuances of organizational culture in the six STEM disciplines. For example, are words about individualism and competition the most central in male-dominated STEM disciplines? Likewise, are words about collaboration or socially-connected science most central in female-dominated STEM disciplines? This analysis does not test any hypotheses, but it provides a rich description of the way words and concepts are situated in each discipline's program texts.

In Study 2, I explore cultural differences across STEM fields by exploring these same texts (STEM program webpages from "about us" or other introductory webpages). I operationalize organizational culture using categorical measure derived from machine learning algorithms, as well as continuous measures. Acker's theory of gendered organizations provides a framework for seeing past the supposed gender neutrality in organizations. I hypothesize the following to test Acker's theory and previous research that has suggested there are different cultures in different STEM disciplines (see Cheryan et al. 2017):

Hypothesis 1: Disciplines with higher proportions of female graduates (e.g., biology) will have more feminine cultures than disciplines with lower proportions of female graduates (e.g., computer science). Likewise, disciplines with lower proportions of female graduates (e.g., computer science) will have more masculine cultures than disciplines with higher proportions of female graduates.

In Study 3, I use the feminine and masculine organizational culture variables I construct in Study 2 to explore how the organizational culture and organizing processes of STEM programs might be related to the proportion of female graduates. We currently know a lot about how individual level factors such as perceptions of STEM disciplines and sense of belonging impact women's experiences in STEM fields (Cech et al. 2011; Cech 2013; Ong 2005; Marra et al. 2012; Stout et al. 2013; Danielak et al. 2014). The review above covered literature about perceptions of STEM culture, organizational interactions and STEM, and gendered subtext that exists in STEM disciplines. This literature leads to the following hypothesis:

Hypothesis 2.1: In each of the six disciplines, STEM programs with more feminine cultures will have higher proportions of female graduates than STEM programs with more masculine cultures.

Besides organizational culture and gendered subtext, this review covered literature showing how organizing processes might play a role in women's inequality in STEM. While no studies have examined how these different organizing processes might be associated with the proportion of female graduates in STEM, this literature finds that women are more inclined to do interdisciplinary research than men and enjoy working in interdisciplinary teams (e.g., Rhoten and Pfirman 2007; van Rijnsoever and Hessels 2011; Smith-Doerr 2005). This research produces the last hypothesis:

Hypothesis 2.2: In each of the six disciplines, STEM programs housed in interdisciplinary departments (e.g., department of biology and chemistry versus a single-disciplinary department of biology) or colleges (e.g., College of Arts and Sciences verses a College of Science) will be associated with higher proportions of women than STEM programs in single-disciplinary departments or colleges.

These analyses will advance our knowledge of potential disciplinary differences in STEM program culture across disciplines, the extent to which STEM disciplines will have a masculine or feminine culture at the program level, and if organizational culture and organizing processes are related to the proportion of female STEM graduates.

Table 1: Study Summaries

Research Question	Data Source(s)	Hypotheses	Support
To what extent are the cultures of STEM disciplines gendered?	STEM program descriptions from websites	N/A	Cheryan et al. 2017
To what extent do the organizational cultures of programs differ across STEM fields?	STEM program descriptions from websites	1	Cheryan et al. 2017
How is organizational culture related to the proportion of female bachelor's graduates in STEM?	Classifications from Study 2 used as measures of organizational culture Institutional, departmental, and program data from IPEDS	2.1	Fox 2000; Rhoten and Pfirman 2007; Kongar et al. 2008; Kasarda et al. 2010
How are the departmental and institutional structures of STEM programs related to the proportion of female bachelor's graduates?	College and department level information from university webpages Institutional, departmental, and program data from IPEDS	2.2	Rhoten and Pfirman 2007; van Rijnsoever and Hessels 2011

CHAPTER 3: DATA COLLECTION AND SAMPLING PROCEDURE

I explain the data collection and sampling procedures in this chapter. I begin this chapter by explaining the data collection process. I then explain why I choose to examine the six STEM disciplines in this dissertation (biology, chemistry, computer science, mathematics, physics, and psychology) and the stratified sampling procedure I use to draw the analytic sample.

Data Collection

The data for this project come from several sources. The first data source is the Integrated Postsecondary Education Data System (IPEDS), which provides data on higher education institutions (IPEDS 2015). To calculate the proportion of women in each discipline's program, I used the IPEDS data available through the U.S. National Science Foundation via the interactive WebCASPAR interface. Specifically, these data come from the 2015 IPEDS Completions survey, which was the most recent data available at the time of analysis (National Science Foundation 2016). I collected each program's data on general majors (e.g., biology) using Classification of Instructional Programs (CIP) codes, which are codes developed by the National Center for Education Statistics to classify different academic majors. I merged the institutional data from IPEDS and the data from the IPEDS completions survey together using the Unit ID.

The second source of data are texts from each program's description, which were obtained from each program's "About Us" webpages, program descriptions, or other introductory webpages. I obtained these texts using a text miner[1] I built via Python 2 (see the appendix for the text miner's code). While I initially considered using other texts found in

[1] For websites that had broken links or did not work with the text miner, I had to manually gather these texts.

previous qualitative studies of STEM texts (e.g., mission statements and course syllabi), I narrowed my focus to program description texts for several reasons. First, while not entirely uniform, most programs had some kind of webpage with a program description, which makes for more accurate comparisons of texts and a larger sample size. Second, it was difficult to obtain course syllabi and mission statements for many of the programs, which would have resulted in a much smaller sample. Lastly, course syllabi would have been difficult to accurately compare since course selection and classification vary by university.

Selection of STEM Disciplines

I specifically chose to compare computer science, biology, mathematics, chemistry, physics, and psychology for several reasons. First, these disciplines are very common majors at universities in the United States, ensuring samples large enough to perform statistical analysis. Second, these disciplines have varying levels of female graduates. Table 2 shows the distribution of female graduates earning bachelor's degrees in each field for the year 2015 (National Science Foundation 2016). These disciplines have also been compared in previous studies and critical reviews examining gender and STEM (e.g., Cheryan et al. 2017; Williams and Ceci 2015; Baram-Tsabari and Yarden 2008), allowing my three studies to build upon previous comparisons. Lastly, these disciplines have varying levels of math intensity, laboratory components, and subject matter, making them interesting to compare.

Table 2: Counts and Percentages of Female Bachelor's Degree Graduates in STEM (2015)

	Total Women Graduates	Total Graduates	Percentage of Women Graduates
Computer Science	2,298	14,510	15.8%
Biology	42,095	69,761	60.3%
Mathematics	7,047	16,393	43.0%
Chemistry	6,203	13,139	47.2%
Physics	1,093	5,488	19.9%
Psychology	79,827	103,909	76.8%

Source: National Science Foundation (2016)

Sampling

The unit of analysis is the program level for six STEM disciplines (computer science, biology, mathematics, chemistry, physics, and psychology). The program level is distinguished from the department level because programs just include pure majors (e.g., mathematics), while departments could include multiple majors in similar fields (e.g., a department of mathematics that contains programs for mathematics, statistics, and mathematics education). I drew a stratified random sample of 1,800 programs (300 per discipline) from the population of universities participating in the IPEDS Completions Survey (National Science Foundation 2016) to ensure adequate statistical power. I selected an additional stratified random sample of 600 programs (100 per discipline) to use as training and test data for the text classifiers. Removing missing data via listwise deletion yielded 1,758 programs for analysis and 580 programs for training and test data. Table 3 gives the final counts for analysis, training, and testing data.

Table 3: Sample Sizes for Analysis and Training, Testing, and Validation Data by Discipline

	Analysis Data	Training and Testing Data
Computer Science	290	98
Biology	297	99
Mathematics	292	92
Chemistry	295	97
Physics	290	95
Psychology	294	99
Total	**1758**	**580**

I collected a sample of academic programs for several reasons. First, it is relatively easy to collect data from this level because of the ability to use CIP codes to separate programs by their general discipline (see the “Data Collection” section above for an explanation of CIP codes). Second, I specifically collected data from general programs in each of the six STEM disciplines, rather than including other sub-disciplinary programs, to rule out potential sub-disciplinary influences and to ensure that each program met the National Science Foundation’s (NSF’s) definition of STEM[2]. Lastly, because programs are housed within departments that often have much autonomy in the university setting, there is potential for organizational cultural differences across programs in the six disciplines (Lee 2007).

[2] The NSF has a broad definition of STEM that includes the social and behavioral sciences, but does not include certain professional and practical disciplines (e.g., medical practice and counseling psychology). Using general program data ensures that majors such as counseling psychology or pre-medicine will not be included in the demographic data.

CHAPTER 4: THE EXTENT TO WHICH THE CULTURES OF STEM DISCIPLINES ARE GENDERED (STUDY 1)

Study 1 examines the following research question: to what extent are the cultures of STEM disciplines gendered? This research question is important because it provides descriptive evidence supporting STEM disciplines as gendered organizations. I explore the first research question by using social network analyses of STEM program texts to examine the relationships between the organizational cultural concepts and to examine which organizational cultural concepts are the most central in the six STEM disciplines. Specifically, I explore the mental models of the six STEM disciplines (biology, chemistry, computer science, mathematics, physics, and psychology) through analyses of centrality scores for one and two-mode networks of organizational cultural concepts in order to examine the extent to which these STEM disciplines are gendered in terms of their organizational culture.

I begin this section by explaining the use of social network analysis in Study 1, specifically, the concept of mental models (Carley and Palmquist 1992). I then explain the structure of the networks used in this analysis that form the mental models. The networks are a one-mode network of words for each organizational cultural concept and their relationships based on their ties to STEM programs, and a two-mode network structure consisting of each STEM program and organizational cultural concept. I also explain the centrality scores I use to quantify which aspects of organizational culture are the most central in the mental models for the six STEM disciplines. I conclude this section with a discussion of the results.

Justification for Social Network Analyses of Texts

The social network analyses of texts in this study are based on work analyzing mental models at the individual level. A *mental model* is an individual's internal representation of reality that consists of a network of associations between concepts (Carley and Palmquist 1992; Diesner et al. 2005). *Concepts* are simply units containing a single idea, which range from single words to more complex phrases (Carley 1997). Mental models make three assumptions (Carley and Palmquist 1992). First, mental models assume that the cognitive structure and texts can be modeled using symbols, which are the concepts themselves. Second, mental models assume that a text is a sample of what is known by the individual. Lastly, mental models assume that the symbolic or verbal structure that is extracted from texts can be represented as networks. Beyond the individual level, *team* or *shared mental models* consist of group members' aggregate understanding and knowledge or beliefs about key elements of their environment (Mohammed et al. 2000).

I examine the shared mental models[3] of organizational culture across six STEM disciplines, where organizational culture consists of beliefs, attitudes, images, values, and behaviors present in organizations (Acker 2012). Specifically, the mental models for each STEM discipline are a network of organizational cultural concepts and the relationships between these concepts (Carley and Palmquist 1992). I treat individual STEM programs as the individuals that form the aggregate shared mental models at the group, or disciplinary level.

Broadly speaking, I complete four steps to analyze the mental models of each STEM discipline's organizational culture as defined by Carley and Palquist (1992). First, I identify concepts of "masculinity" and "femininity" using a confirmatory approach[4]. I code words as belonging to one of these two concepts based on existing research on gender, culture, and STEM

[3] I refer to the shared mental models of each STEM discipline as mental models throughout the dissertation.
[4] While there are words deemed as "masculine" or "feminine," I do not label any words as "gender neutral." Gender neutral is simply the absence of feminine or masculine concepts.

(see table 4 below for the specific words and concepts). Second, I define relationships, or the ties between concepts and words, as both words being present in a STEM program's text (in the case of the one-mode network) and a concept being present in a program's text (in the case of the two-mode network). Third, I extract statements from the texts ("about us" or equivalent introductory webpages), which are defined as a set of two concepts and the relationship between them. Lastly, I display and analyze the networks of the mental models using social network analysis techniques, specifically, measures of centrality from one and two-mode networks of the texts.

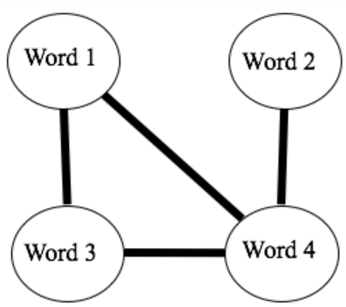
This analysis can be distinguished from the other textual analysis techniques in this dissertation because it goes beyond seeing what concepts of organizational culture are present in each STEM discipline. Instead, I assess whether STEM disciplines exhibit any shared meanings with each other (based on their texts) by extracting the relationships between concepts (Carley 1997). Specifically, I am able to examine which words and organizational cultural concepts are the most central in each discipline's mental model using centrality measures in both one and two-mode networks. The next sub-section describes the two network structures I use in this analysis.

Social Networks Used in Study 1

I use two network structures to examine the extent to which STEM culture is gendered: one-mode and two-mode networks. Generally speaking, a one-mode network consists of a set of actors (or nodes) and their ties to each other (or edges), and a two-mode network consists of a set of actors and their ties to a separate set of actors (or affiliates; Wasserman and Faust 1994). In the one-mode network, the actors (nodes) are words, and the edges (ties) are the shared presence of a word in a STEM program's text. In the two-mode network, the nodes are the STEM programs, the affiliates are the organizational cultural concepts, and the ties are the presence of a concept in a program's text.

The first network structure is a one-mode network consisting of words tied to words, where the words are specific to each organizational cultural concept. The words are tied together based on their shared presence in a STEM program's text; for example, if the words "collaboration" and "team" are in the same program's text, they would have a tie. I use this network structure in study 1 to examine the relationships between words in each concept, which ultimately shows how concepts from organizational culture are related to each other. Figure 2 below is an example visualization of the network and network data matrix I use in this study.

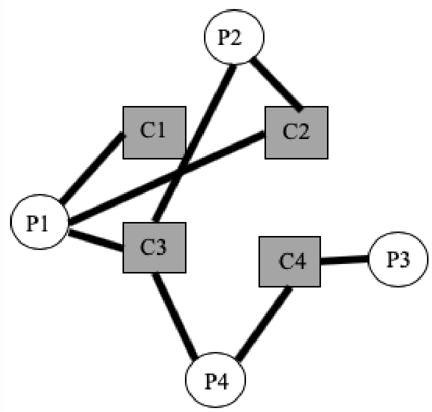
Figure 2: Example of the One-Mode Network Structure and Data Matrix Used in Study 1



	Word 1	Word 2	Word 3	Word 4
Word 1		0	1	1
Word 2	0		0	1
Word 3	1	0		1
Word 4	1	1	1	

The second network structure is a two-mode network consisting of STEM programs tied to organizational cultural concepts by the presence of a concept in the program texts; for example, the computer science program at Washington State University is tied to the organizational cultural concept *expectations of brilliance* if a word from this concept is in this program's text. This network comes from a document-term matrix, where the rows are the texts from each STEM program, and the columns are the organizational cultural concepts. I use this two-mode network structure to examine which concepts from organizational culture are the most central in each of the six STEM disciplines. Figure 3 below is an example of the two-mode network structure and data matrix.

Figure 3: Example of the Two-Mode Network Structure and Matrix Used in Study 1



	Concept 1 (C1)	Concept 2 (C2)	Concept 3 (C3)	Concept 4 (C4)
Program 2 (P1)	1	1	1	0
Program 2 (P2)	0	1	1	0
Program 3 (P3)	0	0	0	1
Program 4 (P4)	0	0	1	1

Justification of Words and Organizational Cultural Concepts

The words and organizational cultural concepts in both networks come from lexicons I created (see table 4 below). I created these lexicons by selecting words from four feminine and masculine concepts that arose from the gender and STEM literature: *socially-connected science* (feminine), *collaboration across disciplines* (feminine), *socially-disconnected science* (masculine), and *competition and expectations of brilliance* (masculine). I expanded the words in each lexicon by using a thesaurus. I justify the four concepts I used to create the lexicon in the sections below (see table 5 for the specific sources justifying each concept).

Table 4: Lexicons of Organizational Culture

Lexicon Title (Concept Measuring Organizational Culture)	**Words in Lexicon**
Collaboration across Disciplines (feminine)	team, collaboration, cooperation, cooperative, collaborative, interdisciplinary, disciplines, teamwork, relationship, communicate, communication, interpersonal, multidisciplinary, multi-disciplinary, integrative, overlap, incorporate, incorporative, versatile, combining, combination multifaceted, synthesize, synthesizing
Socially-Connected Science (feminine)	application, society, justice, value, useful, soft skills, service, service-learning, improve, connect, connection, impact, impactful, meaningful, care, caring, empathy, empathetic, diverse, diversity, friendly
Competition and Expectations of Brilliance (masculine)	competitive, competition, top, world-class, brilliant, genius, star, rank, competitive, competing, aggressive, ambitious, status, rigor, rigorous, ranking, drive, difficult, hard, ambition, ambitious, competing, strive, striving, push, accomplishment accomplish, aptitude, originality, prodigy, prowess, talent, talented, intelligence, intelligent, ability, power, powerful, leading, lead, potential, stellar
Socially-Disconnected Science (masculine)	individual, rational, static, sound, philosophical, fixed, stagnant, constant, central, original, uncommon, special, alone, distinct, specialized, exact, strict, rigid, order, particular, vigor, vigorous, flawless, robust, solid, well-constructed, theory, theoretical, abstract, formal, pure, methodological, complex, technical, indefinite, transcendent, ideal, idea

Collaboration across Disciplines

Collaborative work culture and interdisciplinary research, which I refer to as *collaboration across disciplines*, is an important concept in STEM culture. Women in STEM have been found to perceive a more negative workplace climate than men (Riffle et al. 2013). However, promoting a supportive, inclusive, and collaborative culture is one solution to improving the workplace climate for women in STEM (Fox 2000; Rhoten and Pfirman 2007; Kongar et al. 2008; Kasarda et al. 2010). This is not a surprise, as women tend to value collaborative cultures, interdisciplinary research, and working in teams while conducting their research (Kongar et al. 2008; Fox 2000; Rhoten and Pfirman 2007; Smith-Doerr 2005). Besides collaboration, women also have been found to have proclivities towards interdisciplinary research (Rhoten and Pfirman 2007; van Rijnsoever and Hessels 2011), although bureaucratic structures provide a disadvantage for women in this area. Because collaboration across disciplines has been shown by the literature to be related to women, I label this concept as being a part of feminine organizational culture.

Socially-Connected Science

Besides *collaboration across disciplines*, I label *socially-connected science*, or science done with the purpose of contributing to society, as a part of feminine organizational culture. Women in STEM have been found to value traits such as social consciousness, public welfare, and social responsibility more so than men, who typically have more masculine values such as technological leadership (Cech 2015; Cech 2014; Canney and Bielefeldt 2015). In several studies, women have expressed interest in doing research that impacts society (Smith-Doerr et al. 2016; Kongar et al. 2008; Espinosa 2011). Related to *collaboration across disciplines*, women engage in socially conscious research that crosses disciplines and serves multiple stakeholders and missions outside of academia (Rhoten and Pfirman 2007).

Socially-Disconnected Science

The first masculine cultural concept is *socially-disconnected science*, which is defined as a scientific paradigm that is limited in scope (such as to only theoretical pursuits) and does not consider addressing social issues (Uriarte et al. 2007). This is directly opposed to the paradigm of *socially-connected science*, which I explain in the subsection above. As a result of women's proclivity towards scientific research that helps society, women leave STEM fields in part because of the inability of professors to make STEM education accessible or align with their goals to contribute to society (Espinosa 2011), making *socially-disconnected science* problematic for women in STEM (Uriarte et al. 2007; Rhoten and Pfirman 2007).

Competition and Expectations of Brilliance

The last masculine cultural concept I consider in this study is c*ompetition and expectations of brilliance*. The social science literature has shown that academic fields with the highest expectations of brilliance, or the expectation that one has to be brilliant or a genius to succeed in their field, have the lowest proportions of women (Leslie et al. 2015; Storage et al. 2016). Examples of some of these fields where expectations of brilliance are high are physics and philosophy. Relatedly, it has also been found that women perceive competitive environments in STEM more negatively than men (Gupta 2007), and one study found that one of the biggest factors discouraging female students in engineering was competition in the classroom (Burger 2009). Ultimately, competition has been found to be highly discouraging to women in STEM fields (Uriarte et al. 2007; Smith-Doerr et al. 2016; Burger 2009).

Table 5: Concepts and Indicators to Measure Organizational STEM Culture in Texts	
Masculine STEM Culture	
Concepts/Indicators	**Sources**
Competition and Expectations of Brilliance • Emphasizing a program as rigorous or highly ranked • Competitiveness • Expectations of Brilliance	Parson 2016; Mervis 2011; Gupta 2007; Uriarte et al. 2007; Smith-Doerr et al. 2016; Leslie et al. 2015; Storage et al. 2016
Socially-Disconnected Science • Linear, rational thinking • Knowledge as static and unchanging • Socially disconnected science • Emphasis on theoretical rather than practical application of science	Parson 2016; Smith-Doerr et al. 2016; Kongar et al. 2008; Espinosa 2011
Feminine STEM Culture	
Concepts/Indicators	**Sources**
Collaboration across Disciplines • Teamwork collaboration • Emphasis on interdisciplinary research • Learning as a collaborative process • Personalized education	Kongar et al. 2008; Smith-Doerr 2005; Fox 2000; Rhoten and Pfirman 2007; Sharpe and Sonnert 1999; Sonnert et al. 2007; Griffith 2010; Bettinger and Long 2005; Schneeweis and Zweimuller 2012; Stearns et al. 2016; Robst et al. 1998; Price et al. 2010; Griffith 2010; Marra et al. 2012; Chanderbhan-Forde et al. 2012; Dalrymple and Evangelou 2006; Daily et al. 2007
Socially-Connected Science • Involvement in social justice issues • Diversity and Inclusion • Research that has practical applications • Field is considered dynamic • Emphasis on problem-solving	Smith-Doerr et al. 2016; Espinosa 2011; Kongar et al. 2008; Rhoten and Pfirman 2007; Berjerano and Bartosh 2015; Parson 2016; Heflin 2015

Use of Centrality Scores to Operationalize Organizational Culture

Following Diesner et al's (2005) study of mental models about data privacy and security among people from India, I use centrality measures to examine the relationship between the words from each concept and to determine which concepts of organizational culture are most central in each of the six disciplines. Also similarly to Diesner et al. (2005), I argue that the concepts that are most central in the network are the strongest representations of each discipline's idea of organizational culture. I use four measures of centrality: closeness, betweenness, degree, and eigenvector. I calculate closeness and betweenness centrality for the one-mode network structure, and degree and eigenvector centrality for the two-mode network. I use UCINET software to calculate all of the centrality measures (Borgatti et al. 2002).

In the one-mode network, I first use *closeness centrality* to examine how close a word from each organizational cultural concept is to the other words based on their ties to the STEM program texts. Closeness centrality is the sum of the shortest paths from one node to another node (Wasserman and Faust 1994). The equation for closeness centrality $C_C(n_i)$ is as follows:

$$C_C(n_i) = \left[\sum_{j=1}^{g} d(n_i, n_j)\right]^{-1}$$

where $d(n_i, n_j)$ is the distance between node *i* and node *j*.

Next, I use *betweenness centrality*, which is the probability that a node *j* needs a node *i* (the node whose centrality is being measured) in order to get to node *k* via the shortest path (Wasserman and Faust 1994). In the context of this study, betweenness centrality measures the probability a word is positioned on the shortest path between any other pair of words (see Diesner et al. 2005). The formula for betweenness centrality $C_B(n_i)$ is as follows:

$$C_B(n_i) = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}}$$

where $g_{jk}(n_i)$ is the number of the shortest paths from node *j* to node *k* that contain actor *i*, and g_{jk} is the number of paths that pass through actor *i*.

In the two-mode network, I first use *degree centrality*, which is the number of ties on a node divided by the number of affiliates. In this study, the nodes are the STEM programs, and the affiliates are the organizational cultural concepts. I use this measure to examine which concepts are the most "active" in the network.

Lastly, I use *eigenvector centrality*. Eigenvector centrality says that actors are most central if they are tied to the most central affiliation nodes, which can be conceptualized as a weighted degree measure and measure of influence (Borgatti and Everett 1997). Eigenvector centrality in a two-mode network is calculated by the following equation:

$$\sqrt{\frac{1}{2n_o}}$$

where n_o is the size of the set of vertices to which an actor (or node) belongs.

