ASSESSING IMPACTS OF NEGATIVE STEREOTYPES AND DESIGNING THEORY-DRIVEN INTERVENTIONS TO SUPPORT UNDERREPRESENTED MINORITIES IN STEM

Deborah Wu
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2

Part of the Social Psychology Commons

Recommended Citation
Wu, Deborah, "ASSESSING IMPACTS OF NEGATIVE STEREOTYPES AND DESIGNING THEORY-DRIVEN INTERVENTIONS TO SUPPORT UNDERREPRESENTED MINORITIES IN STEM" (2022). Doctoral Dissertations. 2483.
https://doi.org/10.7275/26417334.0 https://scholarworks.umass.edu/dissertations_2/2483

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
ASSESSING IMPACTS OF NEGATIVE STEREOTYPES AND DESIGNING THEORY-DRIVEN INTERVENTIONS TO SUPPORT UNDERREPRESENTED MINORITIES IN STEM

A Dissertation Presented

by

DEBORAH JEAN WU

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 2022

Department of Psychological and Brain Sciences Social Psychology
Assessing Impacts of Negative Stereotypes and Designing Theory-Driven Interventions to Support Underrepresented Minorities in STEM

A Dissertation Presented

By

DEBORAH JEAN WU

Approved as to style and content by:

Nilanjana Dasgupta, Chair

Jennifer McDermott, Member

Evelyn Mercado, Member

Joya Misra, Member

Farshid Hajir, Department Chair
Psychological and Brain Sciences
ACKNOWLEDGMENTS

During my time here at UMass and Amherst, the support that I have received from friends, family, and advisors have made this dissertation possible. I am also indebted to my collaborators and the research assistants who spent numerous hours in helping me collect my data.

To Buju – thank you for your mentorship these past five and half years. I am endlessly grateful for your patience, communication, and feedback. I have become a much better researcher, writer, and speaker under your guidance. To Jen – thank you for listening to anything that I have on my mind, for welcoming me into the world of developmental science, and for your encouragement. To Jiyoung – thank you for teaching me basically everything I know about social neuroscience and for your always open door during my first year. To Evelyn – thank you for listening, for understanding, and for your support. To Joya – thank you for your thoughtful comments and your time.

I am also incredibly fortunate to have been surrounded by some of the most caring coworkers who are also my dearest friends. To Adrian, Joel, Marielena, Min, Kelsey, Nathan, and SungHa, thank you for your support, and for being there whenever I need your feedback or just a listening ear. To Stelios, thank you for being you. Your never-ending encouragement, attentiveness, and brilliant mind have made my life colorful.

My family has been with me in every step of my life. To my parents and Matt – thank you for always praying and loving me. To my dearest sisters Esther and Priscilla – words cannot describe how proud I am of you and of us, for overcoming everything life has thrown at us. To Lydia – I miss you. Your faith and grace in the face of overwhelming odds will always be my inspiration.
ABSTRACT

ASSESSING IMPACTS OF NEGATIVE STEREOTYPES AND DESIGNING THEORY-DRIVEN INTERVENTIONS TO SUPPORT UNDERREPRESENTED MINORITIES IN STEM

FEBRUARY 2022

DEBORAH JEAN WU, B.A., NORTHWESTERN UNIVERSITY

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Dr. Nilanjana Dasgupta

The underrepresentation of women, racial ethnic minorities, and first-generation college students in science, technology, engineering, and mathematics (STEM) fields in the United States has been well-documented. Members of these minority groups face negative stereotypes casting doubt on their abilities in these fields, which can cause the concern that they will be judged through the lens of the stereotype and devalued. This concern is called social identity threat. This dissertation presents three investigations focusing on the experiences of underrepresented students in STEM, examining when and how altering situational contexts increases or decreases their vulnerability to social identity threat. Chapter 1 is a literature review summarizing theories underlying these investigations. Chapter 2 tested how subtle gender stereotypic cues affect women’s attentional vigilance on a math related task. Chapter 3 examined the long-term impacts of same-sex peer mentorship on women’s academic experiences and retention in engineering. Chapter 4 tested whether immersion in a living learning community designed for first-generation students benefitted their academic experiences and retention
in the biological sciences. Chapter 5 concludes by discussing the theoretical contribution of all three studies, presenting recommendations for application, and posing suggestions for future research.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Social Identity Threat</td>
<td>3</td>
</tr>
<tr>
<td>Overview of Three Papers</td>
<td>8</td>
</tr>
<tr>
<td>2. SUBTLE GENDER CUES ACTIVATE WOMEN’S NEURAL VIGILANCE</td>
<td>12</td>
</tr>
<tr>
<td>Method</td>
<td>17</td>
</tr>
<tr>
<td>Results</td>
<td>23</td>
</tr>
<tr>
<td>Discussion</td>
<td>33</td>
</tr>
<tr>
<td>Supplemental Analyses</td>
<td>38</td>
</tr>
<tr>
<td>Transition to Chapter 3</td>
<td>46</td>
</tr>
<tr>
<td>3. PEER MENTORING FOR WOMEN IN ENGINEERING</td>
<td>48</td>
</tr>
<tr>
<td>Method</td>
<td>55</td>
</tr>
<tr>
<td>Results</td>
<td>61</td>
</tr>
<tr>
<td>Discussion</td>
<td>73</td>
</tr>
<tr>
<td>Supplemental Analyses</td>
<td>78</td>
</tr>
<tr>
<td>Transition to Chapter 4</td>
<td>81</td>
</tr>
<tr>
<td>4. A LIVING LEARNING COMMUNITY FOR FIRST-GENERATION STUDENTS IN BIOLOGY</td>
<td>83</td>
</tr>
<tr>
<td>Study 1 Method</td>
<td>91</td>
</tr>
<tr>
<td>Study 1 Results</td>
<td>97</td>
</tr>
<tr>
<td>Study 1 Discussion</td>
<td>103</td>
</tr>
<tr>
<td>Study 2 Method</td>
<td>104</td>
</tr>
<tr>
<td>Study 2 Results</td>
<td>107</td>
</tr>
<tr>
<td>Discussion</td>
<td>113</td>
</tr>
<tr>
<td>Supplemental Analyses</td>
<td>117</td>
</tr>
</tbody>
</table>
5. GENERAL DISCUSSION

Summary of Key Findings.................................................................123
Theoretical Implications and Applied Considerations............................126
Directions for Future Research............................................................130
Conclusion...........................................................................................133

BIBLIOGRAPHY......................................................................................135
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Descriptive Statistics of the Post-Task Questionnaire Variables</td>
<td>26</td>
</tr>
<tr>
<td>2. Correlations between Exploratory Measures and the Face Priming</td>
<td>43</td>
</tr>
<tr>
<td>Effect on ERN</td>
<td></td>
</tr>
<tr>
<td>3. Descriptive Statistics of Survey Variables</td>
<td>117</td>
</tr>
<tr>
<td>4. Condition Differences of Survey Variables Separated by Cohort</td>
<td>118</td>
</tr>
<tr>
<td>(Having In-Person or Virtual Courses)</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Examples of a neutral, congruent, or incongruent target; sample images of a male face and a scrambled face; a sample trial structure</td>
<td>21</td>
</tr>
<tr>
<td>2. Histograms of the distribution of the STEM career question for women and men</td>
<td>25</td>
</tr>
<tr>
<td>3. Grand averaged event-related brain potentials elicited by errors and correct responses as a function of the prime type at FCz and Cz electrodes for women</td>
<td>29</td>
</tr>
<tr>
<td>4. Grand averaged event-related brain potentials elicited by errors and correct responses as a function of the prime type at FCz and Cz electrodes for men</td>
<td>30</td>
</tr>
<tr>
<td>5. The face priming effect on ERN separated by gender and interest in pursuing STEM careers</td>
<td>33</td>
</tr>
<tr>
<td>6. Grand averaged event-related brain potentials elicited by errors and correct responses as a function of the prime type at Pz for men and women</td>
<td>45</td>
</tr>
<tr>
<td>7. Participants’ feelings of anxiety and motivation in engineering from college entry through graduation</td>
<td>64</td>
</tr>
<tr>
<td>8. Percentage of participants who successfully acquired and completed engineering internships while in college</td>
<td>65</td>
</tr>
<tr>
<td>9. Percentage of participants who obtained Bachelor’s degrees in engineering and STEM</td>
<td>66</td>
</tr>
<tr>
<td>10. Participants’ intentions to pursue graduate school in engineering from college entry through graduation</td>
<td>68</td>
</tr>
<tr>
<td>11. Participants’ emotional health and well-being from college entry through graduation</td>
<td>69</td>
</tr>
<tr>
<td>12. Participants’ confidence in engineering from college entry through graduation</td>
<td>72</td>
</tr>
<tr>
<td>13. Participants’ belonging in engineering from college entry through graduation</td>
<td>73</td>
</tr>
<tr>
<td>14. Item asking participants to describe how close they felt with their mentor</td>
<td>79</td>
</tr>
</tbody>
</table>
15. Participants’ feelings of belonging in Biology 151 in Study 1..............................100
16. Participants’ feelings of anxiety in Biology 151 in Study 1.................................101
17. Participants’ feelings of motivation in Biology 151 in Study 1...............................102
18. Participants’ final grades in Biology 151 in Study 1.............................................103
19. Participants’ feelings of belonging in Biology 152 in Study 2...............................108
20. Participants’ feelings of confidence in their abilities in biology in Study 2.............109
21. Participants’ perceived relevance of biology in the real-world in Study 2..............110
22. Percentage of participants retained in biological sciences majors
    in their second year of college..............................................................................112
CHAPTER 1

INTRODUCTION

Disparities in science, technology, engineering, and mathematics (STEM) pathways in the United States have been well-documented over the past few decades (Landivar, 2013). Women, racial ethnic minorities, and first-generation college students are underrepresented in STEM educational and workforce pathways (Ambrose, 2019; Bettencourt et al., 2020; Landivar, 2013; Xue & Larson, 2015). Multiple factors cause and aggravate these disparities from elementary and secondary education, higher education, and through the workforce, one of which is the prevalence of culturally pervasive stereotypes that characterize science, engineering, and technology as masculine domains and idealize brilliant scientists, engineers, and technologists as White and Asian men (Beasley & Fischer, 2012; Steele & Aronson, Steele et al., 2002; Charlesworth & Banaji, 2019). People who deviate from that idealized stereotype are perceived with doubt and their abilities in STEM are called into question (Beasley & Fischer, 2012; Cheryan & Plaut, 2010).

For example, a large body of research has established that women, African Americans, and Hispanics face negative stereotypes regarding their abilities and intelligence in STEM (e.g., Eaton et al., 2020; Starr, 2018; Steele & Aronson, 1995; Steele et al., 2002., for reviews, see Charlesworth & Banaji, 2019; Dennehy et al., 2018; Schmader, 2010). First-generation college students also face barriers in higher education, making it more difficult for them to enter and complete college in comparison to continuing-generation students (Bettencourt et al., 2020; Stephens et al., 2012). These barriers make it difficult for first-generation students to access highly-skilled STEM
professions that require at least a Bachelor’s degree for eligibility (Fayer et al., 2017). In sum, a combination of negative stereotypes associating STEM with White and Asian men, barriers in higher education, and challenging technical content that requires solid foundational knowledge starting in high school (Rodgers et al., 2014) often signals to women, underrepresented racial minorities (URM), and first-generation students that they do not have the ability necessary to succeed in STEM fields. Young people from underrepresented groups are likely to feel social identity threat in STEM, characterized by increased vigilance and worry about their ability in STEM contexts because of stereotypes casting doubt on their group’s ability, and by extension their own ability in STEM (Schmader, 2010; Steele et al., 2002).

Unlocking solutions to enhance students’ interest in, and pursuit of, STEM education and careers is of urgent national interest given the national demand for a 21st century workforce with scientific and technological skills and a lagging supply of STEM graduates (Zaza et al., 2020). The scarcity of STEM students is particularly pronounced for women, as well as Black, Hispanic, and Indigenous Americans (National Academies of Science, Engineering, and Medicine, 2019; National Center for Science and Engineering Statistics, 2021). Because skilled STEM jobs are significantly more lucrative than other types of jobs, this underrepresentation also contributes to gender- and race-based income disparities, even among people with college degrees (Melguizo & Wolniak, 2012; Xu, 2013). In order to increase the recruitment, retention, and success of underrepresented young people in STEM prior to workforce entry, it is essential to understand how STEM stereotypes impact them as students, and to design and test
interventions that increase their belonging, motivation, confidence, and decrease their worries and anxiety. This is the overarching goal of my dissertation.

My dissertation consists of three empirical papers that focus on the experiences of underrepresented students in STEM, which examine the impact of negative stereotypes on experiences of social identity threat, and tests whether this threat can be mitigated through interventions designed and implemented in naturally existing educational settings. The first paper (Chapter 2) investigates how negative stereotypes in STEM impact minorities’ experiences preconsciously (measured using electroencephalogram or EEG) as they complete a rapid math related task. The second (Chapter 3) and third (Chapter 4) papers design and test two interventions in naturally-existing field settings and examine their impacts on underrepresented students’ academic experiences in college and persistence in STEM pathways.

Together, these three papers use a wide range of method and measures to test interrelated research questions. In the following section, I briefly summarize theory and empirical research on social identity threat, which provides the theoretical framework for the current research. I then provide a description of each paper, followed by a plan to discuss the current work’s theoretical contributions, and propose research questions to guide future work.

Social Identity Threat

Social identity threat, or the concern about being devalued due to negative stereotypes associated with one’s social group (Branscombe et al., 1999; Steele et al., 2002), provides the framework for my research. Early research on the impacts of social identity threat on young people in STEM focused on the effects of negative stereotypes
on individuals’ performance—referred to as stereotype threat (Spencer et al., 1999; Steele et al., 2002). These studies showed that women, Black, and Hispanic students in college underperformed on academic tests when their identity was made salient or when the test was presented as reflective of their intelligence (Major et al., 1998; Quinn & Spencer, 2001; Schmader et al., 2003; Steele & Aronson, 1995). Beyond performance, stereotype threat also causes momentary declines in cognitive capacity and increases stress, resulting in worsened performance on working memory tasks (Beilock et al., 2007; Schmader & Johns, 2003) and increased physiological threat (i.e., sympathetic nervous system activation, Murphy et al., 2007; blood pressure, Blascovich et al., 2001). Collectively, these studies show that the effects of stereotype threat emerge on measures of performance as well as measures of physiology.

Elicitation of stereotype threat can be obtrusive or unobtrusive. In studies that used obtrusive manipulations, researchers informed participants that the tasks they were about to complete were measures of intelligence (e.g., Quinn & Spencer, 2001; Schmader et al., 2003; Steele & Aronson, 1995), thereby raising the specter of ability-based stereotypes about their ingroup. Other studies used situational cues to signal the same stereotypes unobtrusively (Cheryan et al., 2009; Dasgupta et al., 2015; Murphy et al., 2007). For example, women entered academic workspaces that had stereotypically masculine paraphernalia (e.g., videogames, Star Trek posters; Cheryan et al., 2009), asked to envision attending professional conferences and shown pictures of conference attendees who were mostly men (Murphy et al., 2007), or assigned to work teams populated by mostly men (Dasgupta et al., 2015). These studies showed that exposure to unobtrusive situational cues that signaled masculinity or the overrepresentation of men
made women more anxious and less interested in STEM as compared to situations that were gender neutral. These findings suggest that when underrepresented minorities are immersed in STEM-relevant situations, the design of the space, people in it, and the instructions they receive can activate stereotypes, feelings of inclusion or exclusion, and identity threat. These experiences typically unfold slowly over time. What is less known, however, is whether STEM-relevant stereotypic cues activate social identity threat unconsciously, within milliseconds. To answer this question, my first paper presents a study that tests how women interested in STEM process gender stereotypic situational cues while completing a math-related task. This study captures preconscious cognitive processing within milliseconds after exposure to stereotypic cues measured using neurophysiological measures (i.e., electroencephalogram or EEG).