In this study, eigenvector centrality is used to see which STEM programs are affiliated with the most central concepts (i.e., concepts with the greatest degree centrality based on the presence of these concepts in program texts).

One-Mode Network Results

Table 6 below shows the descriptive statistics for each of the six one-mode networks, which consist of words tied to words based on their shared presence in each program's text. The words represent the organizational cultural concepts. Each network ranges in size from 78 to 88 nodes (in this case, the words making up each organizational cultural concept), with physics and computer science having the most words present from the organizational cultural concepts (nodes), and biology having the least. There are also 1648 to 2840 edges (in this case, the ties between words based on their shared presence in a program's text), with computer science having the most ties and biology having the least ties. An edge (or tie) is present if two words appear in the same program's text. The average weighted degree is the average number of ties for each word (node). As shown, even though physics has among the highest number of words (nodes), it has the lowest average weighted degree due to its lower number of ties in the network.

Table 6: Descriptive Statistics for the One-Mode Networks of Words

	Number of Nodes (Words)	Number of Edges (Ties)	Average Weighted Degree
Biology	78	1729	130.821
Chemistry	87	1870	151.448
Computer Science	88	2840	278.818
Mathematics	79	2170	222.532
Physics	88	1932	123.727
Psychology	85	1648	154.541

The purpose of this analysis of one-mode networks is to see which words from the organizational cultural concepts tend to be grouped together in a program's text, forming a network of concepts (mental model). This mental model, in turn, gives descriptive information about the extent to which each STEM discipline's culture is gendered. Table 7 shows the results for the one-mode network *closeness* and *betweenness centrality* scores for the top 5 most central words in the network of concepts in the six disciplines. Recall that *closeness centrality* examines the sum of the shortest paths of one word to another based on their ties to a program's text, which is a measure of how close they are in the network of concepts. *Betweenness centrality* is the probability a word is positioned on the shortest path between any other pair of words. As shown, the most central word by both centrality measures for all of the networks but chemistry is "skills," which is a word from *socially-connected science* (a feminine organizational cultural concept). For chemistry, "society" is the top word in the network of concepts, which is similar to the other five disciplines since "society" is also a word from the concept *socially-connected science*.

The differences in feminine versus masculine organizational cultural concepts start to emerge when looking at the second most central words in the mental models. All of the disciplines with at least gender parity in the proportion of female graduates (biology, chemistry, mathematics, and psychology) have words from the two feminine organizational cultural concepts (*collaboration across disciplines* and *socially-connected science*) as their second most central word for both measures of centrality. However, computer science and physics (male-dominated fields) have at least one word from *socially-disconnected science*, a masculine organizational cultural concept, for their second-most central word. The remaining top five most central words vary by the centrality measure, but every discipline except computer science tends to have words from feminine organizational cultural concepts.

Table 7: One-Mode Network Closeness and Betweenness Centrality for the Top 5 Most Central Words from the Organizational Cultural Concepts (Centrality Scores in Parentheses)

STEM Discipline	Most Central Word		2nd Most Central Word		3rd Most Central Word		4th Most Central Word		5th Most Central Word	
	Close.	**Between.**	**Close.**	**Between.**	**Close.**	**Between.**	**Close.**	**Between.**	**Close.**	**Between.**
Biology	Skills (.775)	Skills (.082)	Diversity (.705)	Society (.062)	Diverse (.699)	Diversity (.043)	Disciplines (.681)	Individual (.042)	Service & Society (tie) (.675)	Rigorous (.041)
Chemistry	Society (.835)	Society (.127)	Skills (.768)	Skills (.082)	Top (.748)	Top (.075)	Disciplines (.748)	Service (.048)	Inter-disciplinary (.735)	Disciplines (.044)
Computer Science	Skills (.861)	Skills (.045)	Theory (.798)	Service (.044)	Technical (.798)	Top (.041)	Top (.798)	Theory (.035)	Inter-disciplinary (.798)	Technical (.032)
Mathematics	Skills (.868)	Skills (.093)	Disciplines (.798)	Disciplines (.067)	Theory (.790)	Theory (.046)	Society (.745)	Society (.029)	Complex (.745)	Diverse (.029)
Physics	Skills (.821)	Skills (.095)	Theoretical (.744)	Theoretical (.078)	Inter-Disciplinary (.737)	Society (.060)	Society (.737)	Inter-Disciplinary (.050)	Diverse (.737)	Collaboration (.049)
Psychology	Skills (.824)	Skills (.130)	Service (.771)	Service (.071)	Care (.706)	Theoretical (.060)	Individual (.694)	Care (.053)	Society (.694)	Society (.053)

Figure 4 is a visualization of the one-mode networks for each of the six disciplines, which I use to illustrate the one-mode network results from table 7 in the context of the broader network. I created all of the network visualizations in this study using Gephi, an open-source network analysis visualization software. I filtered the visualization to the words that had the average number of ties or above for clarity in the network drawings[5] (see table 6 above for these statistics). The size of the circles (nodes containing each word) represent the centrality of words in each network; the larger the nodes, the more central the word is in the network. The darker the lines (ties), the higher the number of ties between each word (i.e., the higher the number of times each pair of words appears in the same program's text).

These network visualizations show that the most central words across each one-mode network (from table 7) are very distinct and central to all six disciplinary one-mode networks. While each one-mode network contains a mix of words from both feminine and masculine organizational cultural concepts, the visualizations show differences in the ties and tie strength between these concepts.

For biology, feminine words such as "care", "diversity", and "service" have very strong ties to the feminine word "skills" (the most central word in the network), but a few masculine words such as "ability" and "rigorous" also have strong ties to "skills" in this network. The largest nodes, which include the words "skills," "diverse," "disciplines," diversity," and "service," represent some of the words that are the most central by the betweenness and closeness centrality measures. Feminine words such as "service," "skills," and "discipline" have heavy ties to "diversity," while the masculine word "status" is also heavily tied to "diversity."

[5] Filtering visualizations of large networks in order to more clearly visualize the network is common practice; see Bastian et al. 2009.

All in all, the feminine words in the biology network tend to be the most central and most heavily tied to other feminine words, although there are still some masculine words with heavy ties to central feminine words, showing that even STEM disciplines with higher representations of women contain a certain degree of masculine organizational culture.

In chemistry, the feminine words “service,” “disciplines,” “interdisciplinary,” and “skills” are heavily tied to “society,” but there are several masculine words with heavy ties to “society,” such as “central,” “top,” “leading,” and “rigorous”. The largest nodes represent the most central words in the network, such as “society,” “skills,” “interdisciplinary,” and “central.” Compared to biology, there is a larger presence of masculine words that are central or are heavily tied to one of the central feminine words, which makes sense given the smaller proportion of women in chemistry than biology.

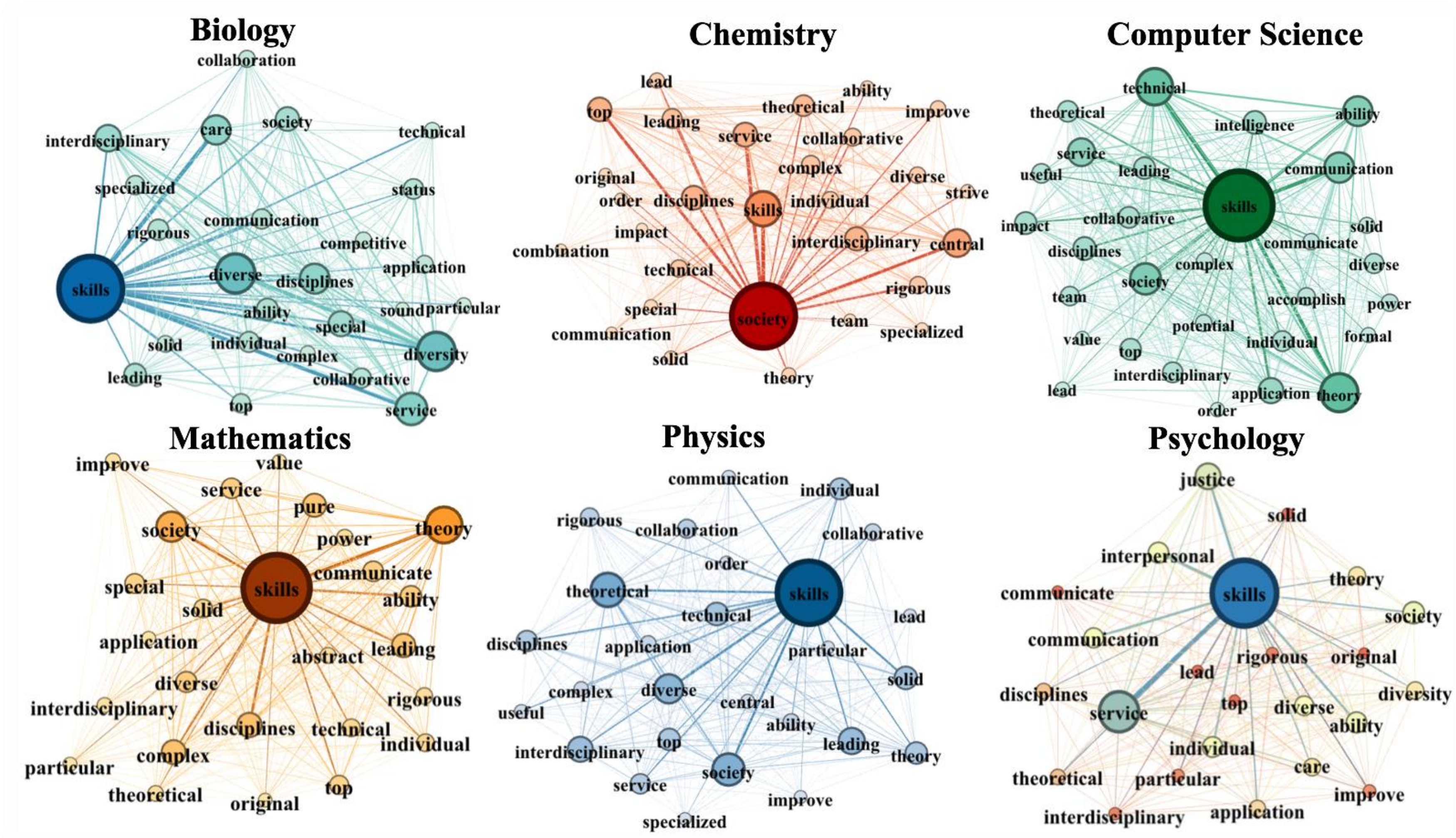
For computer science, there are some feminine words with heavy ties to “skills” (e.g., “service,” “society,” and “collaborative”), but it is notable how heavily tied and central some of the masculine words are positioned in the network due to their large node size in the visualization (e.g., “theory,” “ability,” and “technical”). The masculine words “technical” and “ability” are not only among the most central nodes in the network, but are also very heavily tied to one another, as well as to the most central and feminine word “skills,” based on the thick lines in the network graph. The high centrality of masculine words and lower centrality of feminine words makes sense given the high male dominance of computer science, but the high centrality of some feminine words (e.g., “communication,” “service,” “society,” and “interdisciplinary”) provides some nuance to evaluating the organizational culture of computer science.

In Mathematics, the feminine words “society” and “disciplines” have strong ties to “skills”, but so do several masculine words such as “theory”, “ability”, and “leading”. The largest three nodes in the graph are for the words “skills,” “society,” and “theory,” which are among the top 5 most central words. However, many masculine words (e.g., “top,” “rigorous,” and “technical”) and feminine words (e.g., “interdisciplinary,” “service,” and “communicate) have smaller nodes and lower centrality. These results suggest that mathematics is less gendered than the other STEM disciplines, at least in terms of this mental model. This makes sense given the gender parity of mathematics graduates at the undergraduate level.

Physics has several masculine words that have heavy ties to “skills” (e.g., “theoretical”, “leading”, and “technical”), but has several feminine words with heavy ties as well (e.g., “society”, “interdisciplinary”, and “diverse.” The high centrality of the masculine words “theoretical,” “technical,” and “leading” makes sense given the male dominance of physics, but like computer science, there is a large presence of feminine words in its mental model.

Psychology is perhaps the most distinct of the six networks, most likely because it is a social science. The feminine words of “justice,” “interpersonal,” “communication” and “service” clearly have the strongest ties to the most central word (“skills”), while the masculine words “ability” and “theory” also have strong ties, but are not nearly as central in the network as shown by their node size. The centrality of feminine words in psychology’s one-mode network makes sense given the female dominance of the discipline.

Figure 4: Visualizations of the One-Mode Network of Concepts for the Six STEM Disciplines



The one-mode network results show there are both similarities and differences in the mental models of the six disciplines. While all six mental models have feminine words for at least some of the top five most central words, computer science contains several masculine words that are very central in its network of concepts. Physics, a male-dominated field, surprisingly only contains one masculine word ("theory") in its top five most central words, but "theory" is the second most central word in its network. Interestingly, the two measures of centrality (*closeness* and *betweenness*) agree for most of the results, but tend to diverge for the fourth and fifth most central words. The visualizations confirm these results and show the nuance in centrality, as well as tie strength to the most central words. While most of the disciplines have a mixture of feminine and masculine words with high centrality and tie strengths, a few of the disciplines stand out. Biology and psychology, two disciplines with a higher proportion of female graduates than men, have a larger presence of feminine words that are central and heavily tied to other feminine words in their mental models, while computer science has very central masculine words that are heavily tied to both feminine and masculine words.

Two-Mode Network Results

The two-mode network structure results are in tables 8 through 10, which show the degree and eigenvector centrality of the four organizational cultural concepts (*collaboration across disciplines*, *socially-connected science*, *socially-disconnected science*, and *expectations of brilliance*), as well as the rankings for the centrality of the concepts by discipline and the ranking of centrality scores across disciplines. These results ultimately show which organizational cultural concepts are most central in the network of concepts forming the mental models for the six STEM disciplines. Figure 5 is a visualization of each two-mode network. The large nodes are the four organizational cultural concepts, and the larger and darker the node, the greater the centrality of the organizational cultural concept.

As shown by tables 8 and 9, all of the disciplines except physics have *socially-connected science*, a feminine concept, as their most central organizational cultural concept in their mental model. Physics' most central organizational cultural concept is *socially-disconnected science*, a masculine concept. Interestingly, the least central organizational cultural concept for biology and psychology is *expectations of brilliance and competition*, a masculine concept. The remaining STEM disciplines have *collaboration across disciplines* as their least central concept. It is important to note that chemistry, computer science, and mathematics have the same ordering for the ranking of concepts of *socially-connected science*, *socially-disconnected science*, *expectations of brilliance and competition*, and *collaboration across disciplines*. Biology and psychology also have the same rankings of *socially-connected science, socially-disconnected science*, *collaboration across disciplines,* and *expectations of brilliance*. When ranking the centrality scores across disciplines, computer science has the highest *degree centrality*[6] scores of the six disciplines for three of the four concepts: *collaboration across disciplines*, *socially-disconnected science*, and *expectations of brilliance and competition*. Psychology has the highest degree centrality score for *socially-connected science*. Physics has the lowest centrality scores of the six disciplines for both feminine organizational cultural concepts, while biology has the lowest score for *socially-disconnected science*, and psychology has the lowest score for *expectations of brilliance and competition*.

---

[6] Degree centrality is the number of ties on a node (STEM programs) divided by the number of affiliates (organizational cultural concepts).

As shown by the centrality scores in table 8 and the visualizations of the two-mode networks in figure 5, the mental models of the six disciplines also differ by the strength of the centrality scores. The difference in the centrality scores between *socially-connected science* and *expectations of brilliance* is largest for biology, mathematics, and psychology. Chemistry, computer science, and physics may have *socially-connected science* as their first or second most central concept, but the centrality scores for *expectations of brilliance/competition* are a lot closer to each other than they are for biology, mathematics, and psychology.

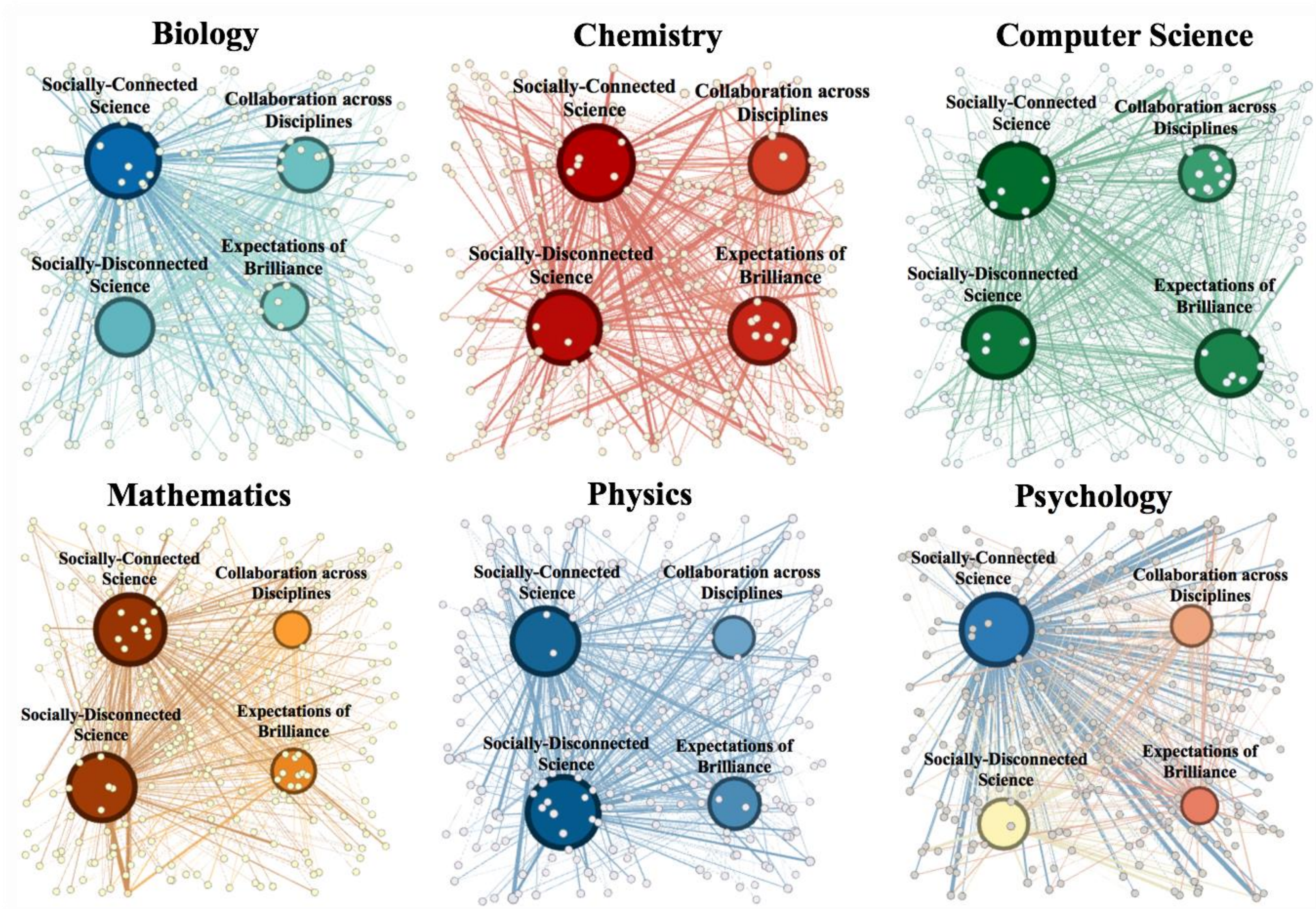
Table 8: Two-Mode Network Degree and Eigenvector Centrality Scores for the Four Organizational Cultural Concepts

	Collaboration across Disciplines		Socially-Connected Science		Socially-Disconnected Science		Expectations of Brilliance/Competition	
	Degree	Eigen.	Degree	Eigen.	Degree	Eigen.	Degree	Eigen.
Biology	.407	.458	.572	.630	.438	.504	.343	.374
Chemistry	.414	.447	.532	.554	.515	.518	.461	.475
Computer Science	.476	.413	.666	.552	.628	.526	.579	.498
Mathematics	.414	.408	.620	.569	.610	.558	.462	.446
Physics	.352	.382	.486	.524	.576	.591	.459	.480
Psychology	.388	.403	.724	.682	.490	.496	.337	.356

Table 9: Rankings of the Four Organizational Cultural Concepts by Centrality for the Six STEM Disciplines				
	Most Central	2nd Most Central	3rd Most Central	4th Most Central
Biology	Socially-Connected Science	Socially-Disconnected Science	Collaboration across Disciplines	Expectations of Brilliance/Competition
Chemistry	Socially-Connected Science	Socially-Disconnected Science	Expectations of Brilliance/Competition	Collaboration across Disciplines
Computer Science	Socially-Connected Science	Socially-Disconnected Science	Expectations of Brilliance/Competition	Collaboration across Disciplines
Mathematics	Socially-Connected Science	Socially-Disconnected Science	Expectations of Brilliance/Competition	Collaboration across Disciplines
Physics	Socially-Disconnected Science	Socially-Connected Science	Expectations of Brilliance/Competition	Collaboration across Disciplines
Psychology	Socially-Connected Science	Socially-Disconnected Science	Collaboration across Disciplines	Expectations of Brilliance/Competition

Table 10: Rankings of Degree Centrality Scores for the Four Organizational Cultural Concepts between Disciplines				
	Collaboration across Disciplines	Socially-Connected Science	Socially-Disconnected Science	Expectations of Brilliance/Competition
1st	Computer Science	Psychology	Computer Science	Computer Science
2nd	Chemistry (tie)	Computer Science	Mathematics	Mathematics
3rd	Mathematics (tie)	Mathematics	Physics	Chemistry
4th	Biology	Biology	Chemistry	Physics
5th	Psychology	Chemistry	Psychology	Biology
6th	Physics	Physics	Biology	Psychology

Figure 5: Visualizations of the Two-Mode Network of Concepts for the Six STEM Disciplines



Comparing Mental Models from the One and Two-Mode Networks

There are many similarities in the mental models of the six STEM disciplines. Both the one-mode and two-mode networks show that *socially connected science*, a feminine organizational cultural concept, tends to be the most central concept or near the top for all six disciplines. However, differences emerge when considering the second most central words or concepts in the mental models. For example, when looking just at the one-mode network of words, computer science has mostly masculine words as the second through fifth most central words in the network, which is what one would expect, but physics has mostly feminine words in its top five most central words. The visualizations of the one-mode networks confirm these results, as the most central words and strongest ties among words tend to be from the masculine organizational cultural concepts for computer science, but not for the other STEM disciplines.

When looking at the two-mode network of programs tied to the four organizational cultural concepts, the rankings of the four concepts in terms of their centrality is very similar across fields, but subtle differences also emerge. Biology and psychology, the STEM fields with the largest proportions of female graduates, both have the feminine concept of *socially-connected science* as their most central concept, and the masculine concept of *expectations of brilliance* as their lowest-ranked concept. Physics' most central concept is *socially-disconnected science*, a masculine concept. When ranking degree centrality scores between disciplines, the feminine organizational cultural concepts generally have STEM disciplines with greater shares of women as having the highest degree centrality scores, and disciplines with lower shares of women with lower centrality scores. This relationship is similar for two masculine organizational cultural concepts: the highest degree centrality scores are both from computer science, the most male-dominated field in the sample, while the lowest scores are from biology and psychology, the disciplines with the highest proportions of female graduates in the sample. I explain these results in context of studies 2 and 3 and speculate on these results in Chapter 7: Discussion, Implications, and Conclusion.

CHAPTER 5: CULTURAL DIFFERENCES BETWEEN PROGRAMS ACROSS STEM FIELDS (STUDY 2)

In study 2, I address the second research question: To what extent are there cultural differences between programs across STEM fields? I justify this research question using Cheryan et al. (2017), a critical review of hundreds of gender and STEM studies published since 1990. In this article, the authors argue that STEM fields should not only be disaggregated when examining reasons for women's unequal representation in certain STEM disciplines over others, but that the masculine culture of certain STEM disciplines plays a large role in shaping the context behind decisions to enter or not enter STEM. My objective in this study is to examine if there are similarities or differences between the organizational cultures in six STEM disciplines: biology, chemistry, computer science, mathematics, physics, and psychology. Specifically, I see if there is any support for hypothesis 1: disciplines with higher proportions of female graduates will have more feminine cultures than disciplines with lower proportions of female graduates, and disciplines with lower proportions of female graduates will have more masculine cultures than disciplines with higher proportions of female graduates.

I begin this chapter by discussing the methods for Study 2, specifically the operationalization of "organizational culture" (the dependent variable) and the models I use in this analysis. I provide an overview of the method I use to classify program texts as masculine, feminine, and gender neutral, which is one way I operationalize organizational culture. Next, I provide more detail on the lexicon scoring method that is used to 1) provide an additional continuous operationalization of organizational culture and 2) to create a series of independent variables used in each machine learning algorithm to generate the organizational culture classifications. I then explain the four machine learning algorithms used to classify the program texts and present the results of an experiment that evaluated the percentage of correct classifications by discipline, machine learning algorithm, and training/test data split. After that, I summarize the analytical techniques that I use to examine if there are differences in organizational culture in the six STEM disciplines. I end this chapter with a discussion of Study 2's results.

Lexicon and Machine Learning Methods Used to Obtain the Dependent Variables

There are a few ways to use texts to create other variables. Lexicon methods use continuous "polarity" scores that are calculated from matching the words in a text to a pre-defined list of words (lexicon), such as scoring texts based on them having positive or negative sentiment (Taboada et al. 2011). On the other hand, supervised machine learning methods first train a statistical model for the use of predicting – in this case, classifying – a text as having a certain level of a dependent variable (e.g., positive versus negative) based on one or more independent variables (James et al. 2017).

I use lexicon and supervised machine-learning methods to operationalize organizational culture in the six STEM disciplines. Organizational culture (whose coding types I describe in more detail in the paragraphs below) is the dependent variable in this set of analyses. Specifically, I operationalize organizational culture in two ways: (1) as a set of continuous variables obtained from the lexicons in Study 1 (*collaboration across disciplines*, *socially-connected science*, *competition and expectations of brilliance*, and *socially-disconnected science*)[7] and (2) as a categorical variable obtained from supervised machine learning

[7]Although some of the dependent variables in Analysis 2 are the continuous measures for *collaboration across disciplines*, *socially-connected science*, *competition and expectations of brilliance*, and *socially-disconnected science*, it is important to note that I also use these variables as independent variables to train the supervised machine learning algorithms used to classify the

classifications (masculine, feminine, or gender neutral). I obtain the categorical measure of organizational culture by training several statistical models from machine learning algorithms with a training dataset, testing the accuracy of these models with a separate testing dataset, and then using the statistical models I trained to classify yet another separate sample of texts, which in this case are the final analysis data (see Figure 6 for a visualization of the supervised machine learning process for text classification). Although the supervised machine learning process results in classifying new texts, to get the initial classifications for the training data, I use the lexicons (see table 4 in Chapter 4 for the specific words used in each lexicon). Operationalizing organizational culture in these two ways is beneficial because it allows me to examine if STEM disciplines are gendered by certain aspects of organizational culture (e.g., *expectations of brilliance*), or if the culture as a whole tends to be masculine, feminine, or gender neutral.

Process for Operationalizing Organizational Culture

Before beginning the supervised machine learning classification of organizational culture, I needed to have a dataset where program texts were already classified as masculine, feminine, or gender neutral. Normally, text classification tasks are much simpler in nature (e.g., classifying texts as having positive or negative sentiment), so it is feasible to have humans classify a smaller sample of texts or use larger, pre-tagged datasets with "positive" and "negative" texts that are already classified (e.g., Ding et al. 2008; Taboada et al. 2011). Since classifying texts as having masculine, feminine, and gender neutral culture is more complicated than many text classification tasks in the statistics and computer science literature and has not been implemented

---

analysis texts as masculine, feminine, and gender neutral. Put differently, the lexicon scores for these variables are used as both the means to obtaining the machine learning classifications and the ends of the operationalization of organizational culture.

(making it impossible to obtain a pre-classified dataset), I combine lexicon and supervised machine learning techniques by first classifying the training/test dataset using lexicons I manually created (Zhang et al. 2011). The specific words and concepts in the lexicons I created are in table 4 of Chapter 4: The Extent to Which the Cultures of STEM Disciplines are Gendered.

I used R to count the number of word matches between the lexicon words and the words in the text, and I initially classified the training and test data as masculine, feminine, or gender neutral by obtaining the following lexicon score:

$$Lexicon\ Score = \frac{\#\ Feminine\ Lexicon\ Words\ -\#\ Masculine\ Lexicon\ Words}{Total\ Words\ in\ Document}$$

$$Initial\ Classification = \begin{cases} Masculine, & Lexicon\ Score < 0 \\ Neutral, & Lexicon\ Score = 0 \\ Feminine, & Lexicon\ Score > 0 \end{cases}$$

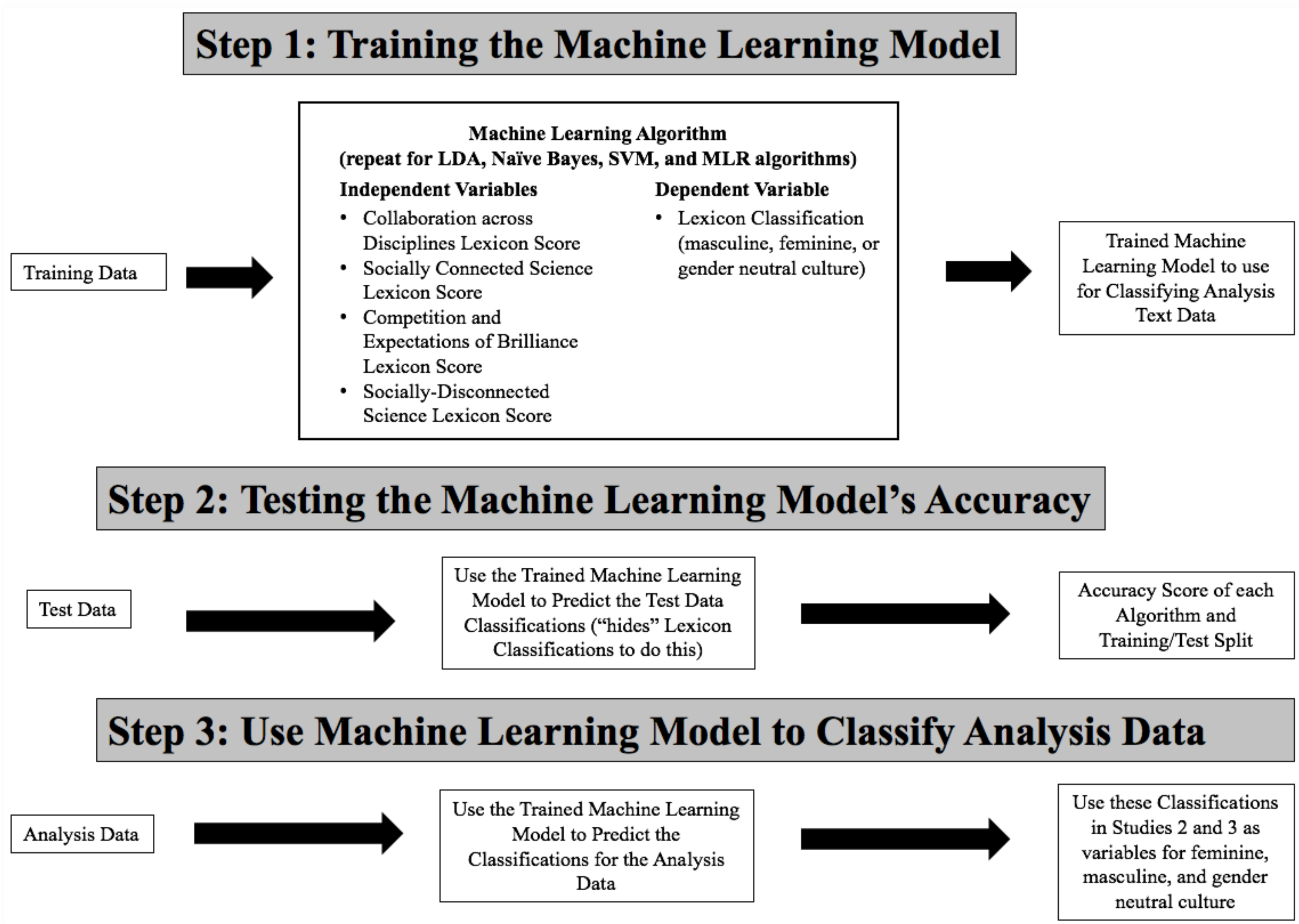
The texts in the training and test datasets were initially classified as feminine if this score was positive, masculine if the score was negative, and gender neutral if the score was equal to zero. The four independent variables in each machine learning algorithm (which also function as the continuous dependent variables in Study 2's regression models) are the individual lexicon scores obtained for *collaboration across disciplines*, *socially-connected science*, *competition and expectations of brilliance*, and *socially-disconnected science* (see table 4 in Chapter 4 for the specific words in each lexicon). These scores were calculated by taking the number of words that matched in each concept's lexicon and dividing it by the number of words in the text.

After I obtained the initial classifications from the training/test data lexicon scores, I randomly split the sample from the training/test data into separate training and testing datasets. I performed these splits with the following common ratios in machine learning applications: 60%

training and 40% testing, 70% training and 30% testing, and 80% training and 20% testing. I then implemented the text classification in three steps: 1) training the machine learning model, 2) testing the machine learning model's accuracy, and 3) using the machine learning model to classify a new set of texts (in this case, the final analysis data). I visualize these steps in Figure 6, and describe them in more detail below.

In step 1 of the text classification process, I trained the machine learning models used to classify the texts as having masculine, feminine, or gender neutral organizational culture. In these machine learning models, I used the lexicon scores for *collaboration across disciplines*, *socially-connected science*, *competition and expectations of brilliance*, and *socially-disconnected science* as the independent variables, and the initial classifications as feminine, masculine, and gender neutral calculated from the lexicon scores as the dependent variable. The output of step 1 resulted in a machine learning model for each algorithm that was trained by the training dataset. In step 2, I used the test data and the newly trained machine learning models to predict the classification of the test data. I tested the accuracy of the models by calculating the percentage of classifications that were correct (i.e., the percentage of predicted classifications that matched the original classifications from the lexicon scores). In step 3, I used the machine learning models that were trained in step 1 to classify a new set of texts, specifically, the final analysis data.

Figure 6: Text Classification Steps



Machine Learning Algorithms

To implement step 1 of the text classification (training the machine learning models), I utilized four machine learning algorithms: Linear Discriminant Analysis (LDA), the Naïve Bayes Classifier, Support Vector Machines (SVM), and Multinomial Logistic Regression. I chose these four algorithms because of their simplicity to implement and their common use in the scientific literature (e.g., Lilleberg et al. 2015; Soelistio and Surendra 2015; Boyd and Pennebaker). I use four algorithms rather than just one to ensure the highest accuracy possible. I summarize and describe the mathematical details of each machine learning algorithm used to classify the analysis data in the next four subsections.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is an algorithm that treats text classification similarly to regression analysis (Li et al. 2006). This algorithm involves 5 steps, which I summarize applied to the current text classification case. I used R to automatically implement steps 1-5 on the training datasets using the MASS package (Venables and Ripley 2002).

Step 1: Compute the Mean Vectors

LDA begins by computing the mean vectors for each independent variable, or the lexicon scores for *collaboration across disciplines* (CAD), *socially-connected science* (SCS), *competition and expectations of brilliance* (CEB), and *socially-disconnected science* (SDS). These mean vectors are calculated for the masculine, feminine, and gender neutral classes. This can be represented as follows:

$$\boldsymbol{\mu_i} = \begin{bmatrix} \mu_i(x_{CAD}\ score) \\ \mu_i(x_{SCS}\ score) \\ \mu_i(x_{CEB}\ score) \\ \mu_i(x_{SDS}\ score) \end{bmatrix} \text{ for } i = \text{masculine, feminine, or gender neutral class}$$

Step 2: Compute the Within-Class Variance-Covariance Matrix S_w and the the Between-Class Variance-Covariance Matrix S_B

Next, LDA computes the Within-Class and Between-Class Variance-Covariance Matrices. This step is done to "partition" the variance similarly to that done in an Analysis of Variance framework; you can think of the Within-Class Variance-Covariance matrix similarly to the sum of squares error, and the Between-Class Variance-Covariance matrix similarly to the sum of squares treatment. The following operations are used to calculate the Within-Class Variance-Covariance Matrix S_w:

$$S_w = \sum S_i$$

where $S_i = \sum(x - \mu_i)(x - \mu_i)'$, which is the sum of the variance-covariance matrices for each ith class, and μ_i is the mean vector for each ith class (calculated in Step 1).

The Between-Class Variance-Covariance Matrix S_B is calculated as follows:

$$S_B = \sum N_i(\mu_i - \mu)(\mu_i - \mu)'$$

where N_i is the sample size of each ith class, μ_i is the mean of the lexicon score for each ith class, and μ is the overall mean for the lexicon scores.

Step 3: Obtain the Linear Discriminants by Solving the Generalized Eigenvalue Problem for the Matrix $S_W^{-1}S_B$

Next, LDA obtains the linear discriminants, which can be conceptualized similarly to regression coefficients because they are linear combinations of the independent variables. The sum of the linear discriminants multiplied by their respective predictors yields a discriminant score that is used in part to classify the texts as masculine, feminine, or gender neutral. LDA computes these linear discriminants by solving the Generalized Eigenvalue Problem, which is characterized by the following equation:

$$\Sigma v = \lambda v$$

where v is an eigenvector of the variance-covariance matrix Σ, and λ is the eigenvalue.

*Step 4: Select the Linear Discriminants for the Linear Discriminant Coefficient Matrix **W***

In this step, LDA first sorts the eigenvectors by decreasing eigenvalues to select the eigenvectors with the largest eigenvalues, which contain the most information. After the eigenvector-eigenvalue pairs are sorted, the LDA constructs a 4×2 eigenvector matrix **W** that is from the two most informative, or largest eigenvector-eigenvalue pairs. Put more intuitively, **W** is a matrix of discriminant coefficients, which are similar to regression coefficients.

Step 5: Formation of the Linear Discriminant Model Used for Classification

At this point, the calculations for the LDA are complete. This step combines all the previous components to form the equation for classifying the text documents. This equation looks very similar to linear regression, and is represented as follows:

$$\mathbf{y} = \mathbf{XW}$$

where **y** is a vector of the new text classifications, **X** is a matrix containing the lexicon scores (predictor variables), and **W** is a matrix containing the discriminant coefficients. Table 11 below summarizes the mathematical notation used to explain Linear Discriminant Analysis.

Table 11: Summary of Mathematical Terms for Linear Discriminant Analysis

Mathematical Concept	**Notation in this Study**	**Definition applied to this study**	**Use in this Study**
Mean vectors	$\boldsymbol{\mu_i}$	Vector of means for each independent variable by the initial classification groups (masculine, feminine, and gender neutral)	Used to compute both the within and between-class variance-covariance matrices
Within-Class Variance-Covariance Matrix	S_W	The variance-covariance matrix (with the variances along the diagonal of the matrix and the covariances on the off-diagonals) that is pooled (within each of the three possible classification groups); comparable concept to the sum of squares error in ANOVA	Used to obtain the linear discriminants by solving the Generalized Eigenvalue Problem for the matrix $S_W^{-1}S_B$
Between-Class Variance-Covariance Matrix	S_B	The variance-covariance matrix between the three classifications; comparable concept to the sum of squares treatment in ANOVA	Used to obtain the linear discriminants by solving the Generalized Eigenvalue Problem for the matrix $S_W^{-1}S_B$
Generalized Eigenvalue Problem	$\Sigma v = \lambda v$	Solves the equation $\Sigma v = \lambda v$ for the matrix $\mathrm{S}_\mathrm{W}^{-1}\mathrm{S}_\mathrm{B}$ by finding a scalar λ (eigenvalue) and a vector v (eigenvector)	Used to obtain the linear discriminants by solving for the matrix $S_W^{-1}S_B$
Coefficient Matrix	$\mathbf{W}$	Matrix of discriminant coefficients; comparable concept to regression coefficients	Used in the model $\mathbf{y} = \mathbf{XW}$ to help classify texts as masculine, feminine, or gender neutral

Naïve Bayes Classifier

The Naïve Bayes Classifier is a simple classification algorithm that is often used in the context of text classification (McCallum and Nigam 1998). I implement the Naïve Bayes Classifier in R using the e1071 package (Meyer et al. 2017). This algorithm is called "naïve" because it makes the often unrealistic assumption that all of the attributes of the data, in this case the words in each text, are independent of each other. It is also based on Bayes' Rule (Bain and Engelhardt 1992), which can be used to find the probability of an event B_j happening given prior knowledge or information A. Bayes' Rule is formulated as follows:

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{P(A)}$$

where $P(A|B) = \frac{A \cap B}{P(B)}$ and $P(A) = \sum P(B_i)P(A|B_i)$, which is also known as the total probability.

More intuitively, you can apply Bayes' Rule to the text classification case at hand, where the goal is to get the posterior probability that the text is classified as masculine, feminine, or gender neutral in order to classify each text with the organizational culture category with the highest posterior probability. I illustrate this idea with the following formula:

$$posterior = \frac{prior \times likelihood}{total\ probability}$$

where the posterior probability is the probability that the text is classified as masculine, feminine, or gender neutral given the independent variables in the machine learning model (in this case, the lexicon scores for *socially-connected science*, *collaboration across disciplines*, *socially-disconnected science*, and *competition and expectations of brilliance*), the prior is the initial probability of drawing a masculine, feminine, or gender neutral text from the dataset, the likelihood is the probability of obtaining the data for each independent variable x_i given the classification as feminine, masculine, or gender neutral, and the total probability is the probability of obtaining the data for each independent variable. Mathematically, this can be written as follows:

$$p(C_j|x_i) = \frac{P(C_j)P(x_i|C_j)}{P(x_i)}$$

where $p(C_j|x_i)$ is the posterior probability, $P(C_j)$ is the prior probability, $P(x_i|C_j)$ is the likelihood, and $P(x_i)$ is the total probability, or $\sum P(C_j)P(x_i|C_j)$. Table 12 provides a review of the notation used in these equations and an explanation of mathematical terms.

Table 12: Summary of Mathematical Terms for the Naïve Bayes Classifier

Mathematical Concept	**Notation**	**Definition applied to this study**	**Use in this Study**
Posterior Probability	$P(C_j\|x_i)$	The probability of the classification C_j (j = feminine, masculine, or gender neutral) given the value of the independent variables x_i in the Naïve Bayes model	The highest posterior probability is used to classify each text as masculine, feminine, or gender neutral
Prior Probability	$P(C_j)$	The probability of the text having an initial lexicon classification C_j	Used in numerator to calculate the posterior probability
Likelihood	$P(x_i\|C_j)$	The probability of obtaining the data for each independent variable x_i given the classification as feminine, masculine, or gender neutral	Used in numerator to calculate the posterior probability
Total Probability	$P(x_i)$	The probability of obtaining the data for the independent variables x_i	Used in denominator to calculate the posterior probability

Support Vector Machines (SVM)

Unlike machine learning algorithms that use linear methods to classify texts (e.g., Linear Discriminant Analysis), Support Vector Machines (SVM) use non-linear decision boundaries for classification. Non-linear decision boundaries are made possible by using a kernel, or a function that quantifies the similarity of two observations (Karatzoglou et al. 2006; James et al. 2017; Meyer and Wien 2017). The support vector machines begin with the inner product of observations x_i and $x_{i'}$ for each pair of observations in the training dataset in order to provide a measure of similarity between points. The inner product is defined as follows:

$$\langle x_i x_{i'} \rangle = \sum x_{ij} x_{i'j}$$

This inner product is the basis for the support vector classifier, which is linear. However, since support vector machines expand the support vector classifier to include non-linear decision boundaries, a kernel function K is used on the inner product, which can be represented as $K(x_i, x_{i'})$. In the R implementation of SVM I use via the e1071 package (Meyer et al. 2017), the radial basis function is used as the kernel. The radial basis kernel K_{RB} used on the inner product of each pair of training data observations is notated as follows:

$$K_{RB}(x_i, x_{i'}) = e^{\frac{||x_i - x_i'||^2}{2\sigma^2}}$$

where $||x_i - x_i'||^2$ is the squared Euclidean distance, or squared length, between each pair of observations from the training dataset, and σ is a free parameter.

In the R implementation of SVM applied to this study, C-classification is used to minimize the errors, which is a similar idea to how errors are minimized in regression analysis. The following equation for the error is minimized:

$$\frac{1}{2} w'w + C \sum_{i=1}^{N} \vartheta_i$$

where *w* is a vector of coefficients, *C* is a constant, and ϑ_i are the parameters for handling the data that are not separable (i.e., cannot be separated by the non-linear decision boundaries). I summarize the mathematical notation for Support Vector Machines in table 13.

While support vector machines can only solve classification tasks with two classes, the implementation I use in R allows for more than two classes, making it appropriate for use in this study. This implementation in R uses the one-against-one technique, which means that it fits binary classifiers for each pair of classes, resulting in $N(N\text{-}1)/2$ binary classifiers (three binary classifiers in the case of this study), and then choosing the classification as masculine, feminine, or gender neutral based on the largest number of "votes" from these binary classifiers.

Table 13: Summary of Mathematical Terms for Support Vector Machines

Mathematical Concept	**Notation**	**Definition applied to this study**	**Use in this Study**
Inner product	$\langle x_i x_{i'} \rangle$	the sum of the products of two equal lists of numbers	Provides a measure of similarity between points as a basis for SVM
Kernel	$K(x_i, x_{i'})$	A function on the inner product that quantifies the similarity of the points	In SVM, used to expand the feature space, or the dimensions containing the variables, to include non-linear decision boundaries
Radial Basis Function	$K_{RB}(x_i, x_{i'})$	A kernel that uses the squared Euclidean distance, or squared length, between points	Used to make non-linear decision boundaries possible in SVM; specific kernel used in the R implementation
C-classification	$\frac{1}{2}w'w + C\sum_{i=1}^{N}\vartheta_i$	An error function that is minimized in SVM when applied to classification	Specific error function that is minimized in the R implementation of SVM

Multinomial Logistic Regression (MLR)

The last machine learning algorithm I use to derive the measures of organizational culture in this study is multinomial logistic regression (MLR). I implement this in R using the nnet package (Venables and Ripley 2002). In the two-class case, which is binomial logistic regression, the probability that a dichotomous variable Y takes on the value of 1 is modeled. To accomplish this, the logistic function must be used because it gives outputs between 0 and 1. The logistic function is defined as follows:

$$P(Y) = \frac{e^{\beta_0+\beta_1 X}}{1+\ e^{\beta_0+\beta_1 X}}$$

where *P(Y)* is the probability that the dichotomous variable takes on the value of 1, β_0 is the intercept, β_1 is the vector of regression coefficients, and *X* are the independent variables.
In logistic regression, we use the logit, or the inverse of the logistic function, as a link function to "linearize" this model since it currently is not linear. To linearize the logistic function, you first multiply the denominator of the right-hand side of the logistic function to obtain the following:

$$P(Y) + P(Y)e^{\beta_0+\beta_1 X} = e^{\beta_0+\beta_1 X}$$

By subtracting $P(Y)e^{\beta_0+\beta_1 X}$ from both sides and factoring out $e^{\beta_0+\beta_1 X}$ on the right-hand side, the equation reduces to the following:

$$P(Y) = e^{\beta_0+\beta_1 X}(1 - P(Y))$$

Finally, dividing by $1 - P(Y)$ and taking the natural logarithm of both sides yields the logistic regression equation resulting from the logit link function:

$$ln\left(\frac{P(Y)}{1 - P(Y)}\right) = \beta_0 + \beta_1 X$$

where $ln\left(\frac{P(Y)}{1-P(Y)}\right)$ is the logit for the dichotomous dependent variable Y, β_0 is the intercept, β_1 is the vector of regression coefficients, and X are the independent variables.