Because the experience of social identity threat or stereotype threat can lead to worsened performance, increased vigilance, and decreased retention (e.g., Beasley & Fischer, 2012; Steele & Aronson, 1995; Cheryan et al., 2009), an important next step is to design and test effective interventions to mitigate these negative effects and promote academic thriving, success, and persistence. Past interventions backed by empirical research include teaching self-affirmation, growth mindset, and emotion regulation skills (Cohen et al., 2009; Miyake et al., 2010; Paunesku et al., 2015; Rozek et al., 2019; Yeager et al., 2019). In these interventions, students are taught to affirm their personal values (self-affirmation; Cohen et al., 2009; Miyake et al., 2010), to reframe intelligence as malleable (growth mindset; Yeager et al., 2019), or to reappraise their anxiety as positive rather than negative (emotion reappraisal; Rozek et al., 2019). While these mental reappraisal interventions have been successful in decreasing performance gaps,
they also place the burden of change on underrepresented individuals who have less power, rather than changing the culture of academic environments through actions of institutional leaders with considerably greater power.

Some interventions have focused on creating culture change through role models, relationships, and community. For example, several studies have demonstrated the efficacy of interventions providing exposure to successful ingroup role models and fostering relationships with successful peers (e.g., Cheryan et al., 2011; Dennehy & Dasgupta, 2017; Herrmann et al., 2016; Ramsey et al., 2013; Stout et al., 2011). This research is consistent with the Stereotype Inoculation Model (Dasgupta, 2011), which posits that exposure to successful ingroup role models and ingroup peers protect underrepresented students from social identity threat, allowing them to thrive in STEM.

In contrast to mental reappraisal techniques, which asks underrepresented minorities to mentally adjust to STEM environments where they face underrepresentation and negative stereotypes, interventions that afford ingroup role models and peers change the STEM environment and expand ingroup social networks.

Previous work on ingroup role model interventions have been beneficial for underrepresented minorities in STEM, however, much of this work was conducted in brief artificial laboratory settings and not in naturally existing environments (e.g., Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; Ramsey et al., 2013). These interventions were also one-time events, with outcomes being measured over a short period of time (e.g., up to two weeks after the interaction). Given these short-term artificial scenarios, students in these studies did not get to develop personal relationships with role models and the long-term impacts of the
intervention are unknown. To address these gaps, the second paper presents a study that examines the impact of a yearlong same-sex peer mentorship intervention on young women students entering college with an interest in engineering and tracks the long-term impacts of these relationships on their academic outcomes through college graduation.

Another type of peer-driven intervention involves creating a community of students with a shared social identity who are part of a living learning community (LLC) in college. Though LLCs have become more prevalent because they have been found to benefit students’ social experiences (e.g., Brower & Inkelas, 2010; Inkelas et al., 2007; Wawrzynski & Jessup-Anger, 2010), past work on the impact of LLCs for underrepresented students has been limited and mixed. One limitation is that in past research on underrepresented students in LLCs, these students were both a numeric minority at their universities and also within their LLCs. While some of these studies found that women and first-generation students benefited from LLCs even when they were in the minority (Inkelas et al., 2007; Szelenyi et al., 2013), other studies found that racial ethnic minorities suffered lower feelings of belonging compared to White students as a numeric minority in LLCs (Eidum et al., 2020; Inkelas et al., 2006; Johnson et al., 2007). Thus, in order to test whether LLCs can function as an effective peer intervention to mitigate social identity threat for underrepresented minorities in STEM, the design of the LLC needs to be specific to only students who share the same marginalized social identity so that they can function as a community of peers with common experiences on an important social dimension.

Another limitation is that previous studies that examine how exposure to ingroup peers affects minorities’ academic experiences have largely involved brief exposure to
such individuals (e.g., Cheryan et al., 2011; Johnson et al., 2019; Pietri et al., 2018; 2019) or a one-on-one relationship with a single person (Dennehy & Dasgupta, 2017). Little work has examined whether a learning community of peers who share a common social identity has benefits for underrepresented minorities in STEM. A third limitation of the previous literature is that almost all interventions investigating the benefits of ingroup peers for minorities in STEM have focused on gender (e.g., Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; Ramsey et al., 2013; for those based on race, see Johnson et al., 2019; Pietri et al., 2018; 2019). None thus far have examined the impacts of a learning community for first-generation students in STEM. To address these gaps in the literature, in the third paper, I present a study that examines the impact of a biological sciences living learning community for first-generation students.

**Overview of Three Papers**

This dissertation presents three separate empirical papers investigating the following research questions:

1. Do momentary cues signaling masculinity elicit preconscious neurological vigilance among women during STEM-related tasks?

2. To what extent does a yearlong peer mentorship have enduring benefits for women engineering students through their college experience until graduation, in terms of academic experiences, behavioral outcomes, emotional health, and retention in STEM? Does the sex of the peer mentor matter?
3. To what extent does a first-year living learning community for first-generation college students in biological sciences enhance their academic experiences, performance, and retention in STEM majors past their first year of college? While the first paper focuses on how negative stereotypes impact underrepresented students’ attentional vigilance in STEM in real-time, the second and third papers test two interventions designed to mitigate the negative effects of stereotypes on students’ real-world experiences and outcomes. Taken together, all three papers highlight how situational cues impact the experiences of underrepresented minorities in STEM. Each chapter presents one of these papers, as well as a short discussion section that transitions into the next manuscript.

In Chapter 2, I present a study that examines whether minimalistic social cues signaling masculinity elicit attentional vigilance from college women in STEM when completing a STEM-related task. In this study, male and female college students who were majoring or intending to major in STEM completed a “math intelligence” task. Before each trial, participants were randomly primed with a subtle masculine cue or a control image. This design tested whether participants would show increased attentional vigilance on trials that had a gender-relevant prime compared to trials when they viewed a control image. Attentional vigilance was measured using the error-related negativity or ERN component, a neurological measure (e.g., Amodio et al., 2004; Botvinick et al., 2004; Carter et al., 1998; Danielmeier et al., 2009). This manuscript was published in *Biological Psychology* in 2020 and is reproduced in this dissertation. In the bridge that transitions from Chapter 2 to Chapter 3, I discuss that just as situational cues signal to
minorities that they are unwelcome in STEM, they can also be designed to make STEM environments more welcoming to underrepresented students.

In Chapter 3, I present a longitudinal field experiment that tests the long-lasting impacts of a one-year peer mentoring intervention for women in engineering through their college experience until graduation. In this experiment, women in their first year of college were randomly assigned to have a female peer mentor, a male peer mentor, or no mentor. I tested whether: 1) being assigned a mentor enhances mentees’ academic experiences in engineering, emotional health and well-being, and STEM retention from first year through graduation as compared to a control condition; and 2) these benefits differ depending on the sex of the mentor (i.e., male or female). This manuscript is in revision at *Nature Communications* and is reproduced in this dissertation. A short discussion bridging from Chapter 3 to Chapter 4 will discuss that while one-on-one interactions with role models or mentors of the same stigmatized social identity has been found to be beneficial, it is also important to examine the impacts of other interventions such as the creation of a cohort or community among students who share the same marginalized identity.

In Chapter 4, I present two longitudinal quasi-experiments that examine the impact of a one-year living learning community (LLC) for first-generation (FG) college students in biological sciences. In both quasi-experiments, I compared three groups of participants, all of whom were students entering college with an interest in the biological sciences: 1) an intervention group with FG students in a biological sciences LLC; 2) a control group with FG students not in the LLC, and 3) an additional comparison group with honors college students, mostly from college-educated families, who participated in
a separate biological sciences LLC. I tested the impact of the FG LLC intervention (relative to the other two conditions) on students’ academic experiences, perceptions of the real-world relevance of biology, and grades in biology in both semesters during their first year. I also examined their retention in biological sciences majors in their second year. This manuscript is currently in preparation for submission to *Nature Human Behaviour* and is reproduced in this dissertation.

Finally, Chapter 5 summarizes each manuscript’s findings and discusses the theoretical implications of this research. Specifically, I argue that the current work adds to the literature by showing that underrepresented students in STEM experience vigilance and threat in ways that are preconscious and occur at early stages of attention processing. I also demonstrate that local academic contexts can be leveraged to signal inclusion, as shown by the latter two papers in my dissertation. Importantly, these interventions were conducted during transition periods, such as the beginning of college, when life feels uncertain to incoming students and different from their high school experience. Fostering peer relationships and building communities for underrepresented minorities during this transition period provides a psychologically secure foundation that set them up for long-term success. Evidence from the latter two studies provides scientific evidence supporting the development and scaling of academic programs at universities that should be used by campus leaders interested in diversity, equity, and inclusion in STEM. Lastly, I suggest directions for future interventions for underrepresented minorities in STEM.
CHAPTER 2

SUBTLE GENDER CUES ACTIVATE WOMEN’S NEURAL VIGILANCE

A central theme in social psychology is that people’s attitudes, beliefs, and behavior are often shaped by factors that lie outside their awareness (Banaji & Dasgupta, 1998; Bargh, 1997; Greenwald & Banaji, 1995; Nisbett & Wilson, 1977). Through immersion in an unequal society and passive observation, human minds learn that social groups are differentially associated with particular roles and attributes that vary in status and power (Dasgupta, 2013). These mental associations are called implicit or automatic stereotypes, which can be passively learned even though they may not be actively endorsed by individuals (Blair et al., 2014; Dasgupta, 2004; Greenwald & Banaji, 1995). For example, with regard to gender, a large literature has documented the ubiquity of implicit gender stereotypes (for a review, see Ellemers, 2018), showing that both women and men more readily associate: (a) women with domestic roles and men with professional roles (Banaji & Hardin, 1996; Blair & Banaji, 1996); (b) women with communal traits and men with agentic traits (Asgari, Dasgupta, & Stout, 2012; Dasgupta & Asgari, 2004; Eagly & Karau, 2002; White & Gardner, 2009); and (c) women with service-oriented careers and men with careers in science, technology, engineering and mathematics or STEM (Miller et al., 2018; Oakhill et al., 2005; Stout et al., 2011).

In particular, implicit stereotypes about gender and STEM have profound effects on girls’ and women’s interest, confidence, and persistence in STEM education and career pathways (Dasgupta & Stout, 2014; Dasgupta, 2011). For example, women who exhibit stronger implicit stereotypes of associating men (more than women) with STEM-oriented professions feel less confident in their ability and are less likely to be interested
in STEM careers (Miller et al., 2015; Nosek, Banaji, & Greenwald, 2002; Stout et al., 2011). Notably, according to studies using large national (Wang et al., 2013) and international samples (Stoet & Geary, 2018), women either outperform men or perform equally well in math and science; however, there continues to be a large confidence gap based on gender, likely due to the cultural prevalence of negative stereotypes that cast doubt on women’s abilities and their place in STEM (Dasgupta & Stout, 2014; Dasgupta, 2011; Dar-Nimrod & Heine, 2006; Leslie et al., 2015).

Due to their concern about negative gender stereotypes, women in STEM fields are highly vigilant to the effects of situational cues signaling potential identity threat. Previous research suggests that situations that activate gender-STEM stereotypes hamper women’s interest and motivation in STEM and increase anxiety about their performance (e.g., Beilock et al., 2007; Cheryan et al., 2009; Dasgupta et al., 2015; Murphy et al., 2007). For example, in one study, when women were explicitly told that men were better at math, they performed worse than men on a math test, whereas in the absence of this explicit statement, women and men performed equally well (Beilock et al., 2007).

Additional evidence suggests that stereotype-eliciting cues do not need to be explicit. When exposed to computer science classroom environments containing stereotypically masculine cues (e.g., videogames, Star-Trek posters), women reported decreased interest in computer science than when they were exposed to classrooms with gender-neutral cues (e.g., nature posters; Cheryan et al., 2009). Other situational cues, such as exposure to work teams and conferences populated by mostly men, also increased women’s anxiety while decreasing their interest in pursuing STEM (Dasgupta et al., 2015; Murphy et al., 2007).
Collectively, these findings suggest that when women are immersed in STEM-relevant situations, the presence of situational cues signaling male dominance activates gender stereotypes. Notably, these studies show how immersion in gender stereotypic situations influence women’s reactions to stereotypically masculine cues slowly over time, which likely involve conscious and deliberate processing of stereotypic cues. The present research complements prior research by examining whether the split-second activation of gender stereotypes impacts women’s reactions at an earlier stage of information processing, when conscious processing is not possible. To do so, we utilized neurophysiological measures (electroencephalogram or EEG) to shed light on automatic modulation of error responses during STEM-relevant performance tasks, well before one’s thoughts can be consciously articulated.

Previous research examining implicit gender stereotyping using EEG has mainly focused on how people process gender stereotypic and counterstereotypic language (e.g., Pesciarelli et al., 2019; Proverbio et al., 2018; White et al., 2009). These studies found that reading counterstereotypic language (e.g., the engineer stained her skirt) elicits greater neurological reactivity on specific event-related brain potentials (ERP; P300, N400, and P600) as compared to stereotypic language, suggesting a surprise response. The present research aims to address a different question which has not yet been examined—i.e., whether the subtle activation of gender stereotypes signaling male dominance in STEM would elicit increased neurological reactivity during a STEM-relevant task, especially when women make task-related errors that are stereotype-consistent. Specifically, we examined whether incidental exposure to a watching male face during a math intelligence task, a social cue signaling male dominance in STEM,
would modulate women’s attentional vigilance to errors captured by neurophysiological signals.

Previous studies suggest that the mere presence of watching eyes can modulate social behaviors and associated neural responses by automatically evoking a concern about potential negative social evaluations (Haley & Fessler, 2005; Hitokoto et al., 2016; Park & Kitayama, 2014; Rigdon et al., 2009). For example, Park and Kitayama (2014) presented an image of a watching face (or a control image) as a priming stimulus on some trials during a flanker task while monitoring participants’ brain activities using EEG. After participants made errors following the face (vs. control) primes, they showed increased attentional vigilance to their errors, indexed by the enhanced magnitude of error-related negativity (ERN), an ERP component of error processing (Gehring et al., 1993). Moreover, this face priming effect was only evident among East Asians, but not among European Americans, consistent with the view that the former group is more interdependent, and thus, more vigilant to errors following subtle social cues implying potential social evaluations.

**Overview of the Present Research**

Building on past evidence, the present research examined whether a watching male face would evoke attentional vigilance to errors among women invested in STEM, insofar as this image could be interpreted as the evaluative presence of a high-status person in the male-dominated STEM context, and thus, activate math-gender stereotypes in a math test-taking situation. Our first aim was to test this hypothesis using a modified paradigm from Park and Kitayama (2014), in which male and female college students performed an alleged math intelligence task while being exposed to an image of a
watching male face as a priming stimulus on some trials. We hypothesized that women would show greater attentional vigilance to their errors when these errors were preceded by male face primes (compared to scrambled faces as control primes).