Multinomial logistic regression, whether it is used for machine learning purposes or statistical inference, can easily be extended from binomial logistic regression (formulated above). Since I use multinomial logistic regression to classify texts as masculine, feminine, or gender neutral, there are two equations for each multinomial logistic regression model. I summarize the mathematical terms for multinomial logistic regression in table 14.

Table 14: Summary of Mathematical Terms for Multinomial Logistic Regression			
Mathematical Concept	**Notation**	**Definition applied to this study**	**Use in this Study**
Logistic function	$\frac{1}{1+\ e^{-(\beta_0+\beta_1 X)}}$	Logistic regression model (before it is linearized with a link function)	Makes it possible to model the probability that a variable Y takes on the value of 1, which can be extended to the multinomial case
Logit	$ln\left(\frac{P(Y)}{1-P(Y)}\right)$	The inverse of the logistic function	Used as a link function to "linearize" the logistic regression model

Accuracy of Machine Learning Algorithms

The machine learning algorithms I use to classify the analysis texts as having masculine, feminine, or gender neutral culture all contain a certain margin of error. I evaluate the accuracy of each machine learning algorithm at each training/test split and for each discipline as the second step of classifying the text documents (see "Figure 6: Text Classification Steps"). I calculate the accuracy as the number of correct classifications divided by the total number of classifications. Table 15 below shows the results for the accuracy ratings, and figure 7 below visualizes these results. As shown, multinomial logistic regression either ties for the most accurate or is the most accurate algorithm for each STEM discipline.

Figure 7: Accuracy Ratings for LDA, MLR, NB, and SVM by Field and Training/Test Split

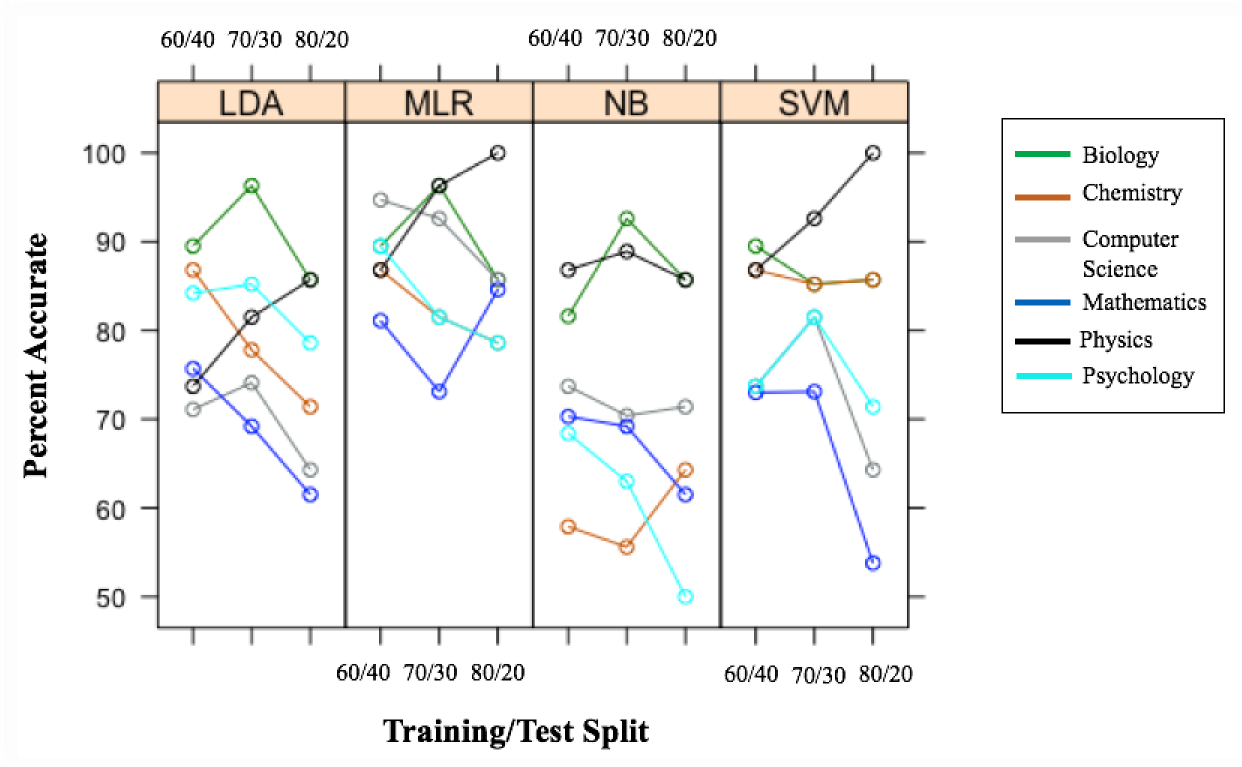


Table 15: Percentage Correct Classifications of 4 Machine Learning Algorithms with Various Training and Test Data Splits

	Machine Learning Algorithm and Training Data/Test Data Split											
	Linear Discriminant Analysis (LDA)			Naïve Bayes Classifier			Support Vector Machines (SVM)			Multinomial Logistic Regression (MLR)		
	60/40	70/30	80/20	60/40	70/30	80/20	60/40	70/30	80/20	60/40	70/30	80/20
Biology	89.5%	96.3%	85.7%	81.6%	92.6%	85.7%	89.5%	85.2%	85.7%	89.5%	96.3%	85.7%
Chemistry	86.8%	77.8%	71.4%	57.9%	55.6%	64.3%	86.8%	85.2%	85.7%	86.8%	81.5%	78.6%
Computer Science	71.1%	74.1%	64.3%	73.7%	70.4%	71.4%	73.7%	81.5%	64.3%	94.7%	92.6%	85.7%
Mathematics	75.7%	69.2%	61.5%	70.3%	69.2%	61.5%	73.0%	73.1%	53.8%	81.1%	73.1%	84.6%
Physics	73.7%	81.5%	85.7%	86.8%	88.9%	85.7%	86.8%	92.6%	100%	86.8%	96.3%	100%
Psychology	84.2%	85.2%	78.6%	68.4%	63.0%	50.0%	73.7%	81.5%	71.4%	89.5%	81.5%	78.6%

To further evaluate how the STEM discipline, machine learning algorithm, and training/test data split are related to the percentage of correct classifications, I conduct an Analysis of Variance (ANOVA) using a split-plot, or a two-way repeated measures design, in a randomized complete block design framework. I accomplish this by using the lmerTest package in R (Kuznetsova et al. 2017).

This ANOVA is a split-plot design because we can think of the six STEM disciplines as six "plots," or the whole plot factor, which are then treated by the four machine learning algorithms and subdivided by the three training/test splits applied to each machine learning algorithm. Put differently, this experimental design is a process of two experiments with different experimental units: one experiment for the four machine learning algorithms applied to each discipline, and another for the three training and test data splits applied to each machine learning algorithm. Figure 8 illustrates this experimental design, and the model for this design is represented as follows:

$$Y_{ijk} = \mu + \alpha_i + \gamma_k + \tau_{ik} + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

where Y_{ijk} is a vector of accuracy scores, μ is the intercept, α_i is the fixed effect for the machine learning algorithm, γ_k is the fixed effect for the discipline (or block), τ_{ik} is the whole discipline error (whole plot error), β_j is the fixed effect for the training/testing data split, $(\alpha\beta)_{ij}$ is the interaction between the machine learning algorithm and the training/testing data split, and ε_{ijk} is the split-plot error.

Figure 8: Illustration of Split-Plot Experimental Design for Evaluating the Percentage of Correct Classifications

BIOLOGY					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

CHEMISTRY					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

COMPUTER SCIENCE					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

MATHEMATICS					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

PHYSICS					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

PSYCHOLOGY					
LDA			Naïve Bayes		
60/40	70/30	80/20	60/40	70/30	80/20
SVM			MLR		
60/40	70/30	80/20	60/40	70/30	80/20

Table 16 shows the results for the ANOVA evaluating the impact of discipline, machine learning algorithm, and training/test data split on the percentage of accurate classifications. As shown, there are statistically significant differences in the mean percentage of correct classifications by discipline ($F(5,15)= 5.34$, $p = .005$) machine learning algorithm ($F(3,15)$, $p = .011$), and training/test split ($F(2,40)$, $p = .026$). The interaction between machine learning algorithm and training/test split was not statistically significant ($F(6,40)$ $p = .938$), implying that the relationship between the type of machine learning algorithm and accuracy did not differ by the training/test split.

Table 16: Split Plot ANOVA Results for Evaluating the Percentage of Correct Classifications

	Sum of Squares	Mean Squares	Degrees of Freedom	F Value
Discipline	902.83	180.56	(5, 15)	5.3441**
ML Algorithm	528.80	176.27	(3, 15)	5.2168*
Training/Test Split	270.09	135.04	(2, 40)	3.9968*
ML Algorithm * Training/Test Split	59.04	9.84	(6, 40)	0.2912

In addition to the split plot ANOVA results, I present the results in table 17 below to show the specific coefficient differences behind the significant fixed effects in this model. Biology, LDA, and the 60% training and 40% testing split are the reference categories[8] used to compare the coefficients. For the discipline main effect, chemistry, computer science, mathematics, and psychology all have a significantly lower percentage of correct classifications than biology. For the machine learning algorithm, the Naïve Bayes Classifier has a marginally

[8] Since this is an exploratory analysis, these reference categories are simply the first ones in alphabetical order, which is the default in R.

significantly lower percentage of correct classifications than linear discriminant analysis, but the other machine learning algorithms are not statistically significant. There are no statistically significant differences in the percentage of correct classifications by training/test split.

Table 17: Split Plot ANOVA Coefficient Differences for Evaluating the Percentage of Correct Classifications

	Coefficient (Std. Error)
Discipline	
Chemistry vs. Biology	-12.075* (4.605)
Computer Science vs. Biology	-12.150* (4.605)
Mathematics vs. Biology	-18.100** (4.605)
Physics vs. Biology	.125 (4.605)
Psychology vs. Biology	-13.142* (4.605)
Machine Learning Algorithm	
Naïve Bayes vs. LDA	-7.050^{+} (4.652)
SVM vs. LDA	.417 (4.652)
MLR vs. LDA	7.900 (4.652)
Training/Test Data Split	
70% Train /30% Test vs. 60% Train/ 40% Test	.517 (3.356)
80% Train /20% Test vs. 60% Train/ 40% Test	-5.633 (3.356)

Variables and Analytical Techniques: Study 2

I use two analytical techniques to examine whether there are cultural differences between programs across STEM disciplines: Multivariate Regression and Multinomial Logistic Regression. For both analyses, the independent variable is the field (computer science is the reference group since it is the most male-dominated field), but the dependent variables differ for the two analytical techniques (see the paragraphs below).

In the first set of models, I use Multinomial Logistic Regression to examine the relationship between STEM disciplines and the classification of a program as having masculine, feminine, or gender neutral organizational culture. This analytical technique is appropriate because one of the dependent variables I use to measure organizational culture is a categorical variable with three different possible codes (masculine, feminine, or gender neutral culture). I use the classifications that were obtained from the multinomial logistic regression algorithm with the 60% training and 40% testing data split since this algorithm and split had the highest accuracy rating (see table 15 above for these accuracy ratings).

Since the probability distribution for multinomial logistic regression is multinomial instead of binomial, there are $J - 1$ regression equations per model, or in this case, two equations per model since the dependent variable has three categories. Both regression equations for each model take this form:

$$ln\left(\frac{P_i(Y)}{1 - P_i(Y)}\right) = X\beta$$

where $P_i(Y)$ is the probability of the classification being feminine culture in the first equation and gender neutral culture in the second equation, X is a design matrix containing a column of 1's for the intercept and columns for each STEM discipline (computer science = 0), and β is a vector containing the intercept and multinomial logistic regression coefficients.

To make the interpretation of the multinomial logistic regression results easier, I use relative risk ratios (RRR) by exponentiating the logistic coefficients with base *e*. Relative risk ratios are interpreted similarly to odds ratios because they represent a positive association if they are above 1, a negative association if they are below 1, and no association if they are equal to 1.

The second set of models use Multivariate Regression to examine the relationship between STEM disciplines and the set of four continuous dependent variables that operationalize different aspects of organizational culture: *collaboration across disciplines* (CAD), *socially-connected science* (SCS), *competition and expectations of brilliance* (CEB), and *socially-disconnected science* (SDS). This method is appropriate because I am examining the relationship between a set of independent variables and multiple dependent variables. The equation for this model is as follows:

$$Y = X\beta + \varepsilon$$

where Y is a matrix representing the dependent variables, X is a design matrix containing a column of 1's for the intercept and columns for each STEM discipline (computer science = 0), β is a matrix containing the intercept and regression coefficients, and ε is a matrix of errors.

One advantage of using multivariate regression, rather than solely running separate OLS models, is the ability to perform tests on the combination of coefficients across the multiple dependent variables. In this case, I use a likelihood ratio test to test the combination of coefficients across the two feminine organizational cultural concepts and two masculine cultural concepts. To accomplish this, I use Multivariate Analysis of Variance (MANOVA) to test the five 1×2 coefficient vectors for each program (computer science is the reference) across the two response variables (one for the combination of the feminine concepts, and another for the combination of the masculine concepts), or $\boldsymbol{\beta}_{(i)}$, $i = 1 \ldots 5$. The null and alternative hypotheses tests on the regression coefficients across each pair of response variables are as follows:

$$H_0\text{: } \boldsymbol{\beta}_{(i)} = 0 \text{ and } H_a\text{: } \boldsymbol{\beta}_{(i)} \neq 0$$

Multinomial Logistic Regression Results

Table 18 presents the first set of results, from a multinomial logistic regression model examining the relationship between STEM discipline and the classification of organizational culture as feminine or gender neutral relative to masculine using relative risk ratios (RRR). The classifications as feminine, masculine, and gender neutral were taken from the multinomial logistic regression algorithm using the 60% training, 40% testing data split because it had the highest accuracy of correct classifications.

As shown, biology is about three times as likely as computer science to be classified as having feminine (RRR = 2.99, $p < .001$) or gender neutral culture (RRR = 2.94, $p < .001$) relative to masculine culture. Chemistry is about 1.7 times as likely as computer science to be classified as having gender neutral culture relative to masculine culture (RRR = 1.69, $p = .017$), but chemistry is not significantly more likely to be classified as having feminine culture relative to masculine culture. Mathematics is only marginally significantly more likely than computer science to be classified as feminine compared to masculine (RRR = 1.43, $p = .054$). While physics is not more likely than computer science to be classified as feminine, it is about twice as likely to be classified as gender neutral compared to masculine (RRR = 1.91, $p = .002$). Lastly, psychology is about 4.5 times as likely as computer science to be classified as having feminine culture (RRR = 4.52, $p < .001$) and gender neutral culture compared to masculine culture (RRR = 4.54, $p < .001$).

Overall, these results provide partial support for hypothesis 1; biology, psychology, and chemistry (three fields that have at least gender parity at the undergraduate level) are more likely than computer science (a male-dominated field) to be classified as having either feminine or gender neutral culture compared to masculine culture. However, math only has one marginally significant difference in being classified as feminine relative to masculine, despite having near gender parity at the undergraduate level. It is not surprising that there are no significant differences between computer science and physics in their classification as feminine compared to masculine since computer science and physics are both male-dominated disciplines.

Table 18: Multinomial Logistic Regression Results with Relative Risk Ratios for the Classification of Feminine and Gender Neutral Culture Relative to Masculine Culture

	Feminine	Gender Neutral
Field (Computer Science = 0)		
Biology	2.99*** (.20)	2.94*** (.23)
Chemistry	1.25 (.19)	1.69* (.22)
Math	1.43^{+} (.19)	1.34 (.23)
Physics	.88 (.20)	1.91** (.21)
Psychology	4.52*** (.21)	4.54*** (.24)
AIC	3698.88	

$^{+}$ *p < .10; * p<.05; ** p<.01; *** p<.001*

Multivariate Regression Results

Table 19 presents the multivariate regression results for the continuous organizational cultural concepts (*collaboration across disciplines, socially-connected science, socially-disconnected science, and competition)* as separate models, and the STEM discipline as the independent variable (computer science is the reference category since it is the most male-dominated field). Since the *collaboration across disciplines* model does not have a significant F test, meaning that we fail to reject the null hypothesis that the fit of the intercept-only model and the model with the variable for discipline are equal, I do not interpret the results for that model. Contrary to hypothesis 1, which is that disciplines with higher proportions of female graduates will have more feminine cultures than disciplines with lower proportions of female graduates, and disciplines with lower proportions of female graduates will have more masculine cultures than disciplines with higher proportions of female graduates, chemistry has a significantly lower score for the feminine cultural concept of *socially-connected science* than computer science ($\beta = -.002$, $p = .004$). However, psychology has both significantly higher scores for *socially-connected science* ($\beta = .003$, $p < .001$) and lower scores for the masculine cultural concepts of *socially-disconnected science* ($\beta = -.002$, $p < .001$) and *competition and expectations of brilliance* ($\beta = -.002$, $p < .001$) than computer science, providing some support for hypothesis 1. Likewise, biology has significantly lower scores for *socially-disconnected science* ($\beta = -.003$, $p < .001$) and *competition and expectations of brilliance* ($\beta = -.002$, $p < .001$) than computer science. Other results indicate that physics has a significantly lower score for *socially-connected science* ($\beta = -.002$, $p = .001$) and a marginally significantly lower score for *competition and expectations of brilliance* ($\beta = -.0007$, $p = .064$) than computer science. Interestingly, although mathematics has about an equal representation of women and men, it only had a marginally significantly lower score for *competition and expectations of brilliance* ($\beta = -.0008$, $p = .054$) than computer science.

Table 19: Multivariate Regression Results for the Continuous Scores of Masculine and Feminine Cultural Concepts

	Collaboration (Feminine)	Socially-Connected Science (Feminine)	Socially-Disconnected Science (Masculine)	Competition (Masculine)
Biology	-.0008^{+} (.0004)	-.001 (.0007)	-.003*** (.0005)	-.002*** (.0004)
Chemistry	-.0009* (.0004)	-.002** (.0007)	-.002*** (.0005)	-.001*** (.0004)
Math	-.0003 (.0004)	.0002 (.0007)	.0004 (.0005)	-.0008^{+} (.0004)
Physics	-.001* (.0004)	-.002** (.0007)	-.0006 (.0005)	-.0007^{+} (.0004)
Psychology	-.0008^{+} (.0004)	.003*** (.0007)	-.002*** (.0005)	-.002*** (.0004)
Adjusted R-Squared	.002	.04	.03	.03
F Statistic	1.739	14.14***	11.02***	10.46***

$^{+}$ *p < .10; * p<.05; ** p<.01; *** p<.001*

Tables 20 and 21 show the MANOVA results, which test if the combination of the multivariate regression coefficients for the two feminine and two masculine organizational cultural concepts are different from zero. This test shows if the combination of the multivariate regression coefficients for the two feminine and two masculine concepts are statistically significant or not while taking into account the correlation of the dependent variables. The combination of the two feminine concepts (*collaboration across disciplines* and *socially connected science*) are statistically significant for chemistry, physics, and psychology, and marginally significant for biology when compared to computer science. The combination of the two masculine organizational cultural concepts (*socially-disconnected science* and *expectations of brilliance and competition*) are statistically significant for biology, chemistry, and psychology when compared to computer science.

Table 20: MANOVA and Likelihood Ratio Tests for the Multivariate Regression Model (Feminine Continuous Organizational Cultural Concepts)

	Wilk's Lambda	*F*	*p*	*Significance Level*
Biology	.003	2.368	.094	+
Chemistry	.006	5.367	.005	**
Mathematics	.003	.229	.796	n.s.
Physics	.008	7.127***	<.001	***
Psychology	.012	11.058***	<.001	***

$^{+}$ *p < .10;* *p<.05*; ** *p<.01;* *** *p<.001* (two-tailed tests)

Table 21: MANOVA and Likelihood Ratio Tests for the Multivariate Regression Model (Masculine Continuous Organizational Cultural Concepts)

	Wilk's Lambda	*F*	*p*	*Significance Level*
Biology	.032	28.837	<.001	***
Chemistry	.014	12.559	<.001	***
Mathematics	.003	2.114	.121	n.s.
Physics	.003	2.2995	.101	n.s.
Psychology	.027	24.036	<.001	***

$^{+}$ *p < .10;* *p<.05*; ** *p<.01;* *** *p<.001* (two-tailed tests)

Overall, the multivariate regression results provide partial support for hypothesis 1 that STEM fields with higher proportions of women (biology, chemistry, psychology, and math) will have more feminine cultures than disciplines with lower proportions of women (computer science and physics), and vice versa. Interestingly, chemistry and physics actually score significantly lower than computer science on the feminine concept of *socially-connected science*, while psychology has a significantly higher score than computer science as I expected. However, for the two masculine organizational cultural concepts, the results are more in-line with hypothesis 1; biology, chemistry, and psychology, three fields with much higher proportions of female graduates than computer science, have significantly lower scores for both *socially-disconnected science* and *expectations of brilliance and competition* than computer science. Another interesting finding is that mathematics does not have significantly different results than computer science for any of the measures of organizational culture, despite having over twice the proportion of women than computer science at the undergraduate level.

CHAPTER 6: ORGANIZATIONAL CULTURE, ORGANIZING PROCESSES, AND THE PROPORTION OF FEMALE STEM GRADUATES (STUDY 3)

In study 3, I address the third and fourth research questions, which ask the following: 3) How is organizational culture related to the proportion of female bachelor's graduates in STEM? and 4) How are the departmental and institutional structures of STEM programs related to the proportion of female bachelor's graduates? I also address the last two hypotheses, which are that STEM programs with higher scores of feminine culture and lower scores of masculine culture (as demonstrated from the measures of organizational culture), or STEM programs housed in interdisciplinary departments or colleges, will have higher proportions of female bachelor's graduates than programs with masculine organizational cultures (hypothesis 2.1) or single-disciplinary departments and colleges of science[9] (hypothesis 2.2). In this chapter, I explain Study 3's variables, analytical technique, and results.

Study 3 Variables

I use the proportion of female STEM bachelor's graduates from the IPEDS Completions Survey (National Science Foundation 2016) in six disciplines (biology, chemistry, computer science, mathematics, physics, and psychology) as the dependent variable. The independent variables in part 1 of this study (organizational culture) are the continuous and categorical measures of organizational culture from Study 2 (*collaboration across disciplines*, *socially-connected science*, *competition and expectations of brilliance*, and *socially-disconnected science*,

[9] Because there were very few psychology departments that were housed in the college of science, the comparison is instead made to departments housed in the college of liberal arts.

and being classified[10] as masculine, feminine, or gender neutral). The independent variables in the second part of this study are the department and college-level structure variables measuring if these organizing processes are interdisciplinary or single-disciplinary. These are a series of dummy variables measuring if STEM programs are housed in interdisciplinary departments or colleges versus single-disciplinary departments or colleges of science (or colleges of liberal arts in the case of psychology). Table 22 gives the full list of department and college level dummy variables, specifying each group that is the reference category.

Table 22: Organizing Processes Variables by Discipline

Discipline	Department Level Structure Variables	College Level Structure Variables
Computer Science	• Computer Science and Engineering • Computer Science and Math • Science • Computer Science (reference category)	• Technology • Engineering • Arts and Sciences • STEM[11] • Science (reference category)
Biology	• STEM • Biology (reference category)	• STEM • Arts and Sciences • Health/Science • Science (reference category)
Mathematics	• Mathematics and Statistics • Mathematics and Computer Science • Science and Math • Math (reference category)	• STEM • Science and Math • Arts and Sciences • Science (reference category)
Chemistry	• STEM • Chemistry and Biology/Biochemistry • Chemistry (reference category)	• STEM • Arts and Sciences • Science (reference category)

[10] Just as what was done in analysis 2, the categorical measure of organizational culture using the masculine classification as the reference category and uses the classification from the multinomial logistic regression algorithm with a split of 60% training data and 40% testing data.
[11] STEM stands for science, technology, engineering, and math.

Table 22: Organizing Processes Variables by Discipline Cont.		
Physics	• STEM • Physics and Astronomy • Physics (reference category)	• STEM • Arts and Sciences • Science (reference category)
Psychology	• Social and Behavioral Sciences • Behavioral Science and Services • Psychology (reference category)	• Social Science • Professional Studies • Arts and Sciences • Science and Health • Liberal Arts (reference category)

The control variables for both parts of Study 3 are the institution's student population (total undergraduate and graduate students enrolled for credit), region (Northeast, Midwest, South, and West = reference), student-to-faculty ratio[12], private versus public (private = 1, public = 0), research 1 versus non-research 1[13] (research 1 = 1, non-research 1 = 0), degree of urbanization (suburb, town, rural, urban = reference), gender of the institution's President or head (female = 1, male = 0), and the proportion of female faculty institution-wide.

Analytical Technique

Since the dependent variable in this study is the proportion of female bachelor's graduates in STEM, using Ordinary Least Squares regression models would not be appropriate because proportions lie on the standard unit interval of [0, 1] and are often asymmetric in their distribution, creating problems with heteroscedasticity (the equal variance assumption) and

[12] Student-to-faculty ratio comes from IPEDS, and is measured as the total full-time students (plus 1/3 the number of part-time students) not in graduate or professional programs divided by the total number of full-time instructional staff (plus 1/3 the number of part-time instructional staff) who are not teaching in graduate or professional programs (IPEDS 2015). Instructional staff include any employees whose primary activity is instruction or instruction/research/public service and who are not medical school employees.

[13] I define Research 1 institutions using the Carnegie Classification from 2015 (IPEDS 2015), where institutions classified as "Doctoral Universities: Highest Research Activity" are research 1, and all other institutions are non-research 1.

inaccurate hypothesis tests in small samples (Cribari-Neto and Zeileis 2009). Using logistic regression to model proportions results in lower deviation, more accurate predictions, and a higher correlation between the predictions and observations than linear regression (Zhao et al. 2001). As a result, I chose to use logistic regression of grouped data (Hilbe 2011) by modeling the number of female bachelor's STEM graduates over the total number of bachelor's STEM graduates. This technique is appropriate to use to model the proportion of female STEM bachelor's graduates because the proportion of females is created by dividing two discrete variables (the number of female graduates and the total number of graduates), which can be thought of as a binomial random variable. This random variable *y/n* is binomially distributed as follows:

$$f(y_i;\ p_i n_i) = \binom{n_i}{y_i} p_i^{y_i}(1 - p_i)^{n_i - y_i}$$

where y_i is the number of successes (in this case, the number of female STEM bachelor's graduates) over the number of trials n_i (in this case, the total number of STEM bachelor's graduates) for each STEM program in the dataset, and p_i is the probability that the graduate is female. As a result, the logistic regression model is as follows:

$$ln\left(\frac{P_i(Y)}{1 - P_i(Y)}\right) = X\beta$$

where Y is a vector containing the binomially distributed random variable for the number of female STEM graduates divided by the total number of STEM graduates, X is a design matrix containing a column of 1's for the intercept and independent and control variable, and β is a vector containing the intercept and regression coefficients.

Alternatively, I could have used beta regression, a technique which assumes the dependent variable is beta distributed (i.e., lies between 0 and 1), to model the proportion of female bachelor's STEM graduates. While the beta distribution is very flexible and is another common method for modeling proportions, I chose logistic regression instead for several reasons. Beta regression does not work if any of the observations have exactly zero or one hundred percent female graduates. Beta regression also only models the proportion and does not take into account that the proportion is the number of female graduates out of the total count of graduates, meaning it would not differentiate the proportions of programs with a large number of STEM graduates and programs with a small number of STEM graduates[14].

Tables 23 through 28 below show the descriptive statistics for each of the six disciplines. For the categorical measure of organizational culture, psychology has the highest proportion of programs classified as feminine (55 percent), while physics has the lowest proportion (27 percent). Computer science has the highest proportion of programs classified as masculine (47 percent), and psychology has the lowest proportion (16 percent). Physics has the highest proportion of programs classified as gender-neutral (31 percent), while computer science has the lowest proportion (18 percent). All six disciplines have the highest proportion of their programs housed in single-disciplinary departments and in the College of Arts and Sciences.

[14] For example, two STEM programs could have 50 percent female graduates, but one might have 10 total graduates (5 females) and another 100 total graduates (50 females). Beta regression would only take into account that both departments have 50 percent female graduates, while logistic regression would model the number of female graduates out of the total number of graduates, taking into account that these programs have different sizes despite having the same proportion of female graduates.

Table 23: Descriptive Statistics - Biology

	Mean	*Std. Deviation*	*Minimum*	*Maximum*
Proportion of Female Bachelor's Graduates	0.613	0.129	0.000	1
Organizational Culture Classification				
Feminine Classification	0.508	0.501	0	1
Neutral Classification	0.263	0.441	0	1
Masculine Classification	0.229	0.421	0	1
Department-Level Structure of Program				
Biology	0.875	0.331	0	1
STEM	0.125	0.331	0	1
College-Level Structure of Program				
Science	0.135	0.342	0	1
STEM	0.219	0.414	0	1
Arts and Sciences	0.593	0.492	0	1
Health and Science	0.057	0.233	0	1
Total Student Population	8,487.178	10,937.330	410	80,494
Geographic Location				
Northeast	0.222	0.416	0	1
Midwest	0.283	0.451	0	1
South	0.360	0.481	0	1
West	0.135	0.342	0	1
Student to Faculty Ratio	14.640	4.290	3	30
Private	0.596	0.492	0	1
Research 1	0.077	0.268	0	1
Degree of Urbanization				
Suburb	0.232	0.423	0	1
Town	0.300	0.459	0	1
Rural	0.037	0.189	0	1
Urban	0.431	0.496	0	1
President is Female	0.226	0.419	0	1
Proportion of Female Faculty	0.463	0.086	0.193	0.763

Table 24: Descriptive Statistics - Chemistry

	Mean	*Std. Deviation*	*Minimum*	*Maximum*
Proportion of Female Bachelor's Graduates	0.478	0.171	0.000	1.000
Organizational Culture Classification				
Feminine Classification	0.359	0.481	0	1
Neutral Classification	0.254	0.436	0	1
Masculine Classification	0.386	0.488	0	1
Department-Level Structure of Program				
Chemistry	0.705	0.457	0	1
STEM	0.115	0.320	0	1
Chemistry and Biochemistry	0.180	0.385	0	1
College-Level Structure of Program				
Science	0.169	0.376	0	1
STEM	0.207	0.406	0	1
Arts and Sciences	0.624	0.485	0	1
Total Student Population	10,995.640	11,835.540	410	63,813
Geographic Location				
Northeast	0.251	0.434	0	1
Midwest	0.278	0.449	0	1
South	0.336	0.473	0	1
West	0.136	0.343	0	1
Student to Faculty Ratio	14.753	4.273	3	28
Private	0.536	0.500	0	1
Research 1	0.146	0.354	0.000	1.000
Degree of Urbanization				
Suburb	0.237	0.426	0	1
Town	0.241	0.428	0	1
Rural	0.024	0.152	0	1
Urban	0.498	0.501	0	1
President is Female	0.186	0.390	0	1
Proportion of Female Faculty	0.450	0.070	0.202	0.622

Table 25: Descriptive Statistics – Computer Science

	Mean	*Std. Deviation*	*Minimum*	*Maximum*
Proportion of Female Bachelor's Graduates	0.189	0.130	0.000	1.000
Organizational Culture Classification				
Feminine Classification	0.348	0.477	0	1
Neutral Classification	0.183	0.387	0	1
Masculine Classification	0.469	0.500	0	1
Department-Level Structure of Program				
Computer Science	0.669	0.471	0	1
Computer Science and Engineering	0.163	0.370	0.000	1.000
Computer Science and Math	0.152	0.359	0	1
Science	0.017	0.130	0	1
College-Level Structure of Program				
Science	0.114	0.319	0.000	1.000
Technology	0.069	0.254	0.000	1.000
Engineering	0.211	0.409	0.000	1.000
Arts and Sciences	0.433	0.496	0.000	1.000
STEM	0.173	0.379	0.000	1.000
Total Student Population	11,456.450	11,524.050	460	63,813
Geographic Location				
Northeast	0.210	0.408	0	1
Midwest	0.297	0.458	0	1
South	0.276	0.448	0	1
West	0.217	0.413	0	1
Student to Faculty Ratio	14.814	4.917	3	32
Private	0.559	0.497	0	1
Research 1	0.160	0.367	0.000	1.000
Degree of Urbanization				
Suburb	0.231	0.422	0	1
Town	0.200	0.401	0	1
Rural	0.021	0.143	0	1
Urban	0.548	0.499	0	1
President is Female	0.197	0.398	0	1
Proportion of Female Faculty	0.439	0.083	0.185	0.781

Table 26: Descriptive Statistics – Mathematics

	Mean	*Std. Deviation*	*Minimum*	*Maximum*
Proportion of Female Bachelor's Graduates	0.485	0.188	0.000	1.000
Organizational Culture Classification				
Feminine Classification	0.411	0.493	0	1
Neutral Classification	0.202	0.402	0	1
Masculine Classification	0.387	0.488	0	1
Department-Level Structure of Program				
Math	0.675	0.469	0	1
Math and Statistics	0.092	0.290	0	1
Math and Computer Science	0.178	0.383	0	1
Math and Science	0.051	0.221	0	1
College-Level Structure of Program				
Science	0.161	0.368	0	1
STEM	0.130	0.337	0	1
Science and Math	0.072	0.259	0	1
Arts and Sciences	0.634	0.483	0	1
Total Student Population	9,662.120	11,293.540	584	63,813
Geographic Location				
Northeast	0.284	0.452	0	1
Midwest	0.264	0.441	0	1
South	0.312	0.464	0	1
West	0.140	0.348	0	1
Student to Faculty Ratio	15.024	4.327	7	30
Private	0.521	0.500	0	1
Research 1	0.100	0.301	0.000	1.000
Degree of Urbanization				
Suburb	0.243	0.430	0	1
Town	0.260	0.440	0	1
Rural	0.034	0.182	0	1
Urban	0.462	0.499	0	1
President is Female	0.226	0.419	0	1
Proportion of Female Faculty	0.459	0.077	0.185	0.760

Table 27: Descriptive Statistics – Physics

	Mean	*Std. Deviation*	*Minimum*	*Maximum*
Proportion of Female Bachelor's Graduates	0.239	0.145	0.000	1
Organizational Culture Classification				
Feminine Classification	0.272	0.446	0	1
Neutral Classification	0.310	0.463	0	1
Masculine Classification	0.417	0.494	0	1
Department-Level Structure of Program				
Physics	0.634	0.482	0	1
Physics and Astronomy	0.262	0.441	0	1
STEM	0.103	0.305	0	1
College-Level Structure of Program				
Science	0.252	0.435	0	1
STEM	0.155	0.363	0	1
Arts and Science	0.593	0.492	0	1
Total Student Population	14,133.520	13,075.530	384	63,813
Geographic Location				
Northeast	0.234	0.424	0	1
Midwest	0.290	0.454	0	1
South	0.276	0.448	0	1
West	0.200	0.401	0	1
Student to Faculty Ratio	14.841	4.659	3.000	30
Private	0.472	0.500	0	1
Research 1	0.247	0.432	0.000	1
Degree of Urbanization				
Suburb	0.214	0.411	0	1
Town	0.193	0.395	0	1
Rural	0.017	0.130	0	1
Urban	0.576	0.495	0	1
President is Female	0.228	0.420	0	1
Proportion of Female Faculty	0.433	0.073	0.185	0.597

Table 28: Descriptive Statistics – Psychology

	Mean	Std. Deviation	Minimum	Maximum
Proportion of Female Bachelor's Graduates	0.771	0.095	0.250	1.000
Organizational Culture Classification				
Feminine Classification	0.548	0.499	0	1
Neutral Classification	0.289	0.454	0	1
Masculine Classification	0.163	0.370	0	1
Department-Level Structure of Program				
Psychology	0.864	0.353	0	2
Behavioral Science and Services	0.031	0.173	0	1
Social and Behavioral Sciences	0.109	0.312	0	1
College-Level Structure of Program				
Liberal Arts	0.122	0.328	0	1
Arts and Sciences	0.626	0.485	0	1
Science and Health	0.065	0.246	0	1
Social Science	0.143	0.351	0	1
Professional Studies	0.044	0.206	0	1
Total Student Population	7,833.935	9,963.451	119	80,494
Geographic Location				
Northeast	0.262	0.440	0	1
Midwest	0.276	0.448	0	1
South	0.310	0.463	0	1
West	0.153	0.361	0	1
Student to Faculty Ratio	14.255	4.314	5	31
Private	0.636	0.482	0	1
Research 1	0.073	0.260	0.000	1.000
Degree of Urbanization				
Suburb	0.238	0.427	0	1
Town	0.272	0.446	0	1
Rural	0.058	0.234	0	1
Urban	0.432	0.496	0	1
President is Female	0.279	0.449	0	1
Proportion of Female Faculty	0.469	0.091	0.167	0.800

Results: Organizational Culture and the Proportion of Female STEM Bachelor's Graduates

Tables 29 and 30 show the results from the logistic regression models examining the relationship between organizational culture and the proportion of female STEM bachelor's graduates. I measure the outcome as the number of female STEM graduates over the total number of graduates.

Table 29 shows the results for organizational culture as a categorical variable (masculine is the reference category). When organizational culture is measured in this way, the only statistically significant difference in the proportion of female bachelor's graduates is that computer science programs classified as gender neutral have a lower proportion of female bachelor's graduates than computer science programs classified as masculine (OR = .821, p = .015). While the independent variables are mostly not significant, some of the control variables are. For example, there are significant differences in the proportion of female graduates by region for biology, computer science, math, and psychology. Private institutions have significantly higher proportions of female STEM graduates in biology, chemistry, physics, and psychology programs than public institutions.

Table 29: Logistic Regression Results Modeling Categorical Measures of Organizational Culture and the Proportion of Female STEM Bachelor's Graduates in Six STEM Disciplines (Odds Ratios with Standard Errors in Parentheses)

	Biology	Chemistry	Computer Science	Math	Physics	Psychology
Organizational Culture (masculine = 0)						
Feminine	1.014	0.948	0.957	0.899	0.899	0.978
	(0.042)	(0.066)	(0.058)	(0.066)	(0.099)	(0.043)
Gender Neutral	0.949	1.085	0.821*	0.920	1.115	0.996
	(0.049)	(0.094)	(0.067)	(0.069)	(0.115)	(0.048)
Controls						
Total Student Population	1.000*	1.000	1.000	1.000	1.000	1.000
	(<.000)	(<.000)	(<.000)	(<.000)	(0.00001)	(0.00000)
Geographic region (West = 0)						
Northeast	1.111^{+}	0.860	0.745**	0.950	0.934	1.178**
	(0.067)	(0.092)	(0.062)	(0.101)	(0.123)	(0.064)
Midwest	0.954	0.833^{+}	0.655**	0.978	0.999	1.181**
	(0.055)	(0.084)	(0.055)	(0.109)	(0.122)	(0.066)
South	1.166**	0.995	0.786**	1.220*	1.146	1.250**
	(0.061)	(0.093)	(0.064)	(0.119)	(0.139)	(0.053)
Student/Faculty Ratio	1.005	1.010	0.942**	1.017^{+}	1.000	1.001
	(0.006)	(0.012)	(0.010)	(0.010)	(0.016)	(0.005)
Private	1.156**	1.361**	1.135	1.166	1.348*	1.099*
	(0.060)	(0.142)	(0.105)	(0.109)	(0.203)	(0.052)
Research 1	1.019	0.898	1.136	0.797*	0.985	0.985
	(0.061)	(0.103)	(0.102)	(0.087)	(0.147)	(0.048)
Degree of Urbanization (Urban = 0)						
Suburb	0.967	0.980	0.934	1.041	1.057	1.037
	(0.042)	(0.074)	(0.067)	(0.086)	(0.115)	(0.043)
Town	0.966	0.949	1.172	1.050	0.869	0.937
	(0.049)	(0.097)	(0.114)	(0.102)	(0.125)	(0.044)
Rural	1.067	0.342**	0.716	0.900	0.903	0.956
	(0.159)	(0.136)	(0.264)	(0.249)	(0.460)	(0.132)
Female University President	1.084^{+}	1.024	1.074	1.002	0.927	0.942^{+}
	(0.049)	(0.079)	(0.073)	(0.080)	(0.094)	(0.034)

Table 29 Cont.

	Biology	Chemistry	Computer Science	Math	Physics	Psychology
Proportion of Female Faculty	2.219** (0.685)	1.975 (1.155)	1.962 (0.834)	11.970** (6.308)	5.434* (3.812)	2.528** (0.676)
Pseudo R-Squared[15]	.045	.037	.125	.078	.045	.055

$^{+}$ *p < .10;* p<.05*; ** *p<.01; **** p<.001* (two-tailed tests)

Table 30 below presents the results for organizational culture measured as four continuous measures (*collaboration across disciplines, socially-connected science*, *socially-disconnected science*, and *expectations of brilliance*) net of controls. I centered the continuous measures of organizational culture at their means for this analysis because they are measured at a very low scale compared to the control variables. Centering eliminated the problem of very inflated odds ratios that arise with the use of their original scales[16]. Contradictory to my hypothesis that programs with stronger feminine cultures will have higher proportions of female graduates (hypothesis 2.1), I find a negative relationship between *socially-connected science* and the proportion of female graduates in biology programs (OR = .971, p = .023). Similarly, there is a negative relationship between *collaboration across disciplines* and the proportion of female

[15] Since these are logistic regression models, the total sums of squares cannot be partitioned into the regression sums of squares plus the error sums of squares like in OLS. As a result, the R-Squared used in OLS is not interpretable in logistic regression, and a Pseudo R-Squared must be used. I use McFadden's R-Squared to measure the Pseudo R-Squared, which is calculated as $1 - \frac{Log\ Likelihood\ (Full\ Model)}{Log\ Likelihood\ (Intercept\ Only\ Model)}$. While many of the Pseudo R-Squared values are low in this study, this is not problematic because 1) I am using these models to make inferences on the parameters, not for prediction and 2) low Pseudo R-Squared values are normal in logistic regression and are even not recommended to be reported since they do not have the same intuitive meaning as the R-Squared in OLS (Hosmer and Lemeshow 1998).

[16] Even though centering the continuous organizational culture variables eliminated the problem of inflated odds ratios, these results were very similar to the un-centered results in terms of statistical significance.

graduates in chemistry programs (OR = .936, p = .026). Similar findings contradict hypothesis 2.1 for the masculine organizational cultural concepts. Computer science programs with higher scores for *socially-disconnected science* actually have significantly higher proportions of female graduates than computer science programs with lower scores (OR = 1.108, p = .009). Similarly, biology programs with higher scores for *expectations of brilliance* have a marginally significantly higher proportion of female graduates (OR = 1.034, p = .053). However, there is a negative relationship between *expectations of brilliance* and the proportion of female bachelor's graduates for chemistry (OR = .934, p = .030) and a marginally significant relationship between *expectations of brilliance* and the proportion of female bachelor's graduates for physics programs (OR = .919, p = .077), providing partial support for hypothesis 2.1. Similarly, there is a positive relationship between *collaboration across disciplines* and the proportion of female psychology graduates (OR = 1.031, p = .070).

The controls present other interesting results. There are at least marginal differences in the proportion of female graduates by region for all of the disciplines but physics. In computer science, there is a negative relationship between student/faculty ratios and the proportion of female graduates (OR = .942, p < .001). Private institutions have a higher proportion of female graduates in biology, chemistry, physics, and psychology programs than public institutions. There is also a significant and positive relationship between the proportion of female faculty institution-wide and the proportion of female graduates in biology, math, physics, and psychology.

Table 30: Logistic Regression Results Modeling Continuous Measures of Organizational Culture and the Proportion of Female STEM Bachelor's Graduates in Six STEM Disciplines (Odds Ratios with Standard Errors in Parentheses)

	Biology	Chemistry	Computer Science	Math	Physics	Psychology
Organizational Culture – Continuous Measures						
Collaboration across Disciplines (feminine)	1.005 (0.016)	0.936* (0.028)	0.996 (0.026)	0.992 (0.034)	0.966 (0.043)	1.031^{+} (0.018)
Socially-Connected Science (feminine)	0.971* (0.013)	1.021 (0.025)	0.968 (0.026)	0.990 (0.036)	0.938 (0.049)	1.013 (0.018)
Socially-Disconnected Science (masculine)	1.009 (0.021)	0.957 (0.033)	1.108** (0.043)	0.997 (0.036)	1.017 (0.045)	1.016 (0.016)
Expectations of Brilliance (masculine)	1.034^{+} (0.018)	0.934* (0.029)	0.997 (0.026)	1.011 (0.030)	0.919^{+} (0.044)	1.013 (0.016)
Controls						
Total Student Population	1.000* (<.000)	1.000 (<.000)	1.000 (<.000)	1.000 (<.000)	1.000 (<.000)	1.000 (0.00000)
Geographic region (West = 0)						
Northeast	1.094 (0.066)	0.860 (0.093)	0.747** (0.063)	0.948 (0.100)	0.932 (0.123)	1.187** (0.065)
Midwest	0.970 (0.056)	0.821^{+} (0.085)	0.665** (0.056)	0.986 (0.109)	0.977 (0.119)	1.155* (0.066)
South	1.140* (0.059)	0.972 (0.094)	0.781** (0.064)	1.214* (0.117)	1.187 (0.144)	1.235** (0.053)
Student/Faculty Ratio	1.004 (0.006)	1.015 (0.012)	0.942*** (0.010)	1.015 (0.010)	1.003 (0.016)	1.001 (0.005)
Private	1.118* (0.057)	1.406** (0.148)	1.105 (0.103)	1.168 (0.110)	1.361* (0.207)	1.101* (0.052)
Research 1	1.060 (0.063)	0.914 (0.109)	1.175^{+} (0.109)	0.808^{+} (0.088)	1.008 (0.152)	0.971 (0.048)

Table 30 cont.

	Biology	**Chemistry**	**Computer Science**	**Math**	**Physics**	**Psychology**
Degree of Urbanization (Urban = 0)						
Suburb	0.951	0.977	0.933	1.046	1.054	1.031
	(0.042)	(0.075)	(0.066)	(0.087)	(0.114)	(0.043)
Town	0.953	0.907	1.171	1.055	0.867	0.945
	(0.049)	(0.094)	(0.115)	(0.103)	(0.126)	(0.044)
Rural	1.090	0.355**	0.680	0.878	1.010	0.945
	(0.163)	(0.141)	(0.252)	(0.243)	(0.523)	(0.131)
Female University President	1.057	1.087	1.047	0.999	0.923	0.942
	(0.048)	(0.086)	(0.071)	(0.080)	(0.094)	(0.034)
Proportion of Female Faculty	2.508**	1.430	1.742	12.144**	5.928*	2.452**
	(0.773)	(0.857)	(0.746)	(6.382)	(4.142)	(0.659)
Pseudo R-Squared	.049	.044	.126	.077	.050	.060

$^{+}$ *p < .10;* p<.05; ** p<.01; *** p<.001* (two-tailed tests)

Results: Organizational Structure and the Proportion of Female STEM Bachelor's Graduates

Tables 31 through 36 show the results examining the relationship between organizing processes – both at the department and college levels – and the proportion of female STEM bachelor's graduates. Like the models examining organizational culture, these are logistic regression models modeling the number of female STEM graduates over the total number of graduates. Recall in hypothesis 2.2 that I hypothesize programs housed in interdisciplinary departments and colleges outside of a college of science (or in the case of psychology, a college of liberal arts) will have higher proportions of female graduates than programs housed in single-disciplinary departments and colleges of science (or colleges of liberal arts for psychology programs).

Table 31 shows the results for biology. All other variables held constant, biology programs in a science, technology, engineering, and math department have a marginally higher proportion of female bachelor's graduates than programs in single-disciplinary biology departments (OR = 1.112, p = .092), providing partial support to my hypothesis that interdisciplinary departments yield higher proportions of female graduates than sole disciplinary departments (hypothesis 2.2). However, contrary to hypothesis 2.2, biology programs housed in a college of arts and science have significantly lower proportions of female graduates than biology programs in a college of science (OR = .873, p = .002). I find no other significant results by department or college level structure for biology.