The degree to which male face priming evoked attentional vigilance to errors was assessed with the ERN, a neural index of early, automatic detection of errors during a speeded reaction time task (e.g., Amodio et al., 2004; Botvinick et al., 2004; Carter et al., 1998; Danielmeier et al., 2009). The ERN is characterized by a negative deflection peaking 50 – 100 ms following error commission at fronto-central electrode sites. Prior research suggests that the ERN may serve as an index of attentional vigilance to errors made in stereotypic domains (Forbes et al., 2008; Schmader et al., 2008). For example, Forbes et al. (2008) found that when racial minority students made errors while performing an alleged intelligence task, this elicited large ERN responses especially among those who valued academics more, presumably due to their increased concern about negative societal stereotypes regarding their intelligence (e.g., Steele & Aronson, 1995). Applied to our experiment, we predicted that if the watchful gaze of a male face is sufficient to activate negative gender-math stereotypes on a randomized trial-by-trial basis and if making errors in this context is perceived as confirming these stereotypes, women should exhibit greater neurological vigilance to their errors, when these errors are preceded by male face (vs. control) primes. Our primary aim was to test this prediction by examining whether women would exhibit larger ERN after making errors on trials in which they were primed with male faces (vs. control images), in comparison to men.

Our second aim was to examine whether women who are especially invested in pursuing STEM careers would be particularly vigilant to errors on face (vs. control)
priming trials. Because proficiency in STEM is especially important to this subgroup of women, they may be attuned to gender stereotypic cues more than other women who are less invested in STEM careers and also men. This prediction is consistent with prior work, which found that women who are highly identified with math were more sensitive to gender stereotypes when taking a math test in comparison to low-math identified women and all men (Lesko & Corpus, 2006). We thus hypothesized that women who are highly invested in pursuing STEM careers would show a stronger face priming effect on ERN than women who are less invested in STEM careers and all men.

**Method**

**Participants**

One hundred and twenty-seven undergraduate students from the University of Massachusetts, Amherst (66 men, 61 women, \( M_{age} = 19.88, SD_{age} = 1.62 \)) participated in this study in exchange for course credit or $20. We recruited students who were either majoring in or interested in majoring in a STEM field through the human participant pool and through fliers posted around campus. In this sample, 57.5% were White, 17.3% Asian, 10.2% Black, 2.4% Hispanic, and 12.6% were multiracial or indicated other races/ethnicities. Our sample size was guided by previous studies that involved similar neurophysiological assessments (Kitayama & Park, 2014; Park & Kitayama, 2014). We sought to recruit a minimum of 25 participants per condition (four conditions based on participant gender and low or high investment in STEM careers), plus an additional 20% to guard against possible data attrition and to ensure that the study would be well-powered (Boudewyn et al., 2018; Button et al., 2013).

**Procedure and Materials**
All procedure and materials were approved by the Institutional Review Board at the University of Massachusetts, Amherst. Upon arrival in the lab, participants provided informed consent and were prepared for EEG recording. Participants then performed a numerical Stroop task that was framed as diagnostic of “math intelligence” to make it relevant to gender-math stereotypes (e.g., Blascovich et al., 2001; Martens et al., 2006). Specifically, participants were led to believe that their brain responses would be monitored during task performance to identify neurophysiological responses associated with math intelligence. In reality, we assessed their neural vigilance to errors, indexed by the ERN. The ERN emerges after the commission of errors in speeded choice reaction tasks that involve response conflict, such as Stroop tasks (e.g., Blascovich et al., 2001; Carter et al., 1998; Danielmeier et al., 2009). In particular, we chose the “numerical” version of the Stroop task to strengthen the believability of the cover story that this task was related to math intelligence. The numerical Stroop task has been used in previous studies as an alleged math task (Ashkenazi, 2018; Suárez-Pellicioni et al., 2013) and has been found to be effective in eliciting ERN in response to errors, especially among individuals who are high in math anxiety (Suárez-Pellicioni et al., 2013).

Following the paradigm used in Suárez-Pellicioni et al. (2013), participants were presented with a pair of numbers (i.e., 1–2, 1–8, 9–2, 9–8) on each trial and asked to judge which number was larger in numerical magnitude while ignoring its physical size. There were three trial types: congruent, incongruent, and neutral (see Figure 1-A for examples of each trial type). For congruent trials, the number of larger numerical magnitude within the pair was also larger in physical size, whereas for incongruent trials,
the number of larger numerical magnitude was smaller in physical size. For neutral trials, the numbers only differed in numerical magnitude but were the same physical size.

On each trial, participants were first presented with a visual prime for 90 ms. Adapting a procedure from prior research (Park & Kitayama, 2014; see also Hitokoto et al., 2016), the prime was either an image of a watching male face or a scrambled face (as a control prime), each presented randomly for half of the trials in each block (see Figure 1-B for sample images). The prime stimuli were borrowed from Park and Kitayama (2014), who created race-neutral, young adult male faces using FaceGen Modeller 3.3 (Singular Inversions Inc.) by morphing Caucasian (50%) and Asian (50%) faces—the two racial groups that are perceived to be superior in STEM fields relative to other racial groups (e.g., Canning et al., 2019). The morphed face images were then scrambled to create the scrambled faces. Participants were told that on each trial, an image would appear right before the presentation of the numbers, but that they should ignore these images because they were unrelated to the numbers that would follow. After the presentation of the prime, a fixation cross was presented for 500 ms, followed by a target pair of numbers, which remained on the screen for 100 ms. Participants were asked to press a designated key on the keyboard to indicate whether the number on the left (F) or right (J) was numerically larger, within a response window of 600 ms. The next trial began after an inter-trial interval of 600–800 ms (see Figure 1-C for a sample trial structure).

After completing one practice block of 24 trials, participants completed 16 blocks of 48 trials each. Prime type (face vs. scrambled face) and target type (congruent vs. incongruent vs. neutral) were varied within each block, resulting in six types of trials that
were each presented eight times per block. Trial order was randomized within block. At the end of each block, participants received feedback on their performance, adopted from Park and Kitayama (2014): participants were asked to respond faster in the next block if their accuracy in that block was higher than 90%, or asked to focus on improving accuracy if their accuracy was below 90%. Participants performed the task using a Cybertron TGM1114C PC with Windows 8.1. The stimuli were presented using a 24-inch ASUS VG248QE HD monitor (1920 × 1080) with a 144 Hz rapid refresh rate and a 1 ms response time.
Figure 1. (A) Examples of a neutral (both numbers are the same in size), congruent (the larger number in magnitude is also larger in size), or incongruent target (the larger number in magnitude is smaller in size). (B) Sample images of a male face and a scrambled face. (C) A sample trial structure.

After completing the numerical Stroop task, participants answered a question about how likely they were to pursue a profession in a field related to STEM, using a 7-point scale (1 = not at all likely, 7 = very likely), which served as our measure of
participants’ investment in STEM careers. In addition, they completed a series of questions assessing how engaging, interesting, boring, and difficult they found the task to be, and how satisfied they were with their task performance, using 7-point scales (1 = not at all, 7 = very much). See Supplemental Analyses for the results from additional measures included for exploratory purposes.

**EEG Recording and Processing**

EEG was recorded with 64 electrodes that were placed according to the extended International 10–20 System in a nylon cap and referenced to the left mastoid. The electro-oculogram (EOG) was also recorded from four additional channels, two placed at the outer canthi of both eyes and two placed above and below the right eye, respectively. Sodium chloride gel was added to each sensor to lower the impedance under 5kOhm.¹ EEG and EOG signals were amplified with a band-pass of DC to 100 Hz with actiCHamp amplifier (Brain Products GmbH, Germany) and were sampled with 1024 Hz. All data were re-referenced to the left and right mastoid off-line and were then resampled at 256 Hz. Response-locked ERPs were extracted 200 ms before and 800 ms after each trial response. The data were baseline-corrected at 200 to 100 ms pre-response voltage and then corrected for ocular artifacts (Gratton et al., 1983). A low-pass filter with a half-amplitude cutoff at 30 Hz was applied to remove high frequency noise, following an approach from prior work (e.g., Gehring & Willoughby, 2004; Nieuwenhuis et al., 2001; Themanson et al., 2014). The data were then inspected to remove trials containing artifacts exceeding ±100μV at three midline centered electrodes (Fz, FCz, and Cz). On average, 1.65% (SD = 6.07) of trials were removed and the percentage of removed trials

¹ Due to equipment malfunction, we were not able to check impedance levels for seven participants (two men and five women). After visual inspection of their data, we decided to include them in the final analysis. The results remained the same regardless of whether these participants were included or excluded.
did not differ by gender and/or prime type, ps ≥ .091. After artifact rejection, a minimum number of six error trials per condition was required to be included in the analysis on ERN (Olvet & Hajcak, 2009). On average, participants had approximately 42 error trials per condition (face primes: \(M = 41.99, SD = 20.82\); control primes: \(M = 41.78, SD = 19.18\)). The artifact-free epochs were then averaged separately based on response type (error vs. correct) and prime type (face vs. control).

Since the ERN peaked an average of 30 ms after incorrect responses across participants, it was quantified as the mean amplitude between 20 ms before and 80 ms after the incorrect response. Our analysis focused on ERN amplitudes from two channels—the fronto-central and central midline electrodes (FCz and Cz)—because these are sites at which ERN amplitudes tend to be largest. Following an approach used in prior research (Brazil et al., 2009), we examined the ERN from these two regions by including channel (FCz, Cz) as a within-participant predictor in addition to our two primary predictors: gender (men vs. women) as a between-participant predictor and prime type (face vs. control) as a within-participant predictor.

**Results**

**Data Attrition and Sensitivity Power Analysis**

We excluded two participants who failed to follow task instructions and focused on a sample of 125 participants (65 men, 60 women) for the analyses on self-reports and behavioral performance. For the analysis on ERN, nine additional participants were excluded because one participant did not make enough errors (i.e., a minimum of six trials per condition; Olvet & Hajcak, 2009) and the remaining eight participants had noisy EEG data, leaving a sample of 116 participants (61 men, 55 women). Participants who

---

2 The data from this study are publicly available in Open Science Framework (OSF) at osf.io/pwt3/.
were dropped from the ERN analysis did not differ from those retained on key variables, such as gender, age, and interest in pursuing a career in STEM, \( ps > .201 \).

A sensitivity power analysis using G*Power (Faul et al., 2009) revealed that our primary analysis on ERN (the Gender x Prime Type interaction) based on the final sample \((N = 116)\) is sufficient to detect a small effect \((d = 0.22; \alpha = .05, \text{power} = .80)\). An observed power analysis also revealed that we reached power > .99 for the ERN results.

**Self-Report Measures**

We first tested whether women and men differed in their interest in pursuing STEM careers. Since we had actively recruited participants interested in STEM, not surprisingly, the overall mean interest in pursuing careers in STEM was high \((M = 5.36 \text{ on the 7-point scale, } SD = 2.08)\) and the distribution of the scores was negatively skewed \((\text{skewness} = -1.03, se = 0.22)\). Due to the skewness, we tested whether the two gender groups differed in their indicated interest in STEM careers by conducting a non-parametric test, which is not constrained by normality violations (Blair & Higgins, 1985). We found that the two gender groups did not differ on this variable, \(U = 1723.50, p = .235\) (men: \(M = 5.20, SD = 2.08\); women: \(M = 5.53, SD = 2.09\)). Upon examining the data further, we found that approximately half of the sample reported that they were highly certain of pursuing a STEM career \((47.2\% \text{ gave a rating of 7 on the 7-point scale})\), while the other half reported less certainty \((53.6\% \text{ gave a rating of 6 or below}; \text{see histograms in Figure 2})\). When we examined the percentage of women and men who were highly certain in pursuing STEM careers, no significant gender difference emerged; the percentage of participants who chose 7 did not vary by gender \((\text{women: } 53.3\%, \text{men: } 41.5\%), \chi^2(1) = 1.74, p = .187\).
Figure 2. Histograms of the distribution of the STEM career question (i.e., how likely are you to pursue a professional career in a field related to STEM?) for (A) women and (B) men (1 = not at all likely, 7 = very likely).

Further, the questionnaire responses showed that men and women were no different in their experience during the task (see Table 1 for statistics). Specifically, the
two groups did not differ in their level of task engagement or satisfaction with their performance, $t_s \leq 1.82, ps \geq .068$. There were also no significant gender differences in how interesting, boring, and difficult they found the task, $t_s \leq 1.81, ps \geq .073$.

Table 1. *Descriptive Statistics of the Post-Task Questionnaire Variables*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Men ($n = 65$)</th>
<th>Women ($n = 60$)</th>
<th>$t(df)$, $p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement</td>
<td>$M = 5.11$, $SD = 1.53$</td>
<td>$M = 4.85$, $SD = 1.45$</td>
<td>$t(123) = 0.97, p = .337$</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>$M = 3.65$, $SD = 1.40$</td>
<td>$M = 3.25$, $SD = 0.99$</td>
<td>$t(115.25) = 1.84, p = .068^1$</td>
</tr>
<tr>
<td>Interesting</td>
<td>$M = 3.98$, $SD = 1.91$</td>
<td>$M = 3.47$, $SD = 1.64$</td>
<td>$t(122) = 1.59, p = .115$</td>
</tr>
<tr>
<td>Boring</td>
<td>$M = 4.48$, $SD = 1.77$</td>
<td>$M = 4.78$, $SD = 1.44$</td>
<td>$t(120.61) = -1.05, p = .296^1$</td>
</tr>
<tr>
<td>Difficult</td>
<td>$M = 4.09$, $SD = 1.40$</td>
<td>$M = 3.62$, $SD = 1.54$</td>
<td>$t(123) = 1.81, p = .073$</td>
</tr>
</tbody>
</table>

$^1$ Levene’s test for equality of variances was violated for this specific analysis; thus, the statistic for unequal variances is reported.

**Behavioral Performance**

Next, we tested whether face priming modulated participants’ behavioral performance. We first conducted a 2 Gender (men vs. women) x 2 Prime Type (face vs. control) mixed analysis of variance (ANOVA) on accuracy, with gender as a between-participant factor and prime type as a within-participant factor. Neither the main effects of gender and prime type nor the interaction effect between the two were statistically significant, $F_s(1, 123) \leq 0.71, ps \geq .350$, suggesting that the two gender groups did not differ in their accuracy both on face priming trials (women: $M = 87.49\%$, $SE = 0.98$; men:
We next conducted a 2 Gender (men vs. women) x 2 Prime Type (face vs. control) x 2 Response Accuracy (correct vs. error) mixed ANOVA on response time with gender as a between-participant factor and prime type and response accuracy as within-participant factors. Only the main effect of response accuracy was statistically significant, $F(1, 123) = 458.71, p < .001, d = 3.87$. Participants were significantly faster on error trials ($M = 189.82$ms, $SE = 3.92$) than on correct trials ($M = 227.13$ms, $SE = 3.85$). All other effects were non-significant, $F_s(1, 123) \leq 1.65, ps \geq .202$.

**Aim 1. Gender Differences in the Face Priming Effect on Error-Related Negativity**

Our primary aim was to examine whether women and men differ in their neural vigilance to face priming. We tested our prediction by conducting a 2 Gender (men vs. women) x 2 Prime Type (face vs. control) x 2 Channel (FCz vs. Cz) mixed ANOVA with gender as a between-participant factor and prime type and channel as within-participant factors. There was a significant main effect of gender, $F(1, 114) = 5.14, p = .025, d = 0.42$. Overall, women exhibited larger ERN ($M = 1.23\mu V, SE = 0.63$) than men ($M = 3.21\mu V, SE = 0.60$), regardless of prime type (note that because ERN is a negative deflection from baseline, mean amplitudes that are less positive correspond to larger ERN). The main effect of channel was also significant, $F(1, 114) = 15.98, p < .001, d = 0.75$, indicating that the overall ERN amplitudes were greater at FCz ($M = 1.78\mu V, SE = 0.47$) than at Cz ($M = 2.66\mu V, SE = 0.44$). The main effect of prime type was not

---

3 Previous studies suggest that after making errors, people often engage in compensatory behaviors in an attempt to improve their performance on subsequent trials, which emerge as increased post-error accuracy (e.g., Amodio et al., 2004; Hajcak, McDonald, & Simons, 2003). In an exploratory analysis, we examined whether face priming modulated post-error accuracy differently between the two gender groups. The results from this analysis are reported in Supplemental Analyses.
significant, $F(1, 114) = 1.76, p = .187$. The two-way interactions between gender and channel and between prime type and channel were also not significant, $F(1, 114) = 1.80, p = .182$ and $F(1, 114) = 0.04, p = .842$, respectively.