For the controls in both the department and college level structure models, biology programs in the South have a significantly higher proportion of female graduates than in the West. Biology programs in private institutions also have a significantly higher proportion of female graduates than public institutions. Lastly, there is a significant and positive relationship between the proportion of female faculty institution-wide and the proportion of female bachelor's biology graduates.

Table 31: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Biology (Odds Ratios with Standard Errors in Parentheses)

	Department Structure	**College Structure**
Department Level Structure (biology department = 0)		
STEM department	1.112^{+} (0.070)	
College Level Structure (College of Science = 0)		
Science, Technology, Engineering, and Math		0.982 (0.056)
Arts and Sciences		0.873** (0.040)
Health and Science		0.917 (0.086)
Controls		
Total Student Population	1.000* (0.00000)	1.000** (0.00000)
Geographic region (West = 0)		
Northeast	1.094 (0.066)	1.107^{+} (0.066)
Midwest	0.954 (0.055)	0.977 (0.056)
South	1.155** (0.060)	1.171** (0.061)
Student/Faculty Ratio	1.005 (0.006)	1.003 (0.006)
Private	1.133* (0.058)	1.147** (0.059)
Research 1	1.040 (0.061)	1.089 (0.068)
Degree of Urbanization (Urban = 0)		
Suburb	0.969 (0.042)	0.960 (0.041)
Town	0.963 (0.049)	0.962 (0.049)

Table 31 Cont.

	Department Structure	College Structure
Rural	1.068 (0.159)	1.015 (0.153)
Female University President	1.088^{+} (0.048)	1.083^{+} (0.048)
Proportion of Female Faculty	2.154* (0.669)	2.488** (0.767)
Pseudo R-Squared	.045	.051

$^{+}$ *p < .10;* * *p<.05*; ** *p<.01;* *** *p<.001* (two-tailed tests)

I present the results for chemistry in table 32 below. For department level structure, chemistry programs in a science, technology, engineering, and math department have a significantly higher proportion of female bachelor's graduates than chemistry programs housed in a single-disciplinary chemistry department (OR = 1.361, *p* = .032), providing partial support for hypothesis 2.2. I find no significant differences in the proportion of female graduates between chemistry programs in a chemistry and biochemistry department and those in a single-disciplinary chemistry department. College level structure has no significant impact on the proportion of female graduates.

For the control variable results, chemistry programs in private institutions have significantly higher proportions of female graduates than chemistry programs in public institutions. Chemistry programs in rural institutions have a significantly lower proportion of female graduates than chemistry programs in urban institutions.

Table 32: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Chemistry (Odds Ratios with Standard Errors in Parentheses)

	Department Structure	**College Structure**
Department Level Structure (chemistry department = 0)		
STEM department	1.361*	
	(0.195)	
Chemistry and Biochemistry	0.940	
	(0.076)	
College Level Structure (College of Science = 0)		
Arts and Sciences		0.876
		(0.073)
Science, Technology, Engineering, and Math		1.002
		(0.097)
Controls		
Total Student Population	1.000	1.000
	(0.00000)	(0.00000)
Geographic region (West = 0)		
Northeast	0.845	0.889
	(0.090)	(0.094)
Midwest	0.821^{+}	0.856
	(0.083)	(0.086)
South	0.968	1.017
	(0.092)	(0.095)
Student/Faculty Ratio	1.010	1.008
	(0.011)	(0.011)
Private	1.365**	1.379**
	(0.140)	(0.142)
Research 1	0.890	0.965
	(0.103)	(0.114)
Degree of Urbanization (Urban = 0)		
Suburb	0.964	0.972
	(0.073)	(0.073)
Town	0.929	0.954
	(0.094)	(0.096)
Rural	0.365*	0.346**
	(0.145)	(0.138)
Female University President	1.045	1.034
	(0.080)	(0.079)
Proportion of Female Faculty	1.797	2.394
	(1.047)	(1.423)

Table 32 Cont.

	Department Structure	**College Structure**
Pseudo R-Squared	.040	.038

$^{+}$ *p < .10;* p<.05; ** p<.01; *** p<.001* (two-tailed tests)

The results for computer science are below in table 33. Regarding the impacts of department level structure, computer science programs in a mathematics and computer science department have a significantly higher proportion of female computer science bachelor's graduates than computer science programs in a single-disciplinary computer science department (OR = 1.384, p = .008). There are no significant differences in the proportion of female graduates between engineering and computer science or science departments and single-disciplinary computer science departments. There are no significant differences in the proportion of female bachelor's graduates between Colleges of Technology, Engineering, Arts and Science, or STEM and Colleges of Science in the proportion of female graduates.

Turning to controls, the models show significantly lower proportions of female computer science graduates in the Northeast, Midwest, and South compared to the West. There is also a significant and negative relationship between student/faculty ratio and the proportion of female computer science graduates; the higher the student/faculty ratio, the lower the proportion of female graduates (OR = .943, p < .01).

Table 33: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Computer Science (Odds Ratios with Standard Errors in Parentheses)

	Department Structure	**College Structure**
Department Level Structure (computer science department = 0)		
Engineering and Computer Science	1.081 (0.072)	
Math and Computer Science	1.384** (0.169)	
Science	0.824 (0.204)	
College Level Structure (College of Science = 0)		
Technology		0.936 (0.116)
Engineering		0.955 (0.104)
Arts and Science		1.049 (0.115)
STEM		0.940 (0.114)
Controls		
Total Student Population	1.000 (0.00000)	1.000 (0.00000)
Geographic region (West = 0)		
Northeast	0.791** (0.068)	0.759** (0.065)
Midwest	0.655** (0.055)	0.658** (0.057)
South	0.788** (0.064)	0.776** (0.063)
Student/Faculty Ratio	0.943** (0.010)	0.943** (0.010)
Private	1.129 (0.107)	1.126 (0.105)
Research 1	1.124 (0.101)	1.127 (0.103)
Degree of Urbanization (Urban = 0)		
Suburb	0.940 (0.067)	0.967 (0.069)
Town	1.170 (0.115)	1.177^{+} (0.115)

Table 33 Cont.

	Department Structure	**College Structure**
Rural	0.762	0.739
	(0.283)	(0.273)
Female University President	1.052	1.048
	(0.071)	(0.071)
Proportion of Female Faculty	2.071	1.779
	(0.926)	(0.780)
Pseudo R-Squared	.129	.124

$^{+}$ *p < .10;* * *p<.05*; ** *p<.01;* *** *p<.001* (two-tailed tests)

For mathematics programs (see table 34), neither department nor college level structures have a significant impact on the proportion of female bachelor's graduates. However, there are a few interesting control variable results. All other variables held constant, mathematics programs in the South have significantly higher proportions of female bachelor's graduates than programs in the West in both models. Mathematics programs housed at research 1 universities have lower proportions of female graduates than programs at non-Research 1 universities. Lastly, there is a highly significant and positive relationship between the proportion of female faculty institution-wide and the proportion of female bachelor's graduates in mathematics.

Table 34: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Mathematics (Odds Ratios with Standard Errors in Parentheses)

	Department Structure	**College Structure**
Department Level Structure (math department = 0)		
Mathematics and Statistics	0.926	
	(0.095)	
Math and Computer Science	0.966	
	(0.094)	
Science and Math	1.040	
	(0.226)	
College Level Structure (College of Science = 0)		
Science, Technology, Engineering, and Math		1.033
		(0.111)
Science and Math		1.052
		(0.131)
Arts and Sciences		1.079
		(0.088)
Controls		
Total Student Population	1.000	1.000
	(0.00000)	(0.00000)
Geographic region (West = 0)		
Northeast	0.953	0.956
	(0.100)	(0.102)
Midwest	0.991	0.977
	(0.108)	(0.110)
South	1.221*	1.219*
	(0.118)	(0.122)
Student/Faculty Ratio	1.015	1.017^{+}
	(0.010)	(0.010)
Private	1.161	1.178^{+}
	(0.111)	(0.111)
Research 1	0.796*	0.802*
	(0.088)	(0.088)
Degree of Urbanization (Urban = 0)		
Suburb	1.045	1.046
	(0.086)	(0.088)
Town	1.048	1.059
	(0.102)	(0.103)
Rural	0.876	0.901
	(0.245)	(0.258)

Table 34 Cont.

	Department Structure	College Structure
Female University President	1.006	0.992
	(0.081)	(0.080)
Proportion of Female Faculty	11.807**	11.591***
	(6.359)	(6.131)
Pseudo R-Squared	.077	.077

$^{+}$ *p < .10;* *p<.05; ** p<.01; *** p<.001* (two-tailed tests)

The results for physics programs are in table 29. Like mathematics programs, there are no significant differences in the proportion of female bachelor's graduates by any of the department or college level structures. Regarding the controls, private institutions have significantly higher proportions of female physics graduates than do public institutions. There is also a significant and positive relationship between the proportion of female faculty institution-wide and the proportion of female physics bachelor's graduates, net of other controls.

Table 35: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Physics (Odds Ratios with Standard Errors in Parentheses)

Physics	**Department Structure**	**College Structure**
Department Level Structure (physics department = 0)		
Physics and Astronomy	1.031	
	(0.100)	
Science, Technology, Engineering, and Math	1.085	
	(0.194)	
College Level Structure (College of Science = 0)		
Science, Technology, Engineering, and Math		1.040
		(0.154)
Arts and Sciences		1.051
		(0.108)
Controls		
Total Student Population	1.000	1.000
	(0.00001)	(0.00001)
Geographic region (West = 0)		
Northeast	0.948	0.951
	(0.124)	(0.125)
Midwest	0.984	0.981
	(0.118)	(0.118)
South	1.126	1.138
	(0.137)	(0.137)
Student/Faculty Ratio	1.001	1.001
	(0.016)	(0.016)
Private	1.356*	1.339^{+}
	(0.204)	(0.203)
Research 1	1.006	1.002
	(0.151)	(0.155)
Degree of Urbanization (Urban = 0)		
Suburb	1.037	1.021
	(0.112)	(0.112)
Town	0.864	0.860
	(0.124)	(0.124)
Rural	0.865	0.873
	(0.439)	(0.443)
Female University President	0.921	0.918
	(0.094)	(0.094)
Proportion of Female Faculty	6.052**	5.504*
	(4.204)	(4.001)

Table 35 Cont.

	Department Structure	College Structure
Pseudo R-Squared	.042	.042

+ $p < .10$;* $p<.05$; ** $p<.01$; *** $p<.001$ (two-tailed tests)

The last set of results for Study 3 are in table 36, which are for psychology net of the control variables. For department level structure, there is a significantly higher proportion of female psychology bachelor's graduates in psychology programs housed in departments of behavioral science and services than programs in single-disciplinary psychology departments (OR = 1.668, p = .0002). There are no significant differences in the proportion of female graduates across any of the college level structures. For the controls, there are significantly higher proportions of female psychology bachelor's graduates in the Northeast, Midwest, and South than in the West. There is a significantly higher proportion of female psychology students at private universities compared to public universities. Lastly, there is a significant and positive relationship between the proportion of female faculty institution-wide and psychology graduates.

Table 36: Logistic Regression Models Examining Department and College Level Structures' Impact on the Proportion of Female Bachelor's Graduates in Psychology (Odds Ratios with Standard Errors in Parentheses)

	Department Structure	**College Structure**
Department Level Structure (psychology department = 0)		
Behavioral Science and Services	1.668**	
	(0.229)	
Social and Behavioral Science	0.960	
	(0.061)	
College Level Structure (College of Liberal Arts = 0)		
Arts and Sciences		0.951
		(0.045)
Science and Health		1.021
		(0.067)
Social Sciences		0.937
		(0.055)
Professional Studies		0.867
		(0.089)
Controls		
Total Student Population	1.000	1.000
	(0.00000)	(0.00000)
Geographic region (West = 0)		
Northeast	1.183**	1.187**
	(0.064)	(0.066)
Midwest	1.200**	1.180**
	(0.068)	(0.068)
South	1.259**	1.253**
	(0.054)	(0.056)
Student/Faculty Ratio	1.000	1.000
	(0.005)	(0.005)
Private	1.090^{+}	1.109*
	(0.051)	(0.053)
Research 1	0.997	0.990
	(0.048)	(0.048)
Degree of Urbanization (Urban = 0)		
Suburb	1.024	1.048
	(0.043)	(0.044)
Town	0.938	0.941
	(0.043)	(0.044)
Rural	0.973	0.965
	(0.134)	(0.134)

Table 36 Cont.

	Department Structure	College Structure
Female University President	0.944	0.935^{+}
	(0.034)	(0.034)
Proportion of Female Faculty	2.688**	2.560**
	(0.726)	(0.686)
Pseudo R-Squared	.066	.058

$^{+}$ *p < .10;* * *p<.05;* ** *p<.01;* *** *p<.001* (two-tailed tests)

CHAPTER 7: DISCUSSION, IMPLICATIONS, AND CONCLUSION

Discussion

In this dissertation, I descriptively examined and tested components of Acker's theory of gendered organizations, specifically organizational culture and organizing processes, applied to six STEM disciplines (biology, chemistry, computer science, mathematics, physics, and psychology). I used text data, specifically each STEM program's "about us" or equivalent introductory webpage, as a way of measuring organizational culture using Acker's idea of gendered subtext. Gendered subtexts consist of written or common practice texts that shape gendered processes and structures in organizations (Acker 1990; Acker 2012). I used text data because they were easy to obtain, allowed me to have a large sample of programs, and since other studies have examined culture and STEM using text data (e.g., De Pillis and De Pillis 2008). From the literature on gender and STEM, I derived four concepts to measure organizational culture in STEM programs: *collaboration across disciplines* (feminine), *socially-connected science* (feminine), *socially-disconnected science* (masculine), and *expectations of brilliance and competition* (masculine).

I conducted three quantitative studies to explore the following research questions:

1) To what extent are the cultures of STEM disciplines gendered? (study 1)
2) To what extent do the organizational cultures in programs differ across STEM fields? (study 2)
3) How is organizational culture related to the proportion of female bachelor's graduates in STEM? (study 3)
4) How are the departmental and institutional structures of STEM programs (organizing processes) related to the proportion of female bachelor's graduates? (study 3)

Study 1 was a descriptive social network analysis in which I explored the mental models of the six disciplines to examine the extent to which the cultures of STEM disciplines are gendered (research question 1). In Study 2, I derived different measurements of organizational culture (both categorical and continuous) using machine learning algorithms and used multinomial logistic and multivariate regression models to examine any cultural differences between the six STEM disciplines to address research question 2. In Study 3, I used binomial logistic regression to examine if a STEM program's culture, department-level structure, and college-level structure were related to the proportion of female bachelor's graduates in each of the six STEM disciplines to address research questions 3 and 4.

Summary and Possible Explanations of Results in Studies 1, 2, and 3

Study 1 showed that the mental models of each STEM discipline were similar in that *socially connected science*, a feminine organizational cultural concept, was most central for biology, chemistry, computer science, mathematics, and psychology, and the second most central concept for physics (as shown in the two-mode network). The one-mode network of words showed that every discipline but chemistry had the word "skills," a word from the *socially-connected science* concept, as the most central word in its mental model. Chemistry also had its most central word from *socially-connected science* ("society"). The differences in the mental models started to emerge when I examined the second most central words and beyond. Even though physics had several "feminine" words for its 2nd to 5th most central words in the one-mode network, computer science had several masculine words in its mental model, such as "theory", "technical", and "top." For the two-mode network, even though computer science had the highest degree centrality scores for three out of the four organizational cultural concepts (*collaboration across disciplines*, *socially-disconnected science*, and *expectations of* *brilliance/competition*), psychology had the highest degree centrality scores for *socially-connected science*.

While all of the disciplines had very similar mental models regarding the feminine organizational cultural concepts of *collaboration across disciplines* and *socially-connected science*, the disciplines with the highest representation of women (biology and psychology) differed the most from the most male-dominated disciplines (computer science and physics) because of biology and psychology's lower centrality scores and rankings of the masculine organizational cultural concepts (*socially-disconnected science* and *expectations of brilliance*). Put differently, these mental models show that a lack of feminine organizational culture, at least how it is defined in this study, is not necessarily driving the masculine organizational culture of computer science relative to biology, psychology, and chemistry. Rather, the much higher presence of masculine concepts and words (especially from *expectations of brilliance*) is driving the organizational cultural differences between these fields. These mental models support prior research that found disciplines with higher *expectations of brilliance* have lower proportions of female PhD graduates (Leslie et al. 2015; Storage et al. 2016).

Analyses in study 2 found partial support for my hypothesis that STEM disciplines with higher proportions of women (e.g., biology and psychology) would have more feminine cultures and less masculine cultures than disciplines with lower proportions of women (e.g., computer science and physics; see hypothesis 1). When I measured organizational culture categorically as feminine, masculine, or gender neutral, biology and psychology were three to five times as likely as computer science to be classified as either feminine or gender neutral relative to masculine. Physics and chemistry were about twice as likely as computer science to be classified as gender neutral instead of masculine. Interestingly, despite the large gap between the proportion of female graduates in math versus computer science, there were no statistically significant differences between math and computer science in terms of their organizational culture. These findings show that the categorical measure for organizational culture is able to distinguish the cultures of most of the STEM disciplines, but was not able to account for organizational cultural differences between mathematics, a field that has achieved gender parity in representation at the undergraduate level, and computer science, a male-dominated discipline.

The multivariate regression results of the four continuous measures of organizational culture (*collaboration across disciplines*, *socially-connected science*, *socially-disconnected science*, and *expectations of brilliance and competition*) were consistent with the results that used the categorical measure of organizational culture. While chemistry had a significantly lower score for *collaboration across disciplines* than computer science (a male-dominated field), biology, chemistry, and psychology (three fields with gender parity in representation) had significantly lower scores than computer science for the two masculine organizational cultural concepts (*socially-disconnected science* and *expectation of brilliance and competition*).

Overall, the findings in Study 3 provided mixed support for hypotheses 2.1 and 2.2. The results using the categorical measures of feminine, masculine, or gender neutral organizational culture were mostly not statistically significant. However, I found that computer science programs classified as gender neutral had significantly lower proportions of female graduates compared to programs classified as masculine. For the continuous measures of organizational culture (*collaboration across disciplines*, *socially-connected science*, *socially-disconnected science*, and *expectations of brilliance*), I found that chemistry programs with higher scores on *expectations of brilliance* had significantly lower proportions of female bachelor's graduates, providing support to hypothesis 2.1 that STEM programs with more feminine (or less masculine) cultures will have higher proportions of female bachelor's graduates. It is possible that the categorical measures of organizational culture were largely not significantly related to the proportion of female bachelor's graduates because you lose information in the organizational culture variable when it is treated as categorical instead of continuous, resulting in lower statistical power and the ability to detect a statistically significant result. This would also explain why the continuous measure of organizational culture (*expectations of brilliance*) was significantly related to the proportion of female graduates in chemistry programs.

However, several results in Study 3 were inconsistent with what I hypothesized; in biology programs, I found a negative relationship between scores on *socially-connected science* and the proportion of female graduates. Likewise, in chemistry programs, I found a negative relationship between scores on *collaboration across disciplines* and the proportion of female graduates. In computer science, I found a positive relationship between scores for *socially-disconnected science* and the proportion of female graduates. It is not clear why these relationships between organizational culture and the proportion of female graduates were the opposite of what I hypothesized (hypothesis 2.1). One possible explanation for these findings is that since the way I measured organizational culture is based on program texts, it is possible that programs with lower proportions of female graduates wrote their "about us" webpages in a way that tried to attract more female students than programs that already had higher proportions of female graduates. This phenomenon of overcompensating for a lack of diversity in STEM programs is similar to other settings outside of STEM, which have found that a majority of university homepages or college admissions viewbooks use photos that over-represent the proportion of minorities at the institution (Wilson and Meyer 2009; Matchett and Pipperet 2008).

For the results examining the department and college-level structures (organizing processes) of each STEM discipline's programs, tests of hypothesis 2.2, I first found that biology programs housed in the College of Arts and Sciences had significantly lower proportions of female graduates than programs housed in a College of Science, which was the opposite of what I hypothesized (hypothesis 2.2). However, I found several results that were in line with my hypothesis that programs housed in interdisciplinary departments or colleges would have greater proportions of female graduates than programs in single-disciplinary departments or colleges of science[17] (hypothesis 2.2). I found that chemistry programs in Science, Technology, Engineering, and Math departments had significantly higher proportions of female graduates than chemistry programs housed in departments of chemistry. Similarly, I found that computer science programs housed in departments of mathematics and computer science had significantly higher proportions of female graduates than programs housed in departments of computer science only. I also found that psychology programs housed in departments of behavioral sciences and services had higher proportions of female graduates than programs in departments of psychology. Neither department nor college-level structures were significantly associated with the proportion of female bachelor's graduates in mathematics or physics.

The finding that computer science programs housed in departments of mathematics and computer science had significantly higher proportions of female graduates compared to programs housed in a department with computer science only is especially telling because this finding shows that the underrepresentation of women in certain math-intensive sciences (e.g., computer science and engineering) is more complicated than simply that women are less confident in their

[17] In the case of psychology programs, the reference category was the College of Liberal Arts since very few programs were housed in a College of Science.

mathematics abilities (Cheryan et al. 2017). However, it is unclear why there is a larger proportion of female graduates in mathematics and computer science departments than sole computer science departments. It is possible that it is because departments of mathematics and computer science are interdisciplinary and departments of computer science are single-disciplinary since research has found that women are more interested in interdisciplinary research than single-disciplinary research (Rhoten and Pfirman 2007; van Rijnsoever and Hessels 2011; Smith-Doerr 2005), but it is unclear what mechanisms in interdisciplinary departments drive the higher proportion of female graduates in these departments versus single-disciplinary computer science departments. Survey research could be used in future studies to examine how students and faculty perceive interdisciplinary versus single-disciplinary environments, and if interdisciplinary environments are perceived as more inclusive to women than single-disciplinary environments.

Although this research contributes new knowledge about how different STEM disciplines are gendered, these three studies raise an additional question: why were there no statistically significant differences in the organizational cultures of mathematics and computer science, despite the fact that mathematics has over twice the proportion of female graduates as computer science (National Science Foundation 2016)? I have several possible explanations for this unusual finding. One is that since I only used program descriptions from "about us" webpages, studies 1, 2, and 3 may have left out sources of organizational culture (e.g., grading on a curve) that would further distinguish the organizational culture of mathematics programs from computer science programs had I been able to include other texts (e.g., course syllabi). Another is that mathematics and computer science programs share many of the same courses in the beginning years, and since I did not look at course syllabi or other course texts, the program "about us" and other introductory webpages I did collect would not necessarily reflect this.

I also could not take into account men's choices alongside women's choices to major or remain in STEM in this dissertation, which could be one reason behind women's greater proportion of graduates in mathematics versus computer science (Cheryan et al. 2017). In 2010, about the same number of women intended to major in computer science as mathematics, but the number for men intending to major in computer science was much higher than it was for mathematics (National Science Foundation 2012), producing the disparity between these disciplines. Men's choices, as well as women's, could be a factor behind the cultural difference between mathematics and computer science, and this dissertation was not able to take this factor into account since I only looked at the organizational level.

Comparison of Results in Studies 1-3

Even though I examined different aspects of organizational culture and organizing processes in these three studies, the results complemented each other and were relatively consistent. In Study 1, I found that the mental models of the six disciplines centered around *socially-connected science*, a feminine organizational cultural concept, but that the mental models differed when it came to the centrality of the other concepts and words representing these concepts. Specifically, the differences in the mental models between computer science and the STEM disciplines with higher representations of women (biology and psychology) in Study 1, especially the much higher centrality of *expectations of brilliance* in computer science, provide nuance behind the results found in Study 2. That is, biology, chemistry, and psychology have significantly lower scores on the two masculine concepts than computer science, despite the fact that not all of these fields have significantly higher scores on the two feminine concepts.

It is not clear how organizational culture and the proportion of female bachelor's graduates are linked. While studies 1 and 2 of this dissertation show clear differences in organizational culture between the male-dominated fields and the fields with at least gender parity in representation, especially among computer science versus biology and psychology, Study 3 did not show a relationship between organizational culture and the proportion of female graduates. However, chemistry, computer science, and psychology programs had select interdisciplinary departmental structures (organizing processes) with higher proportions of female graduates than programs in single-disciplinary departments, which were departments of STEM, mathematics and computer science, and behavioral science and services respectively.