Importantly, as predicted, we found a significant Gender x Prime Type interaction effect, $F(1, 114) = 4.65, p = .033, d = 0.40$. Women displayed significantly larger ERN after being primed with male faces ($M = 0.89\mu V, SE = 0.66$) than control images ($M = 1.57\mu V, SE = 0.64$), $F(1, 114) = 5.76, p = .018, d = 0.45$ (see Figures 3-A and 3-B for waveforms and topographic maps, respectively). In contrast, men’s ERN did not vary as a function of prime type, $F(1, 114) = 0.69, p = .549$ (face primes: $M = 3.13\mu V, SE = 0.61$; control primes: $M = 3.29\mu V, SE = 0.62$) (see Figures 4-A and 4-B). Furthermore, the three-way interaction between gender, prime type, and channel was not significant, $F(1, 114) = 0.23, p = .636$, as we found a significant Gender x Prime Type interaction at each electrode site separately (see Supplemental Analyses for the results from the separate analyses by channel).

\footnote{See Supplemental Analyses for an exploratory analysis we conducted to examine the effects of face priming on error positivity (Pe), an ERP component linked to conscious error awareness or an emotional reaction to an error (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000).}
Figure 3. (A) Grand averaged event-related brain potentials (ERPs) elicited by errors (black lines) and correct responses (gray lines) as a function of the prime type (face vs. control) at FCz and Cz electrodes for women. (B) Topographic maps representing the scalp distribution of the ERN in the time range of -20 ms to 82 ms for the face priming condition, control priming condition, and the difference between the two conditions (face – control). The front of the head is at the top of the maps.
Figure 4. (A) Grand averaged event-related brain potentials (ERPs) elicited by errors (black lines) and correct responses (gray lines) as a function of the prime type (face vs. control) at FCz and Cz electrodes for men. (B) Topographic maps representing the scalp distribution of the ERN in the time range of -20 ms to 82 ms for the face priming condition, control priming condition, and the difference between the two conditions (face – control). The front of the head is at the top of the maps.
Aim 2. The Moderating Effect of Interest in Pursuing a STEM Career

Our second aim was to test whether the gender difference we observed above was more pronounced for women who are highly interested in pursuing STEM careers. As noted above, approximately half of our women participants indicated that they were highly certain that they would pursue a STEM career (see Figure 2). Due to the highly skewed distribution of participants’ responses on this variable, we computed a median split based on the total sample to categorize participants into two groups: those who were very certain that they would pursue a STEM career (i.e., who chose 7 on the 7-point scale) and others who were somewhat less certain that they would pursue a STEM career (i.e., who chose 6 or below on the 7-point scale). This median split resulted in four groups: women who were highly certain about pursuing STEM careers ($n = 31$) vs. less so ($n = 24$) and men who were highly certain about pursuing STEM careers ($n = 24$) vs. less so ($n = 37$).

We had an *a priori* prediction that women highly certain about pursuing STEM careers would be the most vigilant to errors following the face (vs. control) primes compared to everybody else. This prediction was motivated by prior research on stereotype threat suggesting that the effects of negative gender stereotypes in STEM on the self (which largely affect women but not men) are especially stronger for women who are highly identified with the STEM domain (e.g., Lesko & Corpus, 2006; Steele, 1997; Steinberg et al., 2012). We expected that the other three groups (i.e., women who are less interested in pursuing a STEM career and all men) would be less impacted by gender stereotypes, due to their lack of domain identification (i.e., women who are less invested in STEM) or to their superior status in the stereotyped domain (i.e., men regardless of
their likelihood to pursue a STEM career). Consistent with this prediction, these three
groups did not differ from each other in their ERN responses to face (vs. control) primes,
$F_{s}(1, 112) \leq 1.24, ps \geq .268$. We thus conducted a planned contrast comparing women
who were highly certain about pursuing STEM careers (+3) to the remaining three groups
(-1, -1, -1).

To capture the degree to which participants were vigilant to errors made in the
context of a watching male face vs. errors made in the context of control primes, we
computed a difference score by subtracting ERN amplitudes for the face priming trials
from those for the control priming trials. Because the effects of face priming were similar
for both FCz and Cz, we averaged the scores from these two electrode sites to compute
this index of differential vigilance to face (vs. control) priming (see Supplemental
Analyses for the analyses separated by channel). The positive scores on this index, which
we refer to as the face priming effect, indicate greater vigilance to errors made in the
context of face (vs. control) priming.

The planned contrast on this dependent variable was statistically significant, $F(1, 112) = 4.57, p = .035, d = 0.40$. As displayed in Figure 5, women who were highly
invested in pursuing STEM careers showed a greater face priming effect compared to the
other three groups. This indicates that this subgroup of women displayed significantly
larger ERN amplitudes in response to face priming ($M = 0.76\mu V, SE = 0.88$) compared to
control priming ($M = 1.69\mu V, SE = 0.86$), $F(1, 112) = 6.11, p = .015, d = 0.46$. In
contrast, the other three groups’ ERN responses were of similar magnitude, regardless of
prime type, $F_{s}(1, 112) \leq 0.65, ps \geq .422$.\(^5\)

\(^5\) We also tested each of the three groups separately as a single comparison group contrasted with women
highly interested in pursuing a STEM career. Consistent with the main analysis, this subgroup of women
Figure 5. The face priming effect on ERN separated by gender and interest in pursuing STEM careers. This index was computed by subtracting the ERN amplitudes on the face priming trials from the ERN amplitudes on the control priming trials, such that positive values indicate larger ERN on the face (vs. control) priming trials.

Discussion

Two research questions guided our investigation. First, during stereotype-relevant achievement tasks, do women show enhanced vigilance to errors following minimalistic cues signaling male dominance even when these cues appear for a split-second? Second, are women who are highly invested in pursuing STEM careers more attuned to these stereotype-relevant situational cues than their peers?

We found support for our primary hypothesis that minimalistic stereotypic cues such as the mere presence of male watching eyes were sufficient to increase neural vigilance to errors, indexed by enhanced ERN. Notably, this effect was pronounced showed a greater face priming effect than men who were highly interested in pursuing a STEM career, $F(1, 112) = 4.36, p = .029, d = 0.42$. A similar but weaker pattern of group differences emerged when these women were compared with men who reported low interest in pursuing a STEM career, $F(1, 112) = 3.73, p = .056, d = 0.36$, and with women who reported low interest in pursuing a STEM career, $F(1, 112) = 1.07, p = .304, d = 0.19$. 

33
among women but not men, consistent with prior work indicating that men are less attuned to situational cues that prime gender stereotypes in STEM contexts (e.g., Murphy et al., 2007; Spencer et al., 1999). While previous work utilized relatively more explicit situational cues to trigger stereotype activation (e.g., Cheryan et al., 2009; Murphy et al., 2007), our work shows that very subtle visual cues signaling male dominance in the context of STEM are sufficient to result in cognitive modulation of neural responses to errors on a trial-by-trial basis in a matter of milliseconds, thereby suggesting that stereotype activation results in automatic regulation of attentional responses at an early stage of informational processing, well before deliberate and conscious processes can be engaged.

One might argue that women’s enhanced attentional vigilance to errors following male face priming may have been driven by exposure to outgroup faces, rather than male faces in particular. That is, individuals may feel more threatened when evaluated by any outgroup member than an ingroup member. However, this explanation cannot account for why only a subgroup of women—those highly invested in pursuing STEM careers—exhibited enhanced attentional vigilance. If the alternative explanation were valid, one would expect that all women in our sample should show this effect, not just a subgroup of women. Thus, it seems more probable that greater vigilance was evoked among a subgroup of women who found it particularly threatening to make mistakes in a domain that is important to their self-concept. This finding is consistent with previous research showing that members of stigmatized groups who chronically anticipate being a target of stereotypes, such as those who are strongly identified with their social group (Schmader, 2002) or with the stigmatized domain (Spencer et al., 1999), are more vulnerable to social
identity threat (Townsend et al., 2011), likely due to their heightened alertness to identity threat cues in their environments (Kaiser et al., 2006).

We theorized that male face priming should evoke attentional vigilance among women in STEM, insofar as it is interpreted as signaling the evaluative presence of a dominant, high-status individual, whose group is positively regarded in that domain. That said, it remains unclear if our participants perceived the faces as higher in status or dominance in the context of STEM, since we did not measure perceived male dominance or status. Future research should directly address this issue by experimentally manipulating features of primed faces to signal high (vs. low) status, such as head posture and eye gaze. For example, faces with a direct eye gaze and an upward head tilt are perceived as more dominant whereas faces with an averted eye gaze and a downward head tilt are perceived as more submissive (Mignault & Chaudhuri, 2003; Rule et al., 2012). Future work should examine whether women will show enhanced attentional vigilance to errors specifically when primed with more dominant (rather than submissive) male faces or whether any male faces will elicit vigilance.

Future research should also address whether the degree to which women are vigilant to errors is reduced when primed with female faces. Previous work shows that women feel more accepted and show better performance when surrounded by same-sex peers (Griffith, 2010). In particular, exposure to female role models is shown to be particularly beneficial for women in STEM fields (Dasgupta, 2011; Drury et al., 2011). For example, according to the Stereotype Inoculation Model (Dasgupta, 2011; Dennehy & Dasgupta, 2017; Stout et al., 2011), the presence of women role models leads women to feel less anxiety and more belonging in STEM by protecting against negative gender
stereotypes. Thus, we expect that female faces may serve as a cue signaling identity safety (rather than identity threat) to women in STEM, thereby decreasing their attentional vigilance to errors made in the context of gender-stereotyped tasks.

Beyond STEM contexts, it would also be important to examine the effects of female watching eyes in achievement domains in which women are stereotypically perceived to be superior to men (e.g., verbal skills). Previous work shows that when women’s gender identity in female-dominant domains is made salient, they perform better on tasks that favor their group (Shih et al., 2006), while men underperform when reminded of stereotypes that do not favor their group (Koenig & Eagly, 2005). Building on these results, we anticipate that situational reminders of female superiority, in the form of female face priming on a different task, may evoke attentional vigilance to errors among men, but not among women. This is another avenue of future research.

Despite significant gender differences in neural responses to face priming, there was no corresponding effect on behavioral performance. We speculate that these null results were driven by the fact that the task was relatively easy (i.e., 88% accuracy for both women and men), consistent with prior evidence suggesting that women and men perform equally on easy math tasks but women underperform on difficult math tests (e.g., Ben-Zeev et al., 2005; Keller, 2007; O’Brien & Crandall, 2003). Future research should examine whether face priming modulates women’s performance differently if the task was more difficult. Another important future extension would be to test whether women’s increased attentional vigilance has any implications for their subsequent levels of motivation, persistence, and performance in other STEM-related tasks. On the one hand, previous research suggests that women in STEM contexts may experience decreased
working memory capacity when under stereotype threat, which in turn, leads to impaired performance on following tasks (Beilock et al., 2007). Alternatively, other studies have found that increased ERN is often associated with better cognitive control (e.g., Amodio et al., 2004; Hajcak et al., 2003), implying that attentional vigilance may increase motivation to improve performance. Future research should test these competing predictions by examining whether heightened attentional vigilance is helpful or detrimental for women’s performance on a subsequent, STEM-related task.

In conclusion, the current findings illustrate how exquisitely attuned and nimble the mind is in selectively attending to stereotypic cues in one moment and then turning attention away quickly when the cue no longer exists. We show that such strategic allocation of attention occurs swiftly and automatically, when individuals who are deeply invested in a performance domain encounter minimalistic cues signaling the gaze of a high-status member in a stereotyped domain.
Supplemental Analyses

Separate Analyses by Channel

Aim 1. Gender Differences in the Face Priming Effect on the Error-Related Negativity

FCz. A 2 Gender (men vs. women) x 2 Prime Type (face vs. control) mixed ANOVA was conducted on ERN at FCz. Mirroring the analysis we reported in the main document based on both channels combined, there was a significant main effect of gender, $F(1, 114) = 5.99$, $p = .016$, $d = 0.46$, such that women exhibited larger overall ERN amplitudes ($M = 0.64\mu V$, $SE = 0.68$) than men ($M = 2.92\mu V$, $SE = 0.64$). There was no significant main effect of prime type, $F(1, 114) = 1.60$, $p = .208$. As predicted, however, there was a significant Gender x Prime Type interaction, $F(1, 114) = 4.40$, $p = .038$, $d = 0.39$, such that women displayed greater ERN after being primed with male faces ($M = 0.29\mu V$, $SE = 0.70$) than with control primes ($M = 1.00\mu V$, $SE = 0.68$), $F(1, 114) = 5.38$, $p = .022$, $d = 0.43$. In contrast, men showed similar magnitudes of ERN following the face primes ($M = 3.01\mu V$, $SE = 0.67$) and control primes ($M = 2.83\mu V$, $SE = 0.65$), $F(1, 114) = 0.36$, $p = .547$.

Cz. A 2 Gender (men vs. women) x 2 Prime Type (face vs. control) mixed ANOVA for ERN at Cz was also conducted. The main effect of gender was marginal, $F(1, 114) = 3.75$, $p = .055$, $d = 0.36$, indicating that women tended to exhibit a slightly larger size of ERN ($M = 1.81\mu V$, $SE = 0.63$) than men ($M = 3.50\mu V$, $SE = 0.60$). There was no significant main effect of prime type, $F(1, 114) = 1.65$, $p = .202$. Furthermore, there was a significant Gender x Prime Type interaction, $F(1, 114) = 4.16$, $p = .044$, $d = 0.38$, such that women displayed greater ERN following the face primes ($M = 1.50\mu V$, $SE = 0.65$) than the control primes ($M = 2.13\mu V$, $SE = 0.64$), $F(1, 114) = 5.24$, $p = .024$, $d = \ldots$
0.42. Men exhibited similar ERN following the face primes ($M = 3.58\mu V, SE = 0.62$) and control primes ($M = 3.43\mu V, SE = 0.61$), $F(1, 114) = 0.30, p = .584$.

**Aim 2. The Moderating Effect of Interest in Pursuing a STEM Career**

**FCz.** We conducted a planned contrast comparing women who were highly invested in pursuing a STEM career with the remaining three groups on the face priming effect computed based on ERN at FCz. This analysis resulted in a marginally significant result, $F(1, 112) = 3.70, p = .057, d = 0.36$. As shown in the main analysis, women who were highly invested in pursuing STEM careers exhibited greater ERN amplitudes following the face primes ($M = 0.08\mu V, SE = 0.95$) than the control primes ($M = 1.02\mu V, SE = 0.92$), $F(1, 112) = 5.23, p = .024, d = 0.42$. In contrast, all other groups did not differ in their ERN amplitudes on face vs. control primed-trials, $Fs \leq 0.80, ps \geq .373$.

**Cz.** When we conducted the same planned contrast on the face priming effect on ERN at Cz, this yielded a significant result, $F(1, 112) = 4.81, p = .030, d = 0.41$. Women who were highly invested in pursuing STEM careers displayed greater ERN following the face primes ($M = 1.45\mu V, SE = 0.88$) than the control primes ($M = 2.36\mu V, SE = 0.86$), $F(1, 112) = 6.17, p = .014, d = 0.46$. All other groups did not significantly differ in their ERN amplitudes between primes, $Fs \leq 0.82, ps \geq .367$.

**The Face Priming Effect on Post-Error Accuracy**

Previous studies suggest that after making errors, people often engage in compensatory behaviors in an attempt to improve their performance on subsequent trials, which emerge as increased post-error accuracy (e.g., Amodio et al., 2004; Hajcak et al., 2003). This regulatory behavior is known to reflect enhanced executive control motivated by the desire to improve performance after error responses. We examined whether face
priming modulated post-error accuracy differently between the two gender groups. Post-error accuracy was computed as the total percentage of correct responses on trials immediately following an incorrect response, separately for each prime type. These variables were then submitted to a 2 Gender (men vs. women) x 2 Prime Type (face vs. control) mixed ANOVA. There were no significant main effects of either prime type, $F(1, 123) = 0.54, p = .463$, or gender, $F(1, 123) = 0.17, p = .683$. The interaction between gender and prime type was marginal, $F(1, 123) = 3.22, p = .075, d = 0.32$, as women tended to be more accurate on trials following face-primed errors ($M = 87.76\%, SE = 1.30$) than on trials following control-primed errors ($M = 86.34\%, SE = 1.35$), $F(1, 123) = 3.08, p = .082, d = 0.31$. In contrast, there was no such tendency among men, $F(1, 123) = 0.58, p = .447$ (face priming trials: $M = 87.47\%, SE = 1.25$; control priming trials: $M = 88.06\%, SE = 1.30$).

To delve further into the possible source of this gender effect, we tested whether women who were more invested in pursuing a STEM career would be more accurate on trials following face-primed errors rather than control-primed errors in comparison to women who were less invested in pursuing STEM careers and all men. We calculated the effect of face priming on post-error accuracy by subtracting the accuracy on trials following control-primed errors from the accuracy on trials following face-primed errors. We then conducted a planned contrast comparing the face priming effect on post-error accuracy between women highly interested in pursuing STEM careers with all men and women less interested in pursuing STEM careers. This contrast was significant, $F(1, 121) = 3.96, p = .049, d = 0.36$. Women who were highly certain about pursuing STEM careers showed higher accuracy on subsequent trials following errors made on the face
priming trials ($M = 88.46\%, SE = 1.77$) than on the control priming trials ($M = 86.10\%, SE = 1.86$), $F(1, 121) = 4.53$, $p = .035$, $d = 0.38$. The other three groups did not differ in their degree of post-error accuracy as a function of prime type, $Fs \leq 1.23$, $ps \geq .269$. This finding suggests that women who were highly certain about pursuing STEM careers may have been more motivated to improve their performance after making errors following the face (vs. control) primes.

**Exploratory Measures**

We included several measures for filler or exploratory purposes throughout the study. The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) was administered before and after the numerical Stroop task to assess participants’ affective states over time. In addition, we administered two types of exploratory measures: (a) individual difference variables that prior research has found to modulate identity threat effects in STEM contexts (Exploratory Measures I) and (b) personality variables that are shown to correlate with the size of the ERN in previous research (Exploratory Measures II).

**Exploratory Measures I**

First, as potential moderators of identity threat effects, we included (a) the Math Identification Questionnaire (Ben-Zeev et al., 2011), (b) the Abbreviated Math Anxiety Scale (Hopko et al., 2003), and (c) the Implicit Theories of Intelligence Scale (Dweck, 1999), adopted for math intelligence. (d) Participants also indicated their awareness of gender-related math stereotypes (e.g., When other people think of individuals who are very good at mathematics, who comes to their mind?; 1 = mostly women, 3 = men and women equally, 5 = mostly men) as well as their personal endorsement of these
stereotypes (e.g., When you think of individuals who are very good at mathematics, who comes to your mind?).

We tested how these variables were correlated with the effect of face priming (i.e., difference in the ERN between the face and control priming trials, averaged across FCz and Cz) for each gender group. None of these variables predicted the face priming effect for both males and females, $rs \leq |.14|$, $ps \geq .326$ (see Table 2-A). These null results, especially from math identification, might seem at odds from previous findings suggesting that one’s identification with the stereotyped domain often moderates identity threat effects (e.g., Lesko & Corpus, 2006; Steele, 1997; Steinberg et al., 2012). Perhaps our correlation was non-significant because we measured participants’ math identification rather than their identification with specific STEM fields that correspond with their major (e.g., physics, chemistry, biology, neuroscience, etc.). It is possible that women who are interested in pursuing a science field may not always be narrowly math-identified.

**Exploratory Measures II**

Participants also completed (a) the Self-Consciousness Scale (Fenigstein et al., 1975), (b) the Self-Construal Scale (Singelis, 1994), (c) the Behavioral Inhibition System and Behavioral Activation System Scales (BIS/BAS; Carver & White, 1994), and (d) the Ten Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003), as these variables have been shown to correlate with the size of the ERN in prior work (e.g., Boksem et al., 2006; Pailing & Segalowitz, 2004; Park & Kitayama, 2014). We tested whether these variables correlated with the magnitude of the face priming effect. As shown in Table 2-B, agreeableness was correlated with a greater face priming effect for
females, but not for males. However, the reliability for this construct was very low ($\alpha = .14$), and thus, we dismissed this result given that its robustness is questionable. In addition, independent self-construal also moderated the gender difference in the face priming effect, such that it predicted an increased face priming effect for females, but not for males. Notably though, our results remained when the effect of independent self-construal is adjusted in a covariate analysis.

Table 2. 
Correlations Between Exploratory Measures and the Face Priming Effect on ERN

| A. Exploratory Measures I | Reliability | Male | | | Female | | |
|---------------------------|-------------|------|------------------|------------------|
| Math Identification | .72 | .00 | .997 | .09 | .500 |
| Math Anxiety | .87 | .01 | .925 | -.14 | .326 |
| Personal Endorsement of Stereotypes | .60 | -.12 | .361 | -.14 | .324 |
| Perceptions of Others’ Stereotypes | .81 | -.14 | .301 | -.11 | .432 |
| Theories of Math Intelligence | .93 | -.10 | .464 | .12 | .400 |

| B. Exploratory Measures II | Reliability | Male | | | Female | | |
|---------------------------|-------------|------|------------------|------------------|
| Private Self-Consciousness | .64 | .05 | .733 | .13 | .344 |
| Public Self-Consciousness | .76 | .20 | .124 | -.01 | .950 |
| Social Anxiety | .79 | .23 | .074 | -.05 | .736 |
| Independence | .71 | -.17 | .198 | .34 | .011 |
| Interdependence | .61 | -.01 | .921 | -.07 | .590 |
| Extraversion | .77 | -.23 | .069 | .07 | .609 |
| Agreeableness | .14 | -.01 | .967 | .29 | .032 |
| Conscientiousness | .57 | -.01 | .945 | .05 | .717 |
| Emotional Stability | .63 | -.02 | .882 | .12 | .404 |
| Openness to Experience | .44 | .08 | .543 | .02 | .877 |
| Behavioral Activation System | .81 | -.02 | .908 | .22 | .100 |
| Behavioral Inhibition System | .76 | -.02 | .899 | -.05 | .744 |

The Face Priming Effect on Error Positivity
In an exploratory analysis, we examined whether face priming modulated error positivity (Pe), differently for women and men. Pe is a parietal positivity, occurring between 200-400ms after error onset (Falkenstein et al., 1991, Overbeek et al., 2005), which has been suggested to reflect conscious error awareness or to be an emotional reaction to an error (Falkenstein et al., 2000). As shown in Figure 6, Pe peaked at Pz approximately 250ms following error onset for our participants. We thus quantified it as the mean amplitude between 150ms and 350ms after the error responses. This variable was then submitted to a 2 Gender (male vs. female) x 2 Prime Type (face vs. control) ANOVA, which yielded a marginal main effect of prime type, $F(1, 114) = 2.97, p = .088, d = 0.32$. Participants tended to show greater Pe after making errors on the control-primed trials ($M = 8.86, SE = 0.52$) than on the face-primed trials ($M = 8.58, SE = 0.53$). The main effect of gender was not significant, $F(1, 114) = 0.95, p = .331$. Importantly, however, there was a significant Gender x Prime Type interaction, $F(1, 114) = 5.92, p = .017, d = 0.45$. As Figure 6-A displays, men showed a greater magnitude of Pe after making errors on the control-primed trials ($M = 9.56, SE = 0.72$) than on the face-primed trials ($M = 8.89, SE = 0.73$), $F(1, 114) = 9.11, p = .003, d = 0.56$. In contrast and also shown in Figure 6-B, there was no significant effect of prime type among women, $F(1, 114) = 0.24, p = .626$ (face primes: $M = 8.27, SE = 0.77$; control primes: $M = 8.16, SE = 0.75$).

These results may suggest that men show decreased conscious recognition or reduced emotional reactions to errors on a math test when their high status is affirmed through exposure to male faces. Consistent with this interpretation, a meta-analysis on stereotype lift (Walton & Cohen, 2003) suggests that because gender stereotypes in
STEM favor men, thereby reinforcing their higher status in this domain, men sometimes perform better when this stereotype is made explicit. Future research is necessary to examine this possibility further.
Figure 6. Grand averaged event-related brain potentials (ERPs) elicited by errors (black lines) and correct responses (gray lines) as a function of the prime type (face vs. control) at Pz for (A) men and (B) women.

Transition to Chapter 3

The situations and environments that surround us directly impact how we perceive others and ourselves. For example, we learn stereotypes by observing society. These observations teach us that certain social groups (i.e., men) are expected to fulfill professional and technical roles and excel at them, while other groups (i.e., women) are expected to hold domestic and communal roles and excel at them (Banaji & Hardin, 1996; Blair & Banaji, 1996). When these normative expectations are violated, for example, when a woman pursues a technical career, her abilities may be called into question, and she may be seen as not smart enough for that role. This, in turn, may make her more vigilant to situations in which her abilities may be doubted. Chapter 2 provides evidence for this phenomenon and shows that this vigilance is expressed especially by women who are highly invested in pursuing STEM.

Just as situations signal to minorities that they are unwelcome in STEM fields, alternative situations can be designed to make STEM environments more welcoming for them. Specifically, greater representation of ingroup members who are also negatively stereotyped in STEM fields in a given academic context ought to signal greater attention to equity and inclusion and render that environment more welcoming to minority individuals. Based on the Stereotype Inoculation Model (Dasgupta, 2011), the presence of these ingroup members may serve as protection for minorities from the negative effects of stereotypes. Applied to gender, women may feel less vigilant and more
confident in their abilities in STEM if they feel supported by other fellow students who are also women. Chapter 3 examines this issue. In a longitudinal field experiment, I investigate the benefits of female peer mentorship for women college students in engineering.
CHAPTER 3

PEER MENTORING FOR WOMEN IN ENGINEERING

Increasing student engagement, success, persistence, and career development in science, technology, engineering, and mathematics (STEM) is a priority in the United States, given the high demand for a skilled STEM workforce in the 21st century and a short supply of college graduates with skills and interest in STEM careers (President’s Council of Advisors on Science and Technology, 2012; Zaza et al., 2020). The shortage of skilled STEM workers, especially in engineering, is exacerbated by the underrepresentation of women and racial ethnic minorities (Ambrose, 2019; Xue & Larson, 2015). Gender and race disparities also exacerbate income inequality between women and men and between racial minority and White Americans, given lucrative salaries in STEM professions, compared to other professions (Allen et al., 2015; Phelan et al., 2010).

A growing body of research has used randomized controlled trials to test whether social psychological interventions can reduce group-based disparities in students’ academic outcomes. Many of these interventions train individuals to cognitively reappraise their experience, with an emphasis on teaching underrepresented individuals to mentally adapt to their environment (which we call mental reappraisal interventions). For example, self-affirmation, growth mindset, and emotion regulation interventions successfully reduced gender, race, and class disparities in academic performance (Cohen et al., 2009; Miyake et al., 2010; Paunesku et al., 2015; Rozek et al., 2019; Shapiro et al., 2013; Shnabel et al., 2013; Walton & Cohen, 2011; Walton et al., 2015; Yeager et al., 2016; 2019). Brief affirmation interventions in which students reflect on personally-held
values decreased gender and race gaps in academic performance in both lab and field settings (Cohen et al., 2009; Miyake et al., 2010; Shapiro et al., 2013; Shnabel et al., 2013; Walton & Cohen, 2011; Walton et al., 2015). Prompting students to cultivate a growth mindset (i.e., viewing intelligence or performance as malleable rather than fixed) improved grades for previously low-performing students (Paunesku et al., 2015; Yeager et al., 2016; 2019). Reappraising anxiety as positive rather than negative also decreased the achievement gap in grades between high and low-income students (Rozek et al., 2019). These interventions make important contributions pointing toward academic equity, but also have limitations. First, by focusing on mental reappraisal levers of change, these studies place the onus on underrepresented individuals to adjust to academic institutions by changing their mindset rather than placing the onus on academic institutions to change learning environments to fit the needs of diverse students. Second, several of these studies focus on grades and test performance as measures of interest but do not assess students’ subjective experiences in academic spaces (e.g., belonging, confidence, anxiety, motivation; Miyake et al., 2010; Paunesku et al., 2015; Shnabel et al., 2013; Walton et al., 2015; Yeager et al., 2016; 2019). Emphasis on underrepresented students’ subjective experiences is critical because they affect persistence and interest even in the absence of objective performance gaps (Dasgupta, 2011; Stout et al., 2011). Third, much research investigating mental reappraisal interventions did not target STEM education, instead focused on overall academic performance, and many of these studies were conducted with adolescents and not college students (Cohen et al., 2009; Miyake et al., 2010; Paunesku et al., 2015; Rozek et al., 2019; Shapiro et al., 2013; Shnabel et al., 2013; Walton & Cohen, 2011; Walton et al., 2015; Yeager et al., 2016; 2019).
Another body of research shines light on the impacts of increasing the visibility of successful own-group role models for underrepresented students in STEM. This research testing the power of role models in STEM is consistent with the Stereotype Inoculation Model, which posits that just as exposure to biomedical vaccines protect one’s body against noxious viruses, exposure to successful experts and peers from one’s own group protect one’s mind against noxious stereotypes that cast doubt on one’s ability in achievement contexts (Dasgupta, 2011). Research has found that reading about or interacting with successful own-group role models closed gender and race gaps by increasing test performance in lab experiments and class grades for women and racial ethnic minorities, while also bolstering their confidence, belonging, and STEM identification (Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013; Stout et al., 2011). Collectively, this work points to an evidence-based solution that promotes equity in STEM education, both in terms of objective (performance) and subjective metrics (experiences of belonging, confidence, motivation, etc.), thereby amplifying the importance of visible representation of underrepresented groups in STEM spaces. Yet this research also has limitations, as these studies typically assess how role models impact student outcomes at a single time-point or over a short period of time (e.g., up to two weeks after interacting with the role model), providing insufficient information about long-term impacts (Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013; Stout et al., 2011). Moreover, the majority of these studies were laboratory experiments conducted in artificial settings, not in naturally existing environments (Cheryan et al., 2011; 2013;
Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013), with a few notable exceptions (Herrmann et al., 2016; Stout et al., 2011).

The current research aims to use the theoretical framework of the Stereotype Inoculation Model (Dasgupta, 2011) to investigate the long-term impacts of an intervention that leverages peer relationships in STEM educational contexts. Using a randomized controlled trial, we examine whether and how mentoring relationships with peers in engineering in the first year of college influence women engineering students’ subjective and objective academic experiences from college entry through graduation. Our focus is on women in engineering because this group is a small minority in engineering degree programs (21%; NSF, 2019) and the workforce (13%; SWE, 2019), despite being 51% of the American population (United States Census Bureau, 2019). Consistent with the Stereotype Inoculation Model (Dasgupta, 2011) and prior research (Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013; Stout et al., 2011), we predicted that relationships with successful female peer mentors in the first year of college would act as “social vaccines” for young women entering engineering, helping them disregard negative gender stereotypes, allay anxiety, preserve motivation, and increase persistence and professional development in STEM. Importantly, we predicted these benefits would endure through graduation several years after the conclusion of peer mentorship.