Studies 1, 2, and 3 of this dissertation contribute to the interdisciplinary literature examining gender and representation in STEM. Studies 1 and 2 were able to "tease out" potential disciplinary differences in organizational culture, and study 3 showed the relationships between organizational culture, departmental structure, college-level structure, and the proportion of female graduates. By doing so, the research here addresses the current lack of STEM education research that disaggregates disciplines (Cheryan et al. 2017). Based on the three studies in this dissertation, the largest differences in organizational culture, as well as the most significant relationships between organizational culture or organizing processes and the proportion of female bachelor's graduates, emerged in the most male and female-dominated disciplines of computer science, biology, and psychology. This is telling since it supports the notion that STEM disciplines should be studied separately not only because they are not created equally in terms of women's representation, but because there are differences in organizational culture among them, especially among STEM disciplines that are male-dominated or have at least gender parity in representation.

Limitations

Despite this dissertation's contribution to the literature on gender and STEM, these analyses have several limitations. First, it is possible that the way I measured organizational culture, whether categorical or through a series of continuous variables, does not represent if a program actually has feminine, masculine, or gender neutral organizational culture. In other words, my measurements of organizational culture, which form the basis of these studies, have possible validity problems. The validity of the organizational culture measures is questionable because the words that were used to define each categorical or continuous measure were taken out of their original context in each program's text. It is possible that a text could have a word that was deemed as feminine or masculine, but have the opposite meaning when the words is interpreted in context of the document than when the word is by itself. For example, a text could have said something like "our program's goal is not to be competitive, but collaborative," which has a masculine word ("competitive"), but the meaning of the word in context of the sentence is not masculine since the sentence is saying the program is *not* competitive, but collaborative. Future studies could attempt to fix these validity issues by having human coders initially classify the training and testing text data as masculine, feminine, or gender neutral in order to pick up on these nuances.

Second, since I only looked at program introductory websites (e.g., "about us") webpages, I was not able to include other aspects of a program's organizational culture that may have masculine or feminine characteristics, such as course syllabi, faculty descriptions, and departmental handbooks, due to low availability of these texts on the STEM program websites. It is possible that the results of this study could have been different if I was able to include additional texts (e.g., course syllabi). For example, I was not able to include language from “weed out” courses (e.g., grading on a curve) or an emphasis on collaboration or independence when completing assignments. As a result, there is a potential mismatch between the text data I examined (program websites) and texts used in prior qualitative research (e.g., admissions materials, course syllabi, and teaching evaluations) since I was not able to examine other aspects of organizational culture that could have played a role and improved the validity of these studies. Future researchers could attempt to build a database of a larger variety of texts (e.g., course syllabi and admissions materials) by contacting STEM programs, or carry out research on these texts using case studies of a smaller group of institutions.

Third, the gender-neutral category of the categorical measure of organizational culture is also limited. I defined gender neutral culture as the complete absence of feminine or masculine culture (there were no words labeled as “gender neutral” in the lexicon), rather than creating a separate gender neutral category like I did with feminine or masculine culture. Regardless, it is difficult to tell what gender neutral culture specifically consists of in STEM programs. A few experiments involving classroom environments and interest in computer science have a condition that represents a gender-neutral environment (e.g., Cheryan et al. 2009; Master et al. 2016), but there does not appear to be a similar precedent to what gender-neutral culture consists of that could be measured with text data. Survey data or interviews on what students and faculty perceive as being “gender neutral” in STEM could help establish what gender neutral means in this context.

Fourth, I only used cross-sectional data in this study, presenting additional limitations. It is possible that the proportion of female graduates was highly variant from year to year, especially for smaller programs where a woman graduating has a lot of weight in the proportion of female graduates. Even though I had access to longitudinal data of the proportion of female graduates and the institutional data for the control variables, I only had access to the program text data from one time point since I pulled them from each program's most recent website, making it impossible to examine how STEM organizational culture is related to the proportion of female graduates over time. Relatedly, there are problems with causality considering that I cannot assess trends over time or provide explanations for the relationship between organizational culture or organizing processes and the proportion of female graduates because one year's worth of data only provides a small picture. Examining this relationship over time would show if there is a consistent relationship between organizational culture, organizing processes, and the proportion of female graduates. If these data were available, it would also be possible to see if STEM programs have changed their organizational culture or organizing processes over time. Even though programs housed in certain interdisciplinary departments for chemistry, computer science, and psychology had a significantly higher proportion of female graduates than programs housed in single-disciplinary departments, the mechanisms behind this statistical relationship remain uncertain due to a lack of data from STEM programs (e.g., program climate surveys of students and faculty).

Fifth, while I considered how organizational culture and organizing processes are related to the proportion of female STEM graduates, something that has not been considered in prior research, this organizational perspective has limits. Since I only looked at organization level data, I do not consider individual perceptions of the STEM organizational culture, or if the content of STEM program texts are related to individual decisions to major in STEM or choose a certain university for a STEM major. This organizational perspective fails to consider the agency of students working within and experiencing STEM programs. It is possible that examining student and faculty perceptions of the organizational culture and organizing processes present in STEM texts could provide an additional validity check at the very least, but would also be a more comprehensive way of using Acker's theory of gendered organizations. By neglecting the individual level, I was not able to consider the other half of Acker's idea of the gendered substructure, which contained the concepts of gendered identities and organizational interactions. Applied to STEM, these studies left out other ways that gender is a part of STEM programs, such as through identity construction and negotiation, aspirations, role models, and social capital. Because of these omissions, it is possible that there are larger organizational cultural differences between STEM disciplines or that organizational culture has a stronger association with the proportion of female graduates.

Lastly, although I disaggregated STEM disciplines by looking at the specific programs for biology, chemistry, computer science, mathematics, physics, and psychology, I neglected to look at subfields within these programs. Subfields are important to examine because several studies have found gender differences in representation within subfields of STEM disciplines, such as engineering (Brawner et al. 2012; Morton 2016) and computer science (Dray et al. 2013; Zweben and Bizot 2016). For example, the computer science subfield of human-computer interaction has reached gender parity in the United States (Zweben and Bizot 2016), so just studying computer science as a whole does not capture the entire picture of the underrepresentation of women in computer science or other STEM disciplines. Researchers should gather similar text data at the subfield level in future studies, and could break out the proportion of female graduates at the subfield level using CIP codes in the IPEDS Completions Survey data.

Contributions of this Dissertation

Despite these limitations, this dissertation makes several contributions to the sociological literature, as well as to the interdisciplinary literature on gender and STEM. One contribution is that I tested hypotheses based on concepts from Acker's theory of gendered organizations, specifically from organizational culture and organizing processes, which prior research has neglected. Studies 1 and 2 of this dissertation showed that there are organizational cultural differences between STEM disciplines, especially between disciplines with higher proportions of female graduates (biology and psychology) and disciplines with lower proportions of female graduates (computer science), but that there were nuances to these differences in organizational culture. While computer science had feminine words and concepts with high centrality in its mental model in a similar way to biology and psychology, the distinguishing factor between computer science and these disciplines was that computer science also had masculine words and concepts, especially *expectations of brilliance and competition*, as highly central in its mental model.

Another contribution follows from Study 3; even though some of the results in this study (regarding the relationship between organizational culture and the proportion of female graduates) were the opposite of what I hypothesized (for example, biology programs had a significant and negative relationship between scores on *socially-connected science* and the proportion of female graduates), these inconsistent findings could be used to improve Acker's theory of gendered organizations. In light of these inconsistent findings, it is possible that Acker's theory of gendered organizations needs to be modified to distinguish the role of organizational culture and gendered subtext in publicly facing texts (e.g., websites) versus internal organizational texts. It is possible that publicly facing organizational texts, such as STEM program "about us" webpages, attempt to appear to have a more female-friendly culture in order to attract more female students, but Acker's theory does not currently take into account any differences between use of texts.

A different contribution from Study 3 is that I find more evidence linking organizing processes at the department level (e.g., the program being in an interdisciplinary versus single-disciplinary department) than at the college level (e.g., the program being in a College of Science versus a more interdisciplinary college). This shows that the organizing processes from Acker's theory of organizations function differently at different levels of organizations. Since I found more associations between organizing processes at the department level than the college level, it is possible that organizing processes have a greater impact on women's representation at levels closest to the individual. At the very least, my work calls for future researchers to explore how organizing processes might impact women's inequality at different levels of organizations, which would further add to Acker's theory.

Another contribution this dissertation makes is that it disaggregated STEM disciplines. By comparing six STEM disciplines, I was able to show that there are organizational cultural differences between STEM disciplines, and that organizational culture and organizing processes are related to the proportion of female graduates in some disciplines (e.g., psychology and computer science) over others (e.g., mathematics and physics).

Conclusion

To conclude, we gain much knowledge from this dissertation's findings. First, we learn, at least in part, the extent to which the organizational cultures of six STEM disciplines are gendered (research question 1). Most of the disciplines, including the male-dominated field of computer science, have feminine words and concepts as most central in their mental models. However, computer science and physics also have masculine words and concepts that are very central in their mental models (especially related to *expectations of brilliance*), which provide a clear distinction between the organizational cultures of male-dominated STEM disciplines and STEM fields with at least gender parity.

Second, we learn that organizational cultures do differ between programs across STEM fields (research question 2), especially between the fields with the highest proportions of female graduates (biology and psychology) and fields with the lowest proportions of female graduates (computer science). For example, biology and psychology are three to five times as likely as computer science to be classified as either feminine or gender neutral relative to masculine. Both chemistry and physics are more likely than computer science to be classified as gender neutral relative to masculine. In light of these findings, there are no significant differences in the organizational cultures of mathematics and computer science.

Lastly, this dissertation presents findings about how organizational culture is related to the proportion of female bachelor's graduates (research question 3) and how the organizing processes of departmental and college-level structures of STEM programs are related to the proportion of female bachelor's graduates (research question 4). For the most part, I do not find any relationships between organizational culture and the proportion of female graduates. However, in some of the fields (e.g., biology) I found a negative relationship between feminine organizational cultural concepts and the proportion of female graduates, or a positive relationship between masculine organizational cultural concepts and the proportion of female graduates, which was the opposite of what I hypothesized (hypothesis 2.1). We also learn that in chemistry, psychology, and computer science programs, certain interdisciplinary department-level structures (e.g., a department of mathematics and computer science) have significantly higher proportions of female graduates than programs in single-disciplinary departments (e.g., department of computer science). At the college level, I do not find any significant differences in the proportion of female graduates between Colleges of Science and other interdisciplinary colleges, which give us knowledge of the greater importance of the department level than the college level as a factor that positively impacts the proportion of female bachelor's graduates.

Future Research

Findings from the three studies presented in this dissertation set up several avenues for future research. First, future research could examine whether organizational cultural concepts found in STEM program texts are related to individual decisions to major in STEM or to choose specific STEM programs. It has been found in at least one study that university websites that emphasize age diversity increase perceived person-organization fit across age (Ihme et al. 2016), so future research should build upon this finding applied to diversity in STEM fields. At the very least, conducting surveys or qualitative interviews of incoming and current student's reactions to STEM program texts (e.g., admissions materials, program "about us" webpages, and course syllabi), especially with regards to their organizational culture, could provide an additional validity check to future quantitative studies involving STEM program texts.

Other individual and organizational level variables should be considered. After all, gendered organizations contain more than organizational structure and organizing process, but also interactions between people in organizations and gendered identities that people bring into organizations. Future research could compare mental models, or internal representations of reality (in this case, the organizational culture of STEM programs) that consist of a network of associations between concepts, of individuals deciding to or currently majoring in STEM to mental models found at the program level, such as those from STEM faculty and administration. These mental models could be derived from texts collected from interviews or focus groups of students, faculty, and administration in STEM programs. Similarly, individual-level mental models of male and female STEM students from different disciplines could also be examined to see how students perceive their STEM program's culture, and if these perceptions differ by gender or STEM discipline. Future research could also build on my consideration of organizing processes by exploring how students perceive programs housed in interdisciplinary departments versus single-disciplinary departments, as well as men's and women's sense of belonging in STEM in these different contexts.

To fill a gap in prior research and in the broader theory of gendered organizations that I was not able to address in these three studies, future research should now consider how the choices of men are related to the proportion of female STEM students. Men's choices matter just as much as women's choices in shaping women's underrepresentation in STEM disciplines (Miller, Taylor, and Buck 1991; Cheryan et al. 2017), so this is an important idea to consider in future research. These studies could specifically look at the individual mechanisms that drive large numbers of men to certain disciplines (e.g., computer science), and compare these with disciplines that have roughly equal proportions of men and women, but less men and people overall (e.g., mathematics). Studies of this nature could help answer why the organizational culture did not differ between computer science and mathematics programs, despite the fact that computer science is male-dominated at the undergraduate level, and mathematics has reached gender parity.

Lastly, future studies should consider other outcomes in STEM besides the numerical representation of female bachelor's graduates. This is especially important to consider when studying female-dominated STEM disciplines such as biology; one study found that although women made up about 60 percent of the course-takers in biology, they represented less than 40 percent of those heard responding to questions posed by instructors in class, showing that inequities still exist for female students in female-dominated STEM fields (Eddy et al. 2014). Research that addresses the link between organizational structure, organizing processes, and a variety of STEM outcomes beyond the proportion of female graduates will inform researchers and the public on ways to address the needs for individual STEM disciplines, whether they have gender parity in representation or not.

Practical Implications

In addition to implications for future research, research findings in this dissertation pose several practical implications. These three studies could help STEM programs be more aware of how the language used in their departments, whether it is written or spoken, might contribute to the inclusivity of the program. Since I found that the higher presence of masculine concepts – especially expectations of brilliance – differentiated the mental models of some of the male-dominated STEM disciplines from the disciplines with at least gender balance, STEM programs and departments could use this information to create a more inclusive culture for women and minority students and faculty. Likewise, since most of the significant relationships I found between organizational culture and the proportion of female graduates were the opposite of what I hypothesized, STEM programs and departments could also examine if they are "window-dressing" their websites or other texts in order to give the impression that their departments are female-friendly, and instead make more effective changes to their overall organizational culture accordingly.

Several implications emerge from my findings on departmental and college level structure (organizing processes) and their relationships with the proportion of female graduates in STEM. This study brings light to the importance of interdisciplinary research on improving the representation of women in STEM, especially at the department level. At the highest level of change, departments could restructure themselves so that they are interdisciplinary instead of single-disciplinary to encourage interdisciplinary work and collaboration among students in order to provide a more inclusive environment. While completely restructuring departments to be interdisciplinary may be unrealistic, STEM programs could at least use the information from this dissertation to justify giving their program coursework, projects, and collaborations with more interdisciplinary components. While interdisciplinary colleges versus colleges of science were not related to the proportion of female STEM graduates, colleges in academic institutions could use this information to justify efforts at making STEM programs more interdisciplinary at the department level.

Ultimately, the implementation of these suggestions will help produce more inclusive STEM disciplines. While organizational culture and organizing processes are hard to change, this dissertation presents research and findings that could help change the way gender and STEM research is completed, which will in turn provide a larger body of information to help change the inclusivity of STEM fields as a whole, whether it is in academic institutions or in the public and private sectors. Disaggregating STEM fields at the program level provides a great start to learning how and why STEM disciplines are not created equal in terms of women's representation so that more effective policies and practices can be utilized to "undo" gender in these fields.

REFERENCES

Acker, Joan, & Van Houten, D. R. 1974. "Differential Recruitment and Control: The Sex Structuring of Organizations." *Administrative Science Quarterly* 19(2): 152-163.

Acker, Joan. 1990. "Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations." *Gender & Society* 4(2): 139-158.

Acker, Joan. 1992. "From Sex Roles to Gendered Institutions." *Contemporary Sociology* 21(5): 565-569.

Acker, Joan. 2006. "Inequality Regimes: Gender, Class, and Race in Organizations." *Gender & Society* 20(4): 441-464.

Acker, Joan. 2012. "Gendered Organizations and Intersectionality: Problems and Possibilities." *Equality, Diversity and Inclusion: An International Journal* 31(3): 214-224.

Allmendinger, J., and Hackman, J. R. 1995. "The More, the Better? A Four-Nation Study of the Inclusion of Women in Symphony Orchestras." *Social Forces* 74(2): 423-460.

Attewell, P., & Monaghan, D. 2015. *Data mining for the social sciences: An introduction*. University of California Press.

Augustyniak, Lukasz, et al. 2014. "Simpler Is Better? Lexicon-Based Ensemble Sentiment Classification Beats Supervised Methods." *Advances in Social Networks Analysis and Mining* (ASONAM), 2014 IEEE/ACM International Conference on. IEEE.

Bain, L. J., and Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics*. Brooks/Cole.

Baram-Tsabari, Ayelet, and Anat Yarden. 2008. "Girls' Biology, Boys' Physics: Evidence from Free-Choice Science Learning Settings." *Research in Science & Technological Education* 26(1): 75-92.

Barnard, C. I. 1968. *The Functions of the Executive*. Cambridge: Harvard University Press.

Bastian, M., et al. 2009. “Gephi: An Open Source Software for Exploring and Manipulating Networks.” *ICWSM* 8:361-362.

Bear, J. B., and Woolley, A. W. 2011. “The Role of Gender in Team Collaboration and Performance.” *Interdisciplinary Science Reviews* 36(2): 146-153.

Bejerano, Arleen R., and Travis M. Bartosh. 2015. “Learning Masculinity: Unmasking the Hidden Curriculum in Science, Technology, Engineering, and Mathematics Courses.” *Journal of Women and Minorities in Science and Engineering* 21(2): 107-124.

Bettinger, Eric P., and Bridget Terry Long. 2005. “Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students.” *The American Economic Review* 95(2): 152-157.

Bird, S. R. 2011. “Unsettling Universities’ Incongruous, Gendered Bureaucratic Structures: A Case-Study Approach.” *Gender, Work & Organization* 18(2): 202-230.

Blauner, R. 1964. *Alienation and Freedom: The Factory Worker and His Industry*. Oxford, England: Chicago University Press.

Borgatti, S. P., and Everett, M. G. 1997. “Network Analysis of 2-Mode Data.” *Social Networks* 19(3):243-269.

Borgatti, S.P., et al. 2002. UCINET for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Boyd, R. L., and Pennebaker, J. W. (2015). “Did Shakespeare Write Double Falsehood? Identifying Individuals by Creating Psychological Signatures with Text Analysis.” *Psychological Science* 26(5):570-582.

Brawner, Catherine E., et al. 2012. "Women in Industrial Engineering: Stereotypes, Persistence, and Perspectives." *Journal of Engineering Education* 101(2): 288-318.

Britton, D. M., & Logan, L. 2008. "Gendered Organizations: Progress and Prospects." *Sociology Compass* 2(1):107-121.

Budig, Michelle J. 2002. "Male Advantage and the Gender Composition of Jobs: Who Rides the Glass Escalator." *Social Problems* 49(2): 258-277.

Burger, Carol. 2009. "Factors that Impact Persistence of Female Students in Undergraduate Engineering." *American Society for Engineering Education*.

Cain, Cindy L., and Erin Leahey. 2014. "Cultural Correlates of Gender Integration in Science." *Gender, Work & Organization* 21(6): 516-530.

Canney, N., and Bielefeldt, A. 2015. "A Framework for the Development of Social Responsibility in Engineers." *International Journal of Engineering Education* 31(1B): 414-424.

Carley, K., and Palmquist, M. 1992. "Extracting, Representing, and Analyzing Mental Models." *Social Forces* 70(3): 601-636.

Carley, K. M. 1997. "Extracting Team Mental Models through Textual Analysis." *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 18(S1): 533-558.

Carrigan, C., et al. 2011. "The Gendered Division of Labor Among STEM Faculty and the Effects of Critical Mass." *Journal of Diversity in Higher Education* 4(3): 131-146.

Cech, Erin A., and Mary Blair-Loy. 2010. "Perceiving Glass Ceilings? Meritocratic Versus Structural Explanations of Gender Inequality among Women in Science and Technology." *Social Problems* 57(3): 371-397.

Cech, Erin, et al. 2011. “Professional Role Confidence and Gendered Persistence in Engineering.” *American Sociological Review* 76(5): 641-666.

Cech, Erin. 2013. “The Self-Expressive Edge of Occupational Sex Segregation.” *American Journal of Sociology*, 119(3): 747-789.

Cech, Erin. 2014. “Culture of Disengagement in Engineering Education?” *Science, Technology, & Human Values* 39(1): 42-72.

Cech, E. (2015). “Engineers and Engineeresses? Self-Conceptions and the Development of Gendered Professional Identities.” *Sociological Perspectives* 58(1):56-77.

Chamberlain, L. J., et al. 2008. “Sexual Harassment in Organizational Context.” *Work and Occupations* 35(3): 262-295.

Chanderbhan-Forde, et al. 2012. “The Doors are Open but They Don’t Come in: Cultural Capital and the Pathway to Engineering Degrees for Women.” *Journal of Women and Minorities in Science and Engineering* 18(2): 179-198.

Cheryan, S., et al. (2009). “Ambient Belonging: How Stereotypical Cues Impact Gender Participation in Computer Science. *Journal of Personality and Social Psychology* 97(6): 1045-1060.

Cheryan, S., et al. 2013. “The Stereotypical Computer Scientist: Gendered Media Representations as a Barrier to Inclusion for Women.” *Sex Roles* 69(1-2): 58-71.

Cheryan, S., et al. 2015. “Cultural Stereotypes as Gatekeepers: Increasing Girls’ Interest in Computer Science and Engineering by Diversifying Stereotypes.” *Frontiers in Psychology* 6(49): 1-8.

Cheryan, S., Ziegler, et al. 2017. “Why Are Some STEM Fields More Gender Balanced Than Others?.” *Psychological Bulletin.* http://dx.doi.org/10.1037/bul0000052

Chesler, Naomi C., and Mark A. Chesler. 2002. "Gender-Informed Mentoring Strategies for Women Engineering Scholars: On Establishing a Caring Community." *Journal of Engineering Education* 91(1): 49-55.

Collomb, Anaıs, et al. 2014. "A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation." Rapport de recherche RR-LIRIS-2014-002.

Daily, Shaundra Bryant, et al. 2007. "The Development of Social Capital in Engineering Education to Improve Student Retention." ASEE Southeast Section Conference, Louisville, KY.

Dalrymple, Odesma, and Demetra Evangelou. 2006. "The Role of Extracurricular Activities in the Education of Engineers." 9th International Conference on Engineering Education, San Juan, PR.

Danielak, Brian A., et al. 2014. "Marginalized Identities of Sense-Makers: Reframing Engineering Student Retention." *Journal of Engineering Education* 103(1): 8-44.

De Pillis, Emmeline, and Lisette de Pillis. 2008. "Are Engineering Schools Masculine and Authoritarian? The Mission Statements Say Yes." *Journal of Diversity in Higher Education* 1(1): 33-44.

De Welde, K., & Laursen, S. L. 2008. "The 'Ideal Type' Advisor: How Advisors Help STEM Graduate Students Find Their 'Scientific Feet'." *The Open Education Journal* 1(1): 49-61.

Dharmadhikari, Shweta C., et al. 2011. "Empirical Studies on Machine Learning Based Text Classification Algorithms." *Advanced Computing* 2(6): 161-169.

Diesner, J., et al. 2005. "Mental Models of Data Privacy and Security Extracted from Interviews with Indians." In *55th Annual Conference of the International Communication Association (ICA), New York, NY*.

Ding, X., et al. 2008. "A Holistic Lexicon-Based Approach to Opinion Mining." In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (pp. 231-240). ACM.

Dou, Remy, and Eric Brewe. 2014. "Network Centrality and Student Self-Efficacy in an Interactive Introductory Physics Environment." Retrieved from http://www.percentral.org/perc/2014/files/REVISEDFORPERC2014DOUANDBREWE1.pdf

Dray, S. M., et al. 2013. Leveraging the Progress of Women in the HCI Field to Address the Diversity Chasm. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2399-2406). ACM.

Dryburgh, H. 1999. "Work Hard, Play Hard: Women and Professionalization in Engineering—Adapting to the Culture." *Gender & Society* 13(5): 664-682.