The present work extends past findings in several ways. First, although mentorship is a popular intervention commonly discussed in academia, government, and industry (American Association for the Advancement of Science, 2017; Jarrett et al., 2012; Rabionet et al., 2009; Woetzel et al., 2015), many field-based mentorship
initiatives were not experiments, making it impossible to determine whether mentorship was the causal factor responsible for downstream outcomes (Eby et al., 2008). For instance, in studies where participants self-selected into mentorship programs, the success of mentorship interventions could be due to pre-existing individual differences between those who opted into mentoring programs and those who did not, rather than the mentorship program itself. Further, when mentors were chosen by mentees, the unique mentor-mentee relationship (rather than mentorship in general) might be responsible for downstream outcomes. To resolve these causal inference problems, we conducted a randomized controlled field experiment in which participants were unaware that the study was related to mentoring.

Second, our field experiment complements past research on role models. While in theory, role model interventions should create more supportive environments for underrepresented individuals, most prior role model experiments used artificial lab settings (e.g., reading about successful role models) and did not create authentic relationships (Cheryan et al., 2011; 2013; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013). In contrast, we instituted a year-long field intervention to investigate whether and how fostering authentic relationships with peer mentors would change women students’ lived experiences in engineering, scaffold their success during a critical transition period into college, and sustain that success after mentors are gone through the next 2-4 years in college. Previous interventions addressing STEM disparities assessed impacts at a single time-point or over a relatively short period of time (Cheryan et al., 2011; 2013; Herrmann et al., 2016; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Miyake et al., 2010; Rozek et al., 2019; Shapiro et
al., 2013; Stout et al., 2011; Ramsey et al., 2013), providing insufficient information about long-term impacts. Though some research has demonstrated longer-term benefits of academic interventions (i.e., 2-3 years), these studies were not specific to STEM (Cohen et al., 2009; Walton & Cohen, 2011). To the best of our knowledge, no research has used a randomized controlled longitudinal field experiment to test whether an intervention implemented at one point of time enhances STEM engagement, persistence, and career preparation several years later among young adults from underrepresented groups in STEM. Our research addresses this gap.

Finally, our research examines whether mentees would gain benefits from having a mentor who does not share their social identity, by testing whether women engineering students receive benefits from male mentorship. Previous research suggests that women in STEM can benefit from effective male mentors (Cheryan et al., 2011; 2013). However, because these studies were conducted in short-term experiments using experimenters posing as role models, it is unclear whether these findings will generalize to real-world academic settings. It is plausible that having male mentors may help women become more accustomed to working in majority-male spaces, which may then protect their feelings of belonging and confidence in a male-dominated field. Thus, we examined the long-term impacts of male mentorship for female students in engineering.

We conducted a field-based longitudinal randomized controlled experiment that examined women’s subjective and objective academic outcomes and compared the benefits of own-gender and other-gender peer mentors to a control condition. We recruited undergraduate women who were first-year and transfer students (N = 150) who planned to major in engineering at a large public university. Participants were randomly
assigned a female peer mentor \((N = 52)\), a male peer mentor \((N = 51)\), or no mentor \((N = 47)\) for 1 year. Peer mentors were junior or senior undergraduate students who volunteered to be mentors and were trained prior to being matched with mentees. Researchers described the study to mentors and mentees as aiming to identify barriers and opportunities experienced by college students in engineering. This generic description served two purposes. First, it ensured that participants who volunteered for this experiment were not opting in because of an interest in mentoring. Furthermore, it retained participant naiveté (e.g., mentors did not know that all mentees were women; neither mentors nor mentees knew that mentor gender was relevant to hypotheses).

Mentor-mentee dyads met several times (median = 4 meetings) during mentees’ first year in college, after which mentoring interactions ended. We surveyed mentees from entry to graduation, measuring their subjective experiences in engineering (anxiety, motivation, belonging, and confidence); participation in career development internship opportunities in engineering; retention in engineering and STEM majors; aspirations to pursue post-graduate degrees in engineering; and emotional well-being. The first survey was administered before mentor assignment, serving as a baseline. Subsequent surveys were administered at 3 time-points in the first year of the study and then once per year until college graduation. College transcripts were also obtained with student permission to assess participants’ choice of major and academic performance.

This investigation significantly expands an earlier article we published reporting the impact of this peer mentorship intervention during the first year of college when mentoring was active (Dennehy & Dasgupta, 2017). In our earlier article, having a female peer mentor produced substantial benefits for women students compared to having a male
mentor or no mentor during the first year of college. Whereas women with male mentors or no mentors reported significantly greater anxiety and relatively lower motivation, declining belonging and confidence in engineering, and declining aspirations to pursue post-graduate degrees in engineering, women assigned to female mentors showed no changes in anxiety and motivation and reported stable belonging, confidence, and aspirations in engineering.

Absent from the initial article was an investigation of the long-term effects of first-year mentorship. Did the impact of peer mentoring decay once mentorship concluded or did it endure? Do durable benefits of first-year mentoring emerge on subjective (psychological experiences) as well as objective indicators (success in career development opportunities, retention in engineering majors)? We address these questions by following the original sample of women through their entire college experience until Bachelor’s degree completion, while also assessing the impact of mentoring on students’ success in securing career development opportunities during college, their major at graduation, as well as their emotional well-being.

Method

Participants

One hundred fifty female students majoring in engineering at the University of Massachusetts Amherst participated in our study. Data collection took place over 8 years, from 2011-2019. Students were recruited during new student orientation for the 2011-12, 2012-13, 2013-14, and 2014-15 academic years. Most students were entering first-years (80%). The remainder were transfer students joining in their second year (20%). During their first year in the study, participants were randomly assigned to a female mentor (n =
52), male mentor ($n = 51$), or no mentor (control condition; $n = 47$). Mentors were undergraduate students in their junior or senior year in the same engineering major as their mentee (e.g., electrical engineering, mechanical engineering). At the time of the baseline survey, participants were 18.34 years of age, on average ($SD = 1.34$). The sample was 67.3% White, 17.3% Asian, 5.3% multiracial, 2.7% Black, 2.7% Hispanic, and 2% other ethnicity. Participants were paid $20 for the baseline survey, $30 for the second survey, and $35 for the third survey during their first year. Participants were paid $10 for each follow-up survey in subsequent years. Peer mentors were paid $100 for each mentee they had.

**Procedure**

*Mentees*

During their first year at the university, participating mentees completed a survey at 3 time-points: 1) a baseline survey in August or September prior to meeting with their mentors, 2) a mid-year survey in January or February, and 3) an end-of-year survey in April or May. In subsequent years, after the mentoring relationship had officially ended, we asked participants to complete a follow-up survey each year sometime during the spring semester (February to May) or summer. On average, participants completed a total of 4.27 surveys; 79% of participants (118 out of 150) completed 4 or more surveys and 21% completed 3 surveys (32 out of 150). The percentage of participants who completed at least 1 or more surveys after their first year did not significantly differ by mentoring condition ($\chi^2(2, N = 150) = 1.60, p = .450$). Participants were also asked to complete an additional survey at least one year after their graduation, of which 67% (100 out of 150)
completed. Completion for this post-graduation survey also did not differ by mentoring condition ($\chi^2(2, N = 150) = 0.14, p = .932$).

*Mentors*

We recruited 58 student mentors (32 females, 26 males) based on faculty recommendations. Mentors were primarily seniors, with some juniors, who were in good standing in their engineering major and played leadership roles in student-run professional clubs. Before being assigned a mentee, mentors attended a half-day training workshop at the beginning of the academic year. During the training, we emphasized two major points. First, mentors were asked to reflect on their first two years in engineering and asked to identify any difficulties they had and the support or information that they wished they had during that difficult time. Based on their experience, they also identified the factors that encouraged students to stay in engineering and difficulties that discourage students and may lead them to drop out of engineering. Mentors were encouraged to share the information they had generated as a group with their mentees and to be a source of support. This information was provided to all mentors in the form of a mentoring guide with topics of discussion to raise with their mentees (e.g., providing advice on academic coursework; providing tutoring help; helping mentees develop plans for college and careers including finding research assistantships and internships; providing social support; connecting mentees with other students and student clubs). Second, mentors were asked to meet with their mentees once a month throughout the academic year and engage with them on any topic in their mentoring guide based on mentee need. Mentors were told that their primary goal was to be their mentee’s friend and ease their transition into college. They were encouraged to meet over social activities. Finally, mentors were
provided the names of their mentees and asked to reach out to their mentees to initiate contact. They were asked to complete a brief online survey after each meeting to keep a record of the meeting, including a summary of the topics discussed.

**Measures**

*Psychological Experiences in Engineering*

**Anxiety and Motivation.** Ten items were used to measure participants’ experiences of anxiety and motivation in the context of their engineering courses. The five items that measured anxiety were: “My engineering related classes this year are likely to be difficult;” “I feel worried about my engineering related classes this year;” “I feel stressed about my engineering related classes this year;” “I feel unsure about my engineering related classes this year;” and “I feel anxious about my engineering related classes this year.” The five motivation items were: “I have the basic skills and abilities to be successful in my engineering related classes this year;” “I will be able to overcome any difficulties I experience in my engineering related classes this year;” “I have what it takes to deal with my engineering related classes;” “I am prepared to deal with my engineering related classes;” and “I feel confident about my engineering related classes this year.” Participants reported how much they agreed with each statement on a scale from 1 (not at all true) to 7 (very true). Responses to these items were averaged to form an index of anxiety (αs between 0.83-0.90, across years in the study) and motivation (αs between 0.79-0.95). These items were asked at three time-points during participants’ first-year in college and subsequently once a year until graduation.

In keeping with previous literature (Beltzer et al., 2014; Dennehy & Dasgupta, 2017; Harvey et al., 2010; Mendes et al., 2001; Moore et al., 2012; U.S. Department of
Education, 2016), we treated anxiety and motivation as a ratio, representing the degree to which participants’ feelings of anxiety about engineering was counteracted by their motivation to do well in engineering. This ratio is conceptualized based on Lazarus and Folkman’s (1984) theory of stress, which posits that the subjective experience of stress is based on the balance between demands and internal resources. Thus, the higher the demand (i.e., anxiety) relative to resources (i.e., motivation), the more stressed an individual would be.

**Confidence.** Participants completed two items that measured how confident they were in their talents in engineering. These items were: “Do you think you have a talent for engineering?” and “How confident do you feel about your engineering ability?” They answered on scales from 1 (not at all) to 7 (very much so). Responses were averaged to create an index of confidence in engineering (αs between 0.78 to 0.92). These items were asked at three time-points during participants’ first-year in college and subsequently once a year until graduation.

**Belonging.** Participants completed four items to measure how much they felt they belonged in engineering: “I feel connected to my peers in engineering;” “I feel accepted by my peers in engineering;” “I feel like an outsider among my peers in engineering” (reverse coded); and “I feel invisible among my peers in engineering” (reverse coded). Similar to the measures for anxiety and motivation, participants indicated how much they agreed with each statement from 1 (not at all true) to 7 (very true). Responses were averaged to create an index of belonging (αs between 0.72-0.88). These items were asked at three time-points during participants’ first-year in college and subsequently once a year until graduation.
Academic Choices in Engineering

Engineering Internships. At the end of the first year and in each subsequent survey in college, participants were asked to report whether they had an engineering internship in the past year. In the post-college survey, participants reported whether they had an engineering internship during their final year of college. Participants’ responses were compiled to create a dichotomous index of whether they had any engineering internships while in college (0 = no engineering internship, 1 = had an engineering internship).

Major at Graduation. Using participants’ undergraduate transcripts, we created two dichotomous variables. First, we coded whether they earned an engineering degree (0 = no engineering degree, 1 = earned an engineering degree). We also coded whether participants graduated with a STEM degree (0 = no STEM degree, 1 = earned a STEM degree). This variable was created by using the U.S. Department of Education’s (2016) definition of STEM which includes biological sciences, physical sciences, computer sciences, engineering, mathematics and statistics, and science technologies.

Post-Graduate Engineering Degree Aspirations. In each survey during college, we asked participants how much they wanted to pursue a Master’s degree or PhD in engineering on a 7-point scale (from 1 = Not at all to 7 = Very much). These items were asked at three time-points during participants’ first-year in college and subsequently once a year until graduation.

Emotional Health and Well-being

During the last time-point of their mentorship year (i.e., spring of their first year), we asked participants to self-report their mental health was over the past year (“In terms
of your overall mental health, how psychologically healthy and happy did you feel during this past academic year?) on a scale of 1 (not at all) to 7 (very healthy and happy). This item was repeated each year until college graduation.

Results

Data Management and Analysis

Results are described in three sections. First, we present results showing outcomes for which women in engineering experienced significantly stronger benefits of female peer mentorship than male peer mentorship or no mentorship. Second, we identify the psychological mechanism that helps explain the benefits of female mentorship. Third, we present two outcomes for which both types of mentors—female and male—showed benefits over having no mentor.

For continuous dependent variables measured at multiple time-points (anxiety and motivation, confidence, belonging, aspirations to pursue post-graduate degrees, and emotional well-being), we investigated how participants’ responses changed throughout college by coding the time-point at which each response was obtained. Time was centered at the beginning of year 1, before mentor assignment (baseline). Responses for subsequent time-points were scaled in reference to the month in which it was obtained relative to the baseline. The time variable was divided by 12, such that slope coefficients indicate yearly change.

We estimated multilevel models using Mplus 8 (Muthén & Muthén, 2017), utilizing the full information maximum likelihood estimator, which is the recommended practice to handle missing data (Allison, 2010; Preacher et al., 2008). Each statistical model had two levels. Level 1 represented variables measured at multiple time-points
across participants (anxiety and motivation, confidence, belonging, post-graduate aspirations, emotional well-being). Level 2 represented the mentor condition—the independent variable that varied between participants. Each model tested whether mentor condition influenced change over time on the dependent variable. In level 1 of the model, the dependent variable was regressed on time, creating a slope representing change over time for that variable. Random effects were created for the intercept (participants’ baseline responses at college entry) and slope. The intercept and slope were allowed to covary at level 2 to control for variance shared between the actual values on the dependent variable and its slope/change over time. At level 2, the intercept and slope of the dependent variable were both regressed on mentor condition. Given that mentor condition was a multicategorical variable, two dummy codes were created to index the female mentor condition (0 = male and no mentor conditions, 1 = female mentor), male mentor condition (0 = female and no mentor conditions, 1 = male mentor). We then tested the effect of mentor condition on longitudinal change over time on participants’ experiences in two ways: (a) by examining whether the trajectory of each mentor condition significantly changed from college entry to graduation; and (b) by testing if these change trajectories significantly differed between mentor conditions. Dichotomous variables (participation in engineering internships, college major at graduation) were analyzed using chi-square tests to examine percentage differences between mentor conditions.