Eddy, S. L., et al. 2014. "Gender Gaps in Achievement and Participation in Multiple Introductory Biology Classrooms." *CBE—Life Sciences Education* 13(3): 478-492.

Ellis, J., et al. 2016. "Women 1.5 Times More Likely to Leave STEM Pipeline after Calculus Compared to Men: Lack of Mathematical Confidence a Potential Culprit." *PloS One* 11(7): e0157447.

Else-Quest, N. M., et al. 2010. Cross-National Patterns of Gender Differences in Mathematics: a Meta-Analysis. *Psychological Bulletin* 136(1): 103-127.

Espinosa, Lorelle. 2011. “Pipelines and Pathways: Women of Color in Undergraduate STEM Majors and the College Experiences that Contribute to Persistence.” *Harvard Educational Review* 81(2): 209-241.

Feeney, Mary, and Margarita Bernal. 2010. “Women in STEM networks: Who Seeks Advice and Support from Women Scientists?.” *Scientometrics* 85(3): 767-790.

Fisher, Dara R., et al. 2014. “Fostering 21st Century Skills in Engineering Undergraduates through Co-Curricular Involvement.” *American Society for Engineering Education*.

Fiske, S. T., & Markus, H. R. 2012. *Facing Social Class: How Societal Rank Influences Interaction*. New York, NY: Russell Sage Foundation.

Fox, Mary Frank. 2000. “Organizational Environments and Doctoral Degrees Awarded to Women in Science and Engineering Departments.” *Women's Studies Quarterly* 28(1/2): 47-61.

Fox, Mary Frank, et al. 2011. Programs for Undergraduate Women in Science and Engineering: Issues, Problems, and Solutions. *Gender & Society* 25(5): 589-615.

Furstenburg, F. 1968. “Structural Changes in the Working Class: A Situational Study of Workers in the Western German Chemical Industry.” In John A. Jackson (ed.), *Social Stratification*, 145-174. London: Cambridge University Press.

Ganley, C. M., and Lubienski, S. T. 2016. “Mathematics Confidence, Interest, and Performance: Examining Gender Patterns and Reciprocal Relations.” *Learning and Individual Differences* 47: 182-193.

Gappa, J., et al. 2007. *Rethinking Faculty work: Higher education’s Strategic Imperative*. San Francisco: Jossey-Bass.

Gardiner, Maria, et al. 2007. “Show Me the Money! An Empirical Analysis of Mentoring Outcomes for Women in Academia.” *Higher Education Research & Development* 26(4): 425-442.

Gardner, S. K. 2013. “Women Faculty Departures from a Striving Institution: Between a Rock and a Hard Place. *The Review of Higher Education* 36(3): 349-370.

Gill, Judith, et al. 2008. “I Still Wanna Be an Engineer! Women, Education and the Engineering Profession.” *European Journal of Engineering Education* 33(4): 391-402.

Gilmartin, S., et al. 2007. “Gender Ratios in High School Science Departments: The Effect of Percent Female Faculty on Multiple Dimensions of Students’ Science Identities.” *Journal of Research in Science Teaching* 44(7): 980-1009.

Glass, Jennifer L., et al. 2013. “What's so Special about STEM? A Comparison of Women’s Retention in STEM and Professional Occupations.” *Social Forces* 92(2): 723-756.

Glynn, M. A., & Abzug, R. 2002. “Institutionalizing Identity: Symbolic Isomorphism and Organizational Names.” *Academy of Management Journal* 45(1): 267-280.

Griffith, Amanda L. 2010. “Persistence of Women and Minorities in STEM Field Majors: Is It the School that Matters?.” *Economics of Education Review* 29(6): 911-922.

Gupta, Namrata. 2007. “Indian Women in Doctoral Education in Science and Engineering: A Study of Informal Milieu at the Reputed Indian Institutes of Technology.” *Science, Technology & Human Values* 32(5): 507-533.

Haines, V. A., & Wallace, J. E. 2003. “Gender-Role Attitudes, Perceptions of Engineering, and Beliefs about Women in Engineering ‘Having It All’: Are Male and Female Engineering Undergraduates Really so Different?.” *Alberta Journal of Educational Research* 49(4): 376-379.

Hazari, Zahra, et al. 2013a. “Factors that Affect the Physical Science Career Interest of Female Students: Testing Five Common Hypotheses.” *Physical Review Special Topics-Physics Education Research* 9(2): 1-8.

Hazari, Z., et al. 2013b. “The Science Identity of College Students: Exploring the Intersection of Gender, Race, and Ethnicity.” *Journal of College Science Teaching* 42(5): 82-91.

Heaverlo, C. A., et al. 2013. “STEM Development: Predictors for 6th-12th Grade Girls' Interest and Confidence in Science and Math.” *Journal of Women and Minorities in Science and Engineering* 19(2): 121-142.

Heflin, D. L. 2015. “They Care What You Wear: Gendered Practices in University Internship Manuals.” *Disestablishmentarian* 1(1): 32-52.

Herrmann, S. D., et al. 2016. “The Effects of a Female Role Model on Academic Performance and Persistence of Women in STEM Courses.” *Basic and Applied Social Psychology* 38(5): 258-268.

Hilbe, J. M. 2011. “Logistic Regression.” In *International Encyclopedia of Statistical Science* (pp. 755-758). Springer, Berlin, Heidelberg.

Hill, P. W., et al. 2014. “The New STEM Faculty Profile: Balancing Family and Dual Careers.” *Gender Transformation in the Academy* 19: 3-20.

Hirshfield, Laura E. 2010. “She Won’t Make Me Feel Dumb: Identity Threat in a Male-Dominated Discipline.” *International Journal of Gender, Science and Technology* 2(1): 5-24.

Hosmer, D. W., and Lemeshow, S. 1998. *Applied Logistic Regression*. John Wiley & Sons, Inc.: New York, New York.

Huffman, M. L., et al. 2010. “Engendering Change: Organizational Dynamics and Workplace

Gender Desegregation, 1975–2005." *Administrative Science Quarterly* 55(2): 255-277.

Ihme, T. A., et al. 2016. "How University Websites' Emphasis on Age Diversity Influences Prospective Students' Perception of Person-Organization Fit and Student Recruitment." *Research in Higher Education* 57(8): 1010-1030.

Iskander, E. T., et al. 2013. "Gender Differences in Expressed Interests in Engineering-Related Fields ACT 30-Year Data Analysis Identified Trends and Suggested Avenues to Reverse Trends." *Journal of Career Assessment* 21(4): 599-613.

James, Gareth, et al. 2017. *An Introduction to Statistical Learning*. Vol. 112. New York: Springer.

Kanter, Rosabeth Moss. 1977. "Some Effects of Percentages on Group Life: Skewed Sex Ratios and Responses to Token Women." *American Journal of Sociology* 82(5): 965-990.

Karatzoglou, Alexandros, et. al. 2005. "Support Vector Machines in R." *Research Report Series / Department of Statistics and Mathematics*, 21. Department of Statistics and Mathematics, WU Vienna University of Economics and Business, Vienna.

Kasarda, Mary, et al. 2010. "Work in progress—Initial Identification of Program Components Leading to Retention of Women in a Pre-Engineering High School Program, and an Undergraduate Engineering Program." *2010 IEEE Frontiers in Education Conference (FIE)*.

Kegen, Nadine V. 2013. "Science Networks in Cutting-Edge Research Institutions: Gender Homophily and Embeddedness in Formal and Informal Networks." *Procedia-Social and Behavioral Sciences* 79: 62-81.

Kessel, Cathy. 2014. "Understanding Underrepresentation: Women in Mathematics and Other Fields." *The Mathematical Intelligencer* 36(4): 10-18.

Kmec, Julie A. 2013. "Why Academic STEM Mothers Feel They Have to Work Harder than Others on the Job." *International Journal of Gender, Science, and Technology* 5(2): 79-101.

Kongar, Elif, et al. 2008. "Increasing the Participation of Women in the Engineering and Technical Services Industries." *American Society for Engineering Education*.

Kundi, Fazal Masud, et al. 2014. "Lexicon-Based Sentiment Analysis in the Social Web." *Journal of Basic and Applied Scientific Research* 4(6): 238-248.

Kuznetsova, A., et al. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82(13):1-26. doi: 10.18637/jss.v082.i13

Latimer, M., et al. (2014). "Organizational Change and Gender Equity in Academia: Using Dialogical Change to Promote Positive Departmental Climates." In *Gender Transformation in the Academy* (pp. 333-353). Emerald Group Publishing Limited.

Lee, J. J. 2007. "The Shaping of the Departmental Culture: Measuring the Relative Influences of the Institution and Discipline." *Journal of Higher Education Policy and Management* 29(1): 41-55.

Leslie, S. J., et al. 2015. "Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines." *Science* 347(6219): 262-265.

Li, Tao et al. 2006. "Using Discriminant Analysis for Multi-Class Classification: An Experimental Investigation." *Knowledge and Information Systems* 10(4): 453-472.

Lilleberg, J., et al. 2015. "Support Vector Machines and word2vec for Text Classification with Semantic Features." In *Cognitive Informatics & Cognitive Computing (ICCI* CC), 2015 IEEE 14th International Conference on* (pp. 136-140). IEEE.

Manning, C.D., et al. 2009. *An Introduction To Information Retrieval*. Retrieved from nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

Maranto, Cheryl, and Andrea Griffin. 2010. "The Antecedents of a Chilly Climate for Women Faculty in Higher Education." *Human Relations* 64(2): 139-159.

Marra, R.M., et al. 2012. "Leaving Engineering: A Multi-Year Single Institution Study." *Journal of Engineering Education* 101(1): 6-27.

Master, A., et al. 2016. "Computing Whether She Belongs: Stereotypes Undermine Girls' Interest and Sense of Belonging in Computer Science. *Journal of Educational Psychology* 108(3): 424-437.

Matchett, E., and T. Pippert. 2008. "We've Got Minorities, Yes We Do: Race and the College Viewbook." Paper presentated at the 70th annual meeting of the Midwest Sociological Society, St. Louis, MO.

Mayo, E. 1933. *The Human Problems of Industrial Civilization*. New York: Macmillan.

McCallum, A., and Nigam, K. 1998. "A Comparison of Event Models for Naive Bayes Text Classification." In *AAAI-98 workshop on learning for text categorization* (Vol. 752, No.

1 , pp. 41-48).

Mervis, J. 2011. "Weed-Out Courses Hamper Diversity." *Science* 334(6061): 1333-1333.

Meyer, D., and Wien, F. T. 2001. "Support Vector Machines." *R News* 1(3): 23-26.

Meyer, David, et al. 2017. e1071: MiscFunctions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.6-8.

Meyer, D., and Wien, F. T. 2017. "Support Vector Machines." Retrieved from http://cran.revolutionanalytics.com/web/packages/e1071/vignettes/svmdoc.pdf

Miller, D. T., et al. 1991. "Gender Gaps: Who Needs to be Explained?" *Journal of Personality* *and Social Psychology* 61:5–12. http://dx.doi.org/10.1037/0022-3514.61.1.5.

Mills, A. J. 1988. "Organization, Gender and Culture." *Organization Studies* 9(3): 351-369.

Mohammed, S., et al. 2000. The Measurement of Team Mental Models: We Have No Shared Schema." *Organizational Research Methods* 3(2): 123-165.

Moreau, Marie-Pierre, et al. 2010. "Constructions of Mathematicians in Popular Culture and Learners' Narratives: A Study of Mathematical and Non-Mathematical Subjectivities." *Cambridge Journal of Education* 40(1): 25-38.

Moreau, Marie-Pierre, and Heather Mendick. 2012. "Discourses of Women Scientists in Online Media: Towards New Gender Regimes?." *International Journal of Gender, Science and Technology* 4(1): 4-23.

Morton, Sarah. 2016. "The Impact of Women's Faculty and Leadership Representation on Women's Undergraduate Engineering Enrollment in the U.S." Working Paper.

National Science Foundation. 2012. Table 2–12. Freshmen Intending S&E Major, by Field, Sex, and Race/Ethnicity: 1995–2010. Arlington, VA: National Science Foundation. Retrieved from https://www.nsf.gov/statistics/seind12/append/c2/at02-12.pdf

National Science Foundation. 2015. "Women, Minorities, and Persons with Disabilities in Science and Engineering: 2015." Special Report NSF 15-311. Arlington, VA. Retrieved from http://www.nsf.gov/statistics/2015/nsf15311/digest/nsf15311-digest.pdf

National Science Foundation. 2016. Integrated postsecondary education data system, 2015, completions survey. National Center for Science and Engineering Statistics: Integrated Science and Engineering Resources Data System (WebCASPAR). Retrieved from https://webcaspar.nsf.gov

National Science Foundation. 2017. "TABLE 9-16. Median annual salary of scientists and engineers employed full time, by broad occupation, age, highest degree level, and sex: 2015." Retrieved from https://www.nsf.gov/statistics/2017/nsf17310/data.cfm

Ong, Maria. 2005. "Body Projects of Young Women of Color in Physics: Intersections of Gender, Race, and Science." *Social Problems* 52(4): 593-617.

Osei-Kofi, Nana, and Lisette E. Torres. 2015. "College Admissions Viewbooks and the Grammar of Gender, Race, and STEM." *Cultural Studies of Science Education* 10(2): 527-544.

Pang, B., and Lee, L. 2004. "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts." In *Proceedings of the 42nd annual meeting on Association for Computational Linguistics* (p. 271). Association for Computational Linguistics.

Parker, M., & Welch, E. W. 2013. "Professional Networks, Science Ability, and Gender Determinants of Three Types of Leadership in Academic Science and Engineering." *The Leadership Quarterly* 24(2): 332-348.

Parson, Laura. 2016. "Are STEM Syllabi Gendered? A Feminist Critical Discourse Analysis." *The Qualitative Report* 21(1): 102-116.

Price, Joshua. 2010. "The Effect of Instructor Race and Gender on Student Persistence in STEM Fields." *Economics of Education Review* 29(6): 901-910.

Rajani, Nazneen Fatema, and Raymond J. Mooney. 2016. "Combining Supervised and Unsupervised Ensembles for Knowledge Base Population." *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP-16).*

Rhoten, Diana, and Stephanie Pfirman. 2007. "Women in Interdisciplinary Science: Exploring Preferences and Consequences." *Research Policy* 36(1): 56-75.

Richman, Laura Smart, et al. 2011. "How Women Cope: Being a Numerical Minority in a Male-Dominated Profession." *Journal of Social Issues* 67(3): 492-509.

Riegle-Crumb, C., Moore, C., and Ramos-Wada, A. 2011. "Who Wants to Have a Career in Science or Math? Exploring Adolescents' Future Aspirations by Gender and Race/Ethnicity." *Science Education* 95(3): 458-476.

Riegle-Crumb, C., et al. 2012. "The More Things Change, the More They Stay the Same? Prior Achievement Fails to Explain Gender Inequality in Entry into STEM College Majors Over Time." *American Educational Research Journal* 49(6): 1048-1073.

Riegle-Crumb, Catherine, and Chelsea Moore. 2014. "The Gender Gap in High School Physics: Considering the Context of Local Communities." *Social Science Quarterly* 95(1): 253-268.

Riffle, Rebecca, et al. 2013. "A Mixed Methods Study of Gender, STEM Department Climate, and Workplace Outcomes." *Journal of Women and Minorities in Science and Engineering* 19(3): 227-243.

Rindfleish, J., and Sheridan, A. 2003. "No Change from Within: Senior Women Managers' Response to Gendered Organizational Structures. *Women in Management Review* 18(6): 299-310.

Robst, John, Jack Keil, and Dean Russo. 1998. "The Effect of Gender Composition of Faculty on Student Retention." *Economics of Education Review* 17(4): 429-439.

Roth, W. D. and Sonnert, G. 2011. "The Costs and Benefits of 'Red Tape': Anti-Bureaucratic Structure and Gender Inequity in a Science Research Organization." *Social Studies of Science* 41(3): 385-409.

Rudman, Laurie A., and Peter Glick. 2001. "Prescriptive Gender Stereotypes and Backlash toward Agentic Women." *Journal of Social Issues* 57(4): 743-762.

Sadler, P. M., et al. 2012. Stability and Volatility of STEM Career Interest in High School: A Gender Study. *Science Education* 96(3): 411-427.

Sanders, Karin, et al. 2009. "Views from above the Glass Ceiling: Does the Academic Environment Influence Women Professors' Careers and Experiences?." *Sex Roles* 60(5-6): 301-312.

Sargent, C. 2009. "Playing, Shopping, and Working as Rock Musicians: Masculinities in 'De-Skilled' and 'Re-Skilled' Organizations." *Gender & Society* 23(5): 665-687.

Schiebinger, Londa, and Martina Schraudner. 2011. "Interdisciplinary Approaches to Achieving Gendered Innovations in Science, Medicine, and Engineering." *Interdisciplinary Science Reviews* 36(2): 154-167.

Schneeweis, Nicole, and Martina Zweimüller. 2012. "Girls, Girls, Girls: Gender Composition and Female School Choice." *Economics of Education Review* 31(4): 482-500.

Sharpe, Norean Radke, and Gerhard Sonnert. 1999. "Pecentages of Women Faculty and Students in the Mathematical Sciences: A Trend Analysis by Institutional Group." *Journal of Women and Minorities in Science and Engineering* 5(1): 17-27.

Simon, H. A. 1945. *Administrative Behavior*. New York: Free Press.

Skuratowicz, E., and Hunter, L. W. 2004. "Where do Women's Jobs Come from? Job Resegregation in an American Bank." *Work and Occupations* 31(1): 73-110.

Smith-Doerr, L. 2004. “Flexibility and Fairness: Effects of the Network Form of Organization on Gender Equity in Life Science Careers. *Sociological Perspectives* 47(1): 25-54.

Smith-Doerr, Laurel. 2005. “Institutionalizing the Network Form: How Life Scientists Legitimate Work in the Biotechnology Industry.” *Sociological Forum* 20(2): 271-299.

Smith-Doerr, Laurel, et al. 2016. “Doing Gender and Responsibility: Scientists and Engineers Talk About Their Work.” *Journal of Women and Minorities in Science and Engineering* 22(1): 49-68.

Soelistio, Y. E., and Surendra, M. R. S. 2015. “Simple Text Mining for Sentiment Analysis of Political Figure Using Naive Bayes Classifier Method.” *arXiv preprint arXiv:1508.05163*.

Soni, Rishabh, and Mathai, K. James. 2016. “Effective Sentiment Analysis of a Launched Product using Clustering and Decision Trees.” *International Journal of Innovative Research in Computer and Communication Engineering* 4(1): 884-891.

Sonnert, Gerhard, et al. 2007. “Undergraduate Women in Science and Engineering: Effects of Faculty, Fields, and Institutions Over Time.” *Social Science Quarterly* 88(5): 1333-1356.

Spalter-Roth, Roberta M., et al. 2011. “Homosociality or Crossing Race/Ethnicity/Gender Boundaries? Pipeline Interventions and the Production of Scholarly Careers.” Paper presented at the annual meeting of the American Sociological Association, August 21, Las Vegas, NV.

Stainback, K., et al. 2011. “The Context of Workplace Sex Discrimination: Sex Composition, Workplace Culture and Relative Power.” *Social Forces* 89(4): 1165-1188.

Stainback, Kevin, et al. 2016. “Women in Power Undoing or Redoing the Gendered Organization?.” *Gender & Society* 30(1): 109-135.

Stearns, Elizabeth, et al. 2016. “Demographic Characteristics of High School Math and Science Teachers and Girls’ Success in STEM.” *Social Problems* 63(1): 87-110.

Stoet, G., & Geary, D. C. 2018. The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education. *Psychological Science* 29(4): 581-593.

Storage, D., et al. 2016. The Frequency of “Brilliant” and “Genius” in Teaching Evaluations Predicts the Representation of Women and African Americans across Fields. *PloS One* 11(3): e0150194.

Stout, J. G., et al. 2011. “STEMing the Tide: Using Ingroup Experts to Inoculate Women’s Self-Concept in Science, Technology, Engineering, and Mathematics (STEM).” *Journal of Personality and Social Psychology* 100(2): 255-270.

Stout, Jane G., et al. 2013. “How a Gender Gap in Belonging Contributes to the Gender Gap in Physics Participation.” *AIP Conference Proceedings*. Vol. 1513. No. 1.

Taboada, Maite, et al. 2011. “Lexicon-Based Methods for Sentiment Analysis.” *Computational Linguistics* 37(2): 267-307.

Traxler, Adrienne L. 2015. “Community Structure in Introductory Physics Course Networks.” *arXiv preprint arXiv:1507.04695*.

Tidball, M. Elizabeth. 1986. “Baccalaureate Origins of Recent Natural Science Doctorates.” *The Journal of Higher Education* 57(6): 606-620.

Uriarte, María, et al. 2007. “Constructing a Broader and More Inclusive Value System in Science.” *BioScience* 57(1): 71-78.

Van Rijnsoever, F. J., & Hessels, L. K. 2011. “Factors Associated with Disciplinary and Interdisciplinary Research Collaboration.” *Research Policy* 40(3): 463-472.

Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S. (MASS) Fourth

Edition (R Package). Springer, New York. ISBN 0-387-95457-0

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York, NY: Cambridge University Press.

Whitehead, A. L. 2013. "Gendered Organizations and Inequality Regimes: Gender, Homosexuality, and Inequality within Religious Congregations." *Journal for the Scientific Study of Religion* 52(3): 476-493.

Williams, Joan, et al. 2013. "Cultural Schemas, Social Class, and the Flexibility Stigma." *Journal of Social Issues* 69(2): 209-234.

Wingfield, Adia Harvey. 2009. "Racializing the Glass Escalator: Reconsidering Men's Experiences with Women's Work." *Gender & Society* 23(1): 5-26.

Yang, Yuehai, et al. 2014. "A Study of Informal Learning Communities: a Tale of Two Physics Courses." Retrieved from http://www.per-central.org/perc/2014/files/2198_AStudyofInformalLearningCommunitiesaTaleofTwoPhysicsCourses.pdf

Zhang, Lei, et al. 2011. "Combining Lexicon-Based and Learning-Based Methods for Twitter Sentiment analysis." *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE)* 1-7.

Zhao, L., et al. 2001. "Comparison of Logistic Regression and Linear Regression in Modeling Percentage Data." *Applied and Environmental Microbiology* 67(5): 2129-2135.

Zweben, Stuart H., and Elizabeth B. Bizot. 2016. "Representation of Women in Postsecondary Computing: Disciplinary, Institutional, and Individual Characteristics." *Computing in Science & Engineering* 18(2): 40-56.

APPENDIX

PYTHON TEXT MINER CODE

```
#need new print function since using Python 2.7.13
#(can comment out if using Python 3)
from __future__ import print_function

import ssl
from functools import wraps
def sslwrap(func):
    @wraps(func)
    def bar(*args, **kw):
        kw['ssl_version'] = ssl.PROTOCOL_TLSv1
        return func(*args, **kw)
    return bar

ssl.wrap_socket = sslwrap(ssl.wrap_socket)

#creates a text miner that extracts text data from a dataset of URLs
#and stores them into files named by a dataset of path names
#If you receive SSLError message, run pip install -U requests[security] --no-
#cache in the terminal

#import packages (need to download these modules first)
import requests
from bs4 import BeautifulSoup

#get URLs by entering path for URL data and open file in read mode
URL_data = '/Users/sarahmorton/Desktop/URLs PSYC.txt'
URL_names = open(URL_data, 'r')

#enter path for path name data and open file in read mode
main_path = '/Users/sarahmorton/Desktop/Paths PSYC.txt'
file_names = open(main_path, 'r')

#For each URL, extract text data and write to file

for URLs, files in zip(URL_names, file_names):

    #Strip '\n' so that the URLs have valid names
    URL = str(URLs.rstrip('\n'))

    #display the URL name
    print(URL)

    #get URL from online
    URL = requests.get(URL)

    #get URL status code (should be around 200)
    status_code = URL.status_code
    print(status_code)

    #extract text data from URL
    soupURL = BeautifulSoup(URL.content, 'html.parser')

    name_of_file = str(files.rstrip('\n'))
```

```
with open(name_of_file, 'w+') as f:
    for ptag in soupURL.find_all('p'):
        data = ptag.text.encode('utf-8') + '\n'
        f.write(data)

print('Data written to', name_of_file)
```