Female Mentors Protect Women’s Subjective Experiences in Engineering, Objective Choices in Engineering, and Emotional Well-being

Anxiety and Motivation
As in previous studies (Beltzer et al., 2014; Dennehy & Dasgupta, 2017; Harvey et al., 2010; Mendes et al., 2001; Moore et al., 2012), anxiety and motivation in engineering was treated as a ratio score, signifying participants’ feelings of anxiety in engineering relative to their motivation. We compared how the anxiety to motivation ratio experienced by women students changed over time from college entry through graduation for students who had been assigned a female peer mentor in the first-year of college compared to the other two conditions. Whereas women without peer mentors ($B = 0.20$, $SE = 0.06$, $p < .001$) and women assigned male mentors ($B = 0.12$, $SE = 0.05$, $p = .014$) showed a significant increase in anxiety relative to motivation in engineering throughout college, women assigned female mentors showed no change ($B = 0.03$, $SE = 0.05$, $p = .571$). A contrast analysis revealed that the change trajectory for women with female mentors was significantly different from the change trajectories in the other two conditions ($B = 0.13$, $SE = 0.06$, $p = .028$), suggesting that female mentors protected women against anxiety and preserved their motivation. When comparing conditions separately, the change trajectories for women with female mentors versus no mentors were significantly different ($B = 0.17$, $SE = 0.07$, $p = .018$). The change trajectory for women with male mentors fell in-between the slopes of the other two conditions and did not significantly differ from either (no mentor condition: $B = 0.08$, $SE = 0.07$, $p = .247$; female mentor condition: $B = 0.09$, $SE = 0.07$, $p = .181$). See Figure 7.
Figure 7. Participants’ feelings of anxiety and motivation in engineering from college entry through graduation. To emphasize the difference in change trajectories between conditions over time for the anxiety to motivation ratio, the y-axis shows participants’ change over time relative to the baseline survey before mentors were assigned.

**Engineering Internships**

Whereas 61% of women without peer mentors and 65% of those with male mentors participated in engineering internships while in college, the success rate was substantially higher—82% for women with a female mentor (Figure 8). The difference between the female mentor condition and the other two conditions was significant ($\chi^2(1, N = 125) = 4.79, p = .029$), indicating that female peer mentorship in the first year of college increased women’s success in securing professional opportunities in engineering compared to male mentorship or no mentorship. When comparing conditions separately, the percentage of women with female mentors who had secured engineering internships was greater than the percentage in the no mentor condition ($\chi^2(1, N = 82) = 4.58, p = .032$). The percentage in the male mentor condition fell in-between the two other
conditions and did not significantly differ from the no mentor condition ($\chi^2(1, N = 81) = 0.18, p = .669$) or the female mentor condition ($\chi^2(1, N = 87) = 3.12, p = .077$).

![Bar chart showing percentage of participants who successfully acquired and completed engineering internships while in college.](image)

**Figure 8.** Percentage of participants who successfully acquired and completed engineering internships while in college.

**Major at Graduation**

We examined whether peer mentorship in the first year of college affected the proportion of students who graduated with an engineering major or any STEM major 2-4 years later (degrees in physical sciences, biological sciences, computer science, engineering, or mathematics, see Figure 9). Whereas 66% of those with no mentor and 71% of those assigned a male mentor earned Bachelor’s degrees in engineering, 79% of women with a female mentor earned engineering degrees. Although these data trended in the predicted direction, this difference was not statistically significant ($\chi^2(1, N = 150) = 1.85, p = .174$). However, for STEM majors, a significant difference in STEM degrees by mentor condition emerged: 92% of women assigned female mentors graduated with
STEM majors compared to 78% of women assigned male mentors and 81% without mentors ($\chi^2(1, N = 150) = 4.09, p = .043$). Although having female peer mentors in the first year of college did not significantly increase engineering degree completion, it did significantly increase STEM degree completion compared to the other two conditions. The percentage of women STEM graduates in the female mentor condition was greater than in the male mentor condition ($\chi^2(1, N = 103) = 3.99, p = .046$). The no mentor condition did not significantly differ from the female mentor condition ($\chi^2(1, N = 99) = 2.84, p = .092$) or the male mentor condition ($\chi^2(1, N = 98) = 0.09, p = .767$).
Figure 9. Percentage of participants who obtained Bachelor’s degrees in engineering (Panel A) and STEM (Panel B).

**Post-Graduate Engineering Degree Aspirations**

Whereas women without mentors \((B = -0.92, SE = 0.20, p < .001)\) and those assigned male mentors \((B = -0.60, SE = 0.18, p = .001)\) showed declining interest in pursuing post-graduate degrees in engineering throughout college, those assigned female mentors held steadfast in their intentions to pursue post-graduate degrees in engineering \((B = -0.22, SE = 0.16, p = .173)\). The change trajectory for the female mentor condition was significantly different from the no mentor and male mentor conditions \((B = 0.54, SE = 0.21, p = .009)\), showing that having a female peer mentor in the first year of college protected against declining post-graduate aspirations in engineering. When comparing conditions separately, the female mentor change trajectory was significantly different from the no mentor change trajectory \((B = 0.70, SE = 0.25, p = .005)\). The male mentor slope fell in between and was non-significantly different from the female mentor slope \((B \ldots\)
= -0.39, \( SE = 0.24, p = .100 \) and the no mentor slope (\( B = 0.32, SE = 0.25, p = .210 \)). See Figure 10.

\[
\begin{array}{|c|c|c|}
\hline
\text{Year} & \text{Female Mentor} & \text{Male Mentor} \\
\hline
0 & 0 & 0 \\
1 & -0.5 & -1 \\
2 & -1 & -1.5 \\
3 & -1.5 & -2 \\
4 & -2 & -2.5 \\
\hline
\end{array}
\]

\textbf{Figure 10.} Participants’ intentions to pursue graduate school in engineering from college entry through graduation.

\textbf{Emotional Health and Well-Being}

We also examined the impact of mentorship condition on emotional health and well-being throughout college. Whereas women without mentors and those assigned male mentors declined in their emotional well-being over time (no mentor: \( B = -0.42, SE = 0.20, p = .018 \); male mentor: \( B = -0.36, SE = 0.15, p = .014 \)), emotional well-being for women with a female mentor held steady throughout college (\( B = 0.13, SE = 0.14, p = .352 \)). The difference in change trajectories between the female mentor condition and the other two conditions was significant (\( B = 0.52, SE = 0.18, p = .004 \)), suggesting that female mentorship protected women’s emotional well-being throughout college. The female mentor change trajectory was significantly different from the no mentor change
trajectory \( B = 0.55, \ SE = 0.23, \ p = .015 \) and the male mentor change trajectory \( B = 0.49, \ SE = 0.20, \ p = .016 \). The no mentor and male mentor conditions were statistically equivalent \( B = 0.06, \ SE = 0.23, \ p = .812 \). See Figure 11.

Figure 11. Participants’ emotional health and well-being from college entry through graduation.

Anxiety and Motivation Mediate the Effect of Female Peer Mentors on Academic Choices in Engineering

Multilevel mediation analyses were conducted to test whether the effect of female peer mentors on objective academic outcomes (successfully securing engineering internships, STEM major at graduation) were explained by the change trajectory in anxiety relative to motivation in engineering. We used 2-1-2 mediation models (Preacher et al., 2010) with mentor condition as the independent variable that varied between participants (level 2), change in the ratio of anxiety to motivation as the mediator that varied within participants (level 1), and participation in engineering internships or
graduating with a STEM degree as the dichotomous dependent variable that varied between participants (level 2).

In each statistical model, the ratio of anxiety to motivation was regressed on time, creating random effects for the intercept and slope. The dependent variable (internship or STEM degree) was then regressed on the mediator slope (change in the ratio of anxiety to motivation) and the mediator slope was regressed on the independent variable (mentor condition). The intercept of the mediator was controlled for in each regression, accounting for participants’ actual values on the mediator. In order to compare the female mentor condition with the other two, condition was coded as a dichotomous variable (0 = no mentor and male mentor conditions and 1 = female mentor condition). Since Mplus does not allow for bootstrapping for indirect effects in a two-level model, each model was conducted with the Bayes estimator with 10,000 iterations, yielding near-identical results to bootstrapping (Muthén et al., 2017).

**Engineering Internships**

A significant indirect effect indicated that the effect of the female mentor condition versus the other two conditions on participation in engineering internships was mediated through change in the ratio of anxiety to motivation ($B = 0.14$, 95% CI [.01, .41]). Women who had female mentors during their first year of college had less change in anxiety relative to motivation compared to women in the other two groups ($B = -0.13$, 95% CI [-.25, -.01]). Having a smaller change in anxiety relative to motivation across time in college was associated with more success in securing engineering internships during college ($B = -1.10$, 95% CI [-2.61, -0.43]).

**Major at Graduation**
Another significant indirect effect indicated that the effect of the female mentor condition versus the other two conditions on graduating with STEM degrees was mediated through change in the ratio of anxiety to motivation ($B = 0.16, 95\% \text{ CI} [.02, .33]$). Compared to the other two conditions, having a female mentor in the first year of college elicited less change in anxiety relative to motivation across years in college ($B = -0.13, 95\% \text{ CI} [-.25, -.01]$), which in turn predicted increased likelihood of graduating with a STEM degree ($B = -1.26, 95\% \text{ CI} [-1.94, -.84]$).

**Male and Female Peer Mentors Produce Similar Benefits in Two Cases**

In most cases, our aforementioned results show that having a female peer mentor had significantly more benefits compared to the control condition whereas having a male peer mentor fell in-between, non-significantly different from both conditions. However, for two subjective outcomes, male mentors produced a benefit similar to that of female mentors, compared to those without mentors.

**Confidence**

Whereas women without mentors showed a decline in their feelings of confidence in engineering across their years in college ($B = -0.29, SE = 0.09, p = .001$), women with male mentors ($B = -0.08, p = 0.07, p = .252$) or female mentors ($B = -0.06, SE = 0.07, p = .379$) maintained confidence in engineering across time without decline (see Figure 12). The change trajectories of both male and female mentor conditions were significantly different from the change trajectory of the no mentor condition ($B = 0.22, SE = 0.10, p = .026$), suggesting that peer mentorship, regardless of mentor gender, protected women’s confidence throughout college. Upon comparing each mentor condition to the control condition, we found that women without mentors reported greater decline in confidence
than those with male ($B = 0.21, SE = 0.11, p = .056$) or female mentors ($B = 0.23, SE = 0.11, p = .040$). There was no difference between the change trajectories of the two mentor conditions ($B = 0.02, SE = 0.10, p = .859$).

![Figure 12](image)

*Figure 12.* Participants’ confidence in engineering from college entry through graduation.

**Belonging**

Whereas women without a mentor displayed a significant decline in feelings of belonging in engineering during college ($B = -0.21, SE = 0.09, p = .019$), those with male mentors ($B = -0.11, SE = 0.07, p = .107$) or female mentors ($B = -0.01, SE = 0.07, p = .902$) did not show significant changes in belonging (see Figure 13). However, we interpret these results with caution, because the change trajectories of the three conditions did not significantly differ from each other (female mentor vs. no mentor, $B = 0.20, SE = ...
0.12, $p = .081$; male mentor vs. no mentor, $B = 0.10$, $SE = 0.11$, $p = .389$; female mentor vs. male mentor, $B = 0.10$, $SE = 0.10$, $p = .303$).

![Change in Belonging](image)

**Figure 13.** Participants’ belonging in engineering from college entry through graduation.

**Discussion**

Our results indicate that same-sex peer mentorship relationships in the first year of college benefit women in engineering, throughout their college experience until graduation, long after mentoring relationships have ended. These benefits emerge for both subjective experiences and objective academic outcomes. The long-term impacts of this intervention are notable, especially given its low-cost and light-touch nature.

Four findings are particularly noteworthy. First, being assigned female peer mentors consistently prevented an increase in women’s anxiety relative to motivation. Being assigned male mentors or no mentors, however, significantly increased women’s anxiety relative to motivation. Second, in comparison to the other two conditions, being
assigned female peer mentors increased women’s success at securing and participating in engineering internships, increased their graduation rate for Bachelor’s degrees in STEM, and protected their aspirations to pursue post-graduate degrees in engineering. Third, female mentorship preserved mentees’ emotional well-being, whereas women with male mentors or no mentors declined in their emotional well-being throughout college. Fourth, protection against anxiety and preservation of motivation was the psychological mechanism that explained how female peer mentorship promoted success in securing engineering internships and success in completion of STEM degrees.

We also discovered two benefits that emerged in both the male and female mentor conditions. In contrast to our first-year results (Dennehy & Dasgupta, 2017), tracking mentees across their entire college experience revealed that having any mentor in the first year of college (male or female) preserved women’s confidence and belonging in engineering, whereas women without peer mentors displayed significant declines on both these variables. Mirroring other educational interventions (Barnett, 2011; Mattanah et al., 2010; Oppenheimer, 1984), this suggests a delayed benefit of male peer mentorship (a “sleeper effect”). One reason for this unpredicted finding may be that women who had male mentors in the first year of college became accustomed to working with male peers through their relationship with their male mentors. Greater feelings of familiarity with male peers may have promoted their confidence as they continued in upper-level engineering classes, labs, and teams where most of their peers were men. This is a speculative explanation and future research should examine whether this delayed benefit of male mentorship on belonging and confidence can be replicated and investigate if increased familiarity working with male peers is the underlying mediating mechanism.
Future research should also compare the roles of belonging, anxiety, and motivation as mediating mechanisms over a long educational time course. Our earlier article showed that during the first year of college, women’s feelings of belonging in engineering consistently mediated the effect of female peer mentors on career aspirations in engineering (Dennehy & Dasgupta, 2017). In contrast, our present investigation shows that when the entire college experience is analyzed as a whole, lower anxiety relative to higher motivation mediated the effect of female peer mentors on women’s academic outcomes. These results suggest that the mechanism that predicts women’s persistence and success in engineering may change over time in college. During the first year of college, it may be particularly important for women to feel a sense of belonging in engineering because the culture of college classes is very different from high school. However, as classes become more difficult and students advance to their second year and beyond, the relative balance of worries and motivation may ultimately predict their long-term persistence in engineering. More research is needed to replicate these findings and explore the role of different types of psychological mechanisms that play important roles at different developmental stages of students’ academic careers.

Another discrepancy emerged in terms of student outcomes during the first year of college versus looking at students’ entire college career. In our previous article, we found that at the end of the first year, 100% of women assigned a female mentor remained in engineering majors, compared to 82% of women assigned a male mentor and 89% of women assigned no mentor (Dennehy & Dasgupta, 2017). Our present research shows that this difference was no longer significant at graduation; however, a closer examination of the data suggests that some women who had female mentors switched out
of engineering after their first year to a different STEM field—specifically computer science, mathematics, chemistry, or biology. Indeed, women with female mentors were more likely to graduate with STEM degrees compared to women in the other two conditions. Thus, it appears that having female mentors in engineering during their first year proved beneficial for women’s retention in STEM, broadly, which is crucial as women face underrepresentation across multiple STEM disciplines.

It is also worth noting that women’s engineering and STEM grade point average (GPA) in college did not differ between mentor conditions ($p \geq .457$), nor did their GPA trajectories significantly change over time ($p \geq .161$) or differ by mentor condition ($p \geq .504$). Thus, the benefit of female mentorship on women’s academic behaviors did not occur by improving mentees’ grades, but by inoculating them against subjective experiences of anxiety and preserving their motivation in engineering through their college years.

Some important future directions include replicating these findings across other STEM disciplines where women are also a minority (e.g., computer science, physics; NSF, 2019) and examining mentoring interventions at later points in women’s STEM careers (e.g., graduate or postdoctoral traineeship in STEM or early career work experiences in STEM). Additionally, although our research focused on women, an important future direction is to investigate the generalizability of these findings to other identity groups that are underrepresented and negatively stereotyped in STEM (e.g., Black, Hispanic, Indigenous, and first-generation college students).

It is time to turn national attention to implementing research-driven solutions in academic and professional institutions to expand the training and success of
underrepresented students in STEM pathways and track the longitudinal impacts of these solutions. Our findings indicate that a low-cost, light-touch mentoring relationship with a successful own-group peer during the transition to college yields dividends long after it has concluded, both for subjective indicators of anxiety, motivation, and emotional well-being and objective indicators of skills and persistence. Such a relational intervention, initiated in the transition to college, constitutes an important step towards equalizing the representation of tomorrow’s scientists and engineers.
Supplemental Analyses

Relationships with Peer Mentors

To assess whether mentees had similar experiences with female compared to male mentors or not, we examined if there were any differences in: 1) the content of conversations between mentees and their mentors and 2) mentees’ subjective reports of their relationship with their mentors. Conversational content was extracted from short surveys completed by mentors; whenever mentors met with their mentees, they reported what they discussed during their get-together. These meeting descriptions were coded by research assistants (blind to mentor condition) as being primarily social (e.g., how college was going, how to make friends) or primarily academic (e.g., which classes to take, how to get internships or jobs). Independent t-tests revealed no differences between mentor conditions in the percentage of meetings that were primarily social (female mentors: $M = 38.12\%, SD = 26.26$; male mentors: $M = 36.01\%, SD = 25.69$; $t(90) = 0.39$, $p = .698$) or primarily academic (female mentors: $M = 45.22$, $SD = 15.50$; male mentors: $M = 47.53$, $SD = 23.21$; $t(90) = -0.56$, $p = .576$).

Mentees’ subjective reports of their relationship with mentors was extracted from a short survey that mentees completed at the end of their mentorship year. Mentees completed 7 items that asked: “How much do you identify with your peer mentor?”, “How similar do you feel to your peer mentor?”, “Can you imagine yourself achieving a similar level of success as your peer mentor?”, “Do you feel personally connected to your peer mentor?”, “Do you feel your mentor-mentee relationship has good chemistry?”, “How much do you admire your peer mentor?”, “How much support have you been getting from your peer mentor?” Mentees responded on 7-point scales ranging from 1
(not at all) to 7 (very much). An eighth item asked mentees to identify how close they felt to their mentor using overlapping Venn diagrams, with one circle representing themselves and the other representing their mentor (from 1 indicating zero overlap between themselves and their mentor and 7 indicating high overlap; drawn from Aron et al., 1992; see Figure 14). The greater the overlap, the more mentees felt closer with their mentors.

Which of the pictures below best describes your mentor-mentee relationship? Please type the number that best corresponds to your response.

Figure 14. Item asking participants to describe how close they felt with their mentor. This question is based on The Inclusion of Other in the Self (IOS) Scale (Aron et al., 1992).

These 8 items had high reliability ($\alpha = .96$) and were averaged to form an index of mentee’s subjective identification with their mentor. Mentor identification did not significantly differ by condition (female mentor: $M = 5.04, SD = 1.44$; male mentor: $M = 4.67, SD = 1.66$; $t(99) = 1.19, p = .236$). Because mentees’ relationships with their mentors did not differ on the basis of content discussed during mentoring meetings nor subjective identification with their mentor, we concluded that the benefits of female mentors were due to their shared social identity, rather than systematic differences in the
quality of mentorship provided by female vs. male mentors or how connected mentees felt to their mentors.

**Race and Generation Status**

Although the primary goal of the current research was on the impacts of gender similarity vs. dissimilarity between mentors and mentees, mentor-mentee similarity vs. dissimilarity on other social identities matter as well. For example, it is important to identify whether underrepresented racial minorities (URMs) and first-generation students benefit more from having a mentor who shares their racial identity or generation status (vs. not). However, our current sample was primarily comprised of White and continuing-generation students. In each condition, White students comprised between 65% to 73% of mentees (female mentor condition: 38 out of 52; male mentor condition: 33 out of 51; control condition: 30 out of 47). Black, Hispanic, and Indigenous students only made up 6% to 10% of mentees (female mentor condition: 4 out of 52; male mentor condition: 5 out of 51; control condition: 3 out of 47). All these participants had mentors who were White. Thus, we were unable to test whether having a mentor of the same race had benefits for URMs.

Students’ first-generation status was the proxy for their family social class. First-generation students made up 14% to 17% of each condition (female mentor condition: 9 out of 52; male mentor condition: 8 out of 51; control condition: 7 out of 47). We did not ask mentors to report whether or not they were first in their family to attend college, making it impossible to examine whether having a mentor who was also first-generation might provide more benefits than a continuing-generation mentor for first-generation mentees.
Transition to Chapter 4

Interventions to promote the success of underrepresented minorities in STEM in college are crucial both from the perspective of fairness and equity and also pragmatically to meet the demands of the STEM workforce. In Chapter 3, I show that a one-year peer mentor intervention during their first year of college had long-term benefits for women in engineering, particularly when they had a female mentor. In other words, these findings provide evidence that same-sex peer mentorship has benefits for women in disciplines where they are underrepresented, especially during transition periods. These results are consistent with the Stereotype Inoculation Model (Dasgupta, 2011), and show that having a single mentor of the same stigmatized social identity can protect against the negative effects of social identity threat, just as a biomedical vaccine can protect individuals from viruses. This data should inform universities’ development of programs to promote the success of underrepresented students in STEM majors because it demonstrates that these types of interventions (i.e., ingroup community-building) are effective.

It is also important to build on community-building research in several ways. The focus of the study in Chapter 3 was to foster one-on-one relationships with ‘near peers’ who were a few years senior to mentees and successful in their majors. Explored in the next chapter, a different type of intervention examines the impact of creating a community of peers, all at the same stage of their academic journey, all of whom share the same stigmatized identity (i.e., a peer cohort). Specifically, Chapter 4 expands upon the previous chapter by examining the impacts of a peer cohort for underrepresented students, this time with a focus on a different social identity (students’ social class) and a
different type of STEM field (biological sciences). The studies in Chapter 4 complement Chapter 3 by also employing the Stereotype Inoculation Model (Dasgupta, 2011) in testing whether the presence of ingroup peers has protective effects for underrepresented students’ subjective experiences and retention in STEM.
CHAPTER 4

A LIVING LEARNING COMMUNITY FOR FIRST-GENERATION STUDENTS IN BIOLOGY

Increasing student engagement, success, and professional development in science, technology, engineering, and mathematics (STEM) is a national priority for the United States given the high demand for a skilled STEM workforce and a shortage of college graduates pursuing STEM education (President’s Council of Advisors, 2012; Zaza et al., 2020). This gap exists in part due to the underrepresentation of women, Black, Hispanic, Indigenous, and first-generation college students in STEM majors (Ambrose, 2019; Bettencourt et al., 2020; Landivar, 2013; Xue & Larson, 2015). As a case in point, though first-generation (FG) students (i.e., individuals whose parents did not complete a 4-year college degree) make up 56% of undergraduates (RTI International, 2019), they make up 36% of STEM college graduates (Bettencourt et al., 2020). Specifically, FG college students are 1.8 times less likely to attain bachelor’s degrees in STEM compared to their peers from continuing-generation (CG) families, whose parents completed four-year college degrees (Bettencourt et al., 2020).

Several hurdles deter FG college students’ pursuit of STEM in particular, and college completion in general. They are more likely to come from working class families that have limited financial resources to support them in college (Reardon, 2011). They are also more likely to have attended high schools with fewer resources and, thus, may be underprepared in foundational pre-college STEM classes (Warburton et al., 2001). FG college students are also less likely to get college-specific advice and coaching from their parents on how to navigate higher education, which CG students routinely get from their
parents (Guryan et al., 2008; Ramey & Ramey, 2010). Furthermore, FG students often come from families that emphasize interdependence and community learning, which are dissonant with typical middle-class university values that emphasize independent self-directed learning, creating a cultural mismatch between students’ family norms and college norms (Fryberg & Markus, 2007; Stephens et al., 2012). This cultural mismatch results in increased stress and lower grades for FG college students (Phillips et al., 2020; Stephens et al., 2012). In sum, though college degrees, particularly in STEM, provide pathways to upward social mobility (Bowen et al., 2005; Melguizo & Wolniak, 2012), the economic and social capital barriers faced by FG students often hinder them from achieving this dream. Moreover, class-based disparities in college and in STEM continue to exacerbate income inequality between the working class and the middle class (Dika & D’Amico, 2016).

Given the large numbers of working-class students who are the first in their families to attend college in the United States, and who have the potential to be the next generation of scientists and engineers, developing and testing theory-driven interventions to promote their success is of national interest. Yet, surprisingly few studies have rigorously tested evidence-based interventions targeting FG student success in STEM. A few notable exceptions are studies that used values-oriented interventions to reduce achievement gaps in grades between FG and CG students in STEM (Harackiewicz et al., 2014; 2016; Hecht et al., 2021; Tibbetts et al., 2016). Specifically, when FG students were invited to reflect on personal values that were important to them (value affirmation induction) or invited to describe how their course content in biology was personally meaningful to them (utility value induction), their subsequent grades in introductory
biology were significantly higher, and they were more likely to persist in biology the following semester in comparison to FG students who did not complete these interventions (Harackiewicz et al., 2014; 2016; Tibbetts et al., 2016). Another study found that a value affirmation intervention improved FG students’ performance on a math test in a laboratory situation (Hecht et al., 2021).

While these studies make important contributions by showing that brief interventions can have potent benefits for FG students in STEM, they also have limitations. First, by focusing on changing students’ mental appraisals, these studies place the onus on underrepresented students to adjust to academic institutions by changing their mindset rather than placing the onus on academic institutions to change learning environments to fit the needs of diverse students. Second, key outcomes in these studies were academic performance (grades, test performance) and retention in the major. Students’ subjective experiences in STEM courses such as belonging, confidence, motivation, and anxiety in STEM were either unexamined (Harackiewicz et al., 2016; Hecht et al., 2021) or were unchanged by the intervention (Harackiewicz et al., 2014; Tibbets et al., 2016). The present research addresses these limitations by designing and field-testing an intervention that modifies the traditional STEM learning environment in biology and examines its impacts on students’ psychological experiences (e.g., belonging, confidence, anxiety, and motivation), performance (grades in biology), and long-term retention in biological science pathways.

Our research is informed by the Stereotype Inoculation Model (Dasgupta, 2011), which proposes that features of local environments in which people are immersed (e.g., academic and professional spaces) have profound impacts on their feelings of belonging,
confidence, motivation to persist, and future aspirations in these fields. These environmental features include exposure to, and authentic relationships with, ingroup members, both peers and professionals, in fields where students face stereotypes casting doubts on their group’s ability. Just as exposure to biomedical vaccines protect and inoculate one’s body against noxious viruses, exposure to ingroup experts and peers protect and inoculate one’s mind against noxious stereotypes. The model proposes that the impact of negative stereotypes is particularly harmful during periods of transition when individuals move from familiar environments to new ones with unfamiliar norms, customs, and expectations. Thus, interventions inserted in these transition periods are predicted to be particularly impactful. Informed by the Stereotype Inoculation Model, the present research rearranged a traditional learning environment in undergraduate biology to make it more welcoming and inclusive. Notably, our intervention differs from others that train students to alter their mindsets to help them better navigate the traditional college environment (e.g., self-affirmation exercises, Cohen & Sherman, 2014; mindsets about intelligence, Yeager & Dweck, 2012). Our focus on changing the “business as usual” educational environment is driven by the acknowledgement that colleges and universities possess more structural power and resources to create change than individual students. The responsibility for change resides with institutions.

Consistent with the Stereotype Inoculation Model, prior research found that situations that afford exposure to successful same-sex peers in STEM (e.g., Cheryan et al., 2011; Dasgupta et al., 2015; Marx & Roman, 2002; Stout et al., 2011) and opportunities to form mentoring relationships with same-sex peers (Dennehy & Dasgupta, 2017; Wu et al., 2021) enhanced women’s academic experiences and
performance in STEM fields where women are underrepresented. Similarly, exposure to high-level women scientists, mathematicians, and engineers have positive impacts on women students, both when encountered in person and when their stories are heard through media (Stout et al., 2011). However, this research has its own limitations. One limitation is that stereotype inoculation interventions have almost exclusively focused on gender (see Johnson et al., 2019; Pietri et al., 2018; 2019 for those based on race). Interventions that examine educational outcomes for other identity groups that face negative stereotypes in STEM, such as FG college students, have not been examined. A second limitation is that stereotype inoculation interventions for minorities in STEM typically involved brief exposure to ingroup members (e.g., Cheryan et al., 2011; Johnson et al., 2019; Pietri et al., 2018; 2019) or a one-on-one mentoring relationship with a single person (Dennehy & Dasgupta, 2017; Wu et al., 2021). Little work has examined whether immersion in a group of peers (e.g., a learning community) all of whom share a common identity has benefits for minorities in STEM. None have empirically tested the impacts of this type of learning community specifically for FG students in STEM.

The current research addresses these limitations by designing a living learning community (LLC) for FG students in the biological sciences. This work also expands upon prior research on LLCs, which have been limited and mixed. First, although LLCs are popular at some universities because participating students report greater belonging and social connection compared to non-participating students (e.g., Brower & Inkelas, 2010; Inkelas et al., 2007; Wawrzynski & Jessup-Anger, 2010), little research has examined their impact on underrepresented students in STEM. Most LLC participants have been students from majority identity groups—i.e., White and continuing-generation
students (Inkelas et al., 2007; Sriram & Diaz, 2016). Second, even when underrepresented students were in LLCs, they continued to be in the numeric minority in those communities (Inkelas et al., 2006; 2007; Johnson et al., 2007; Sriram & Diaz, 2016), which may account for mixed results. While some studies found that women and FG students benefitted from LLCs even when they were in the minority (Inkelas et al., 2007; Szelenyi et al., 2013), other studies found that racial ethnic minorities suffered lower feelings of belonging compared to White students when they were in the minority in LLCs (Eidum et al., 2020; Inkelas et al., 2006; Johnson et al., 2007). Most LLCs were not designed around shared identity groups (race, gender, social class, etc.) for students interested in STEM, with a few notable exceptions. One exception is a study that found women in a one-year all-female LLC in STEM were more likely to graduate with a Bachelor’s degree in STEM (Maltby et al., 2016). Another example is the Meyerhoff Scholars Program at the University of Maryland, a 4-year learning community (with a living component during the first year of college) for individuals who were highly committed to: 1) pursuing STEM (e.g., had participated in STEM research in high school) and 2) the advancement of minorities in STEM. Black students in this program earned Bachelor’s degrees and PhDs in STEM at higher rates than Black students who were not in the program (Domingo et al., 2019; Lee & Harmon, 2013).

We built on this sparse literature by conducting two longitudinal quasi-experiments in a naturally existing academic setting to test the impact of a one-year living learning community in the biological sciences on FG working-class students in their transition to college. Study 1 followed students through their first semester in college
while they took their first introductory biology class. Study 2 followed them through their
second semester in college while they took the second biology class in the sequence.

Undergraduate students at a large public research university were recruited the
summer before their first year in college. Eligible participants were first in their families
to attend college, and all had expressed interest in majoring in the biological sciences
(i.e., biology, microbiology, biochemistry and molecular biology, neuroscience, animal
sciences, or veterinary technology). All were scheduled to take introductory biology in
their first semester of college. In the summer before starting college, eligible new
students received a letter inviting them to apply for a living learning community for FG
students. Recruitment occurred in three academic years: 2018-2019, 2019-2020, and
2020-2021. Each year, among those who applied to the LLC, 19-38 FG students were
quasi-randomly assigned to the LLC (intervention condition) or the control condition
with one constraint: Black, Hispanic, and low-income students were oversampled into the
LLC condition. FG students who were not selected for the LLC and others who had not
applied served as our control condition. Participants in the LLC lived in the same
residential college, took introductory biology and a first-year seminar as a cohort, and
had access to a peer mentor who was also a FG student in biology. In contrast, FG
students in the control condition took introductory biology in a large lecture class and
took their first-year seminars with CG students. They lived in residential halls with CG
students and did not have a dedicated peer mentor who was a FG student. In other words,
first-generation students in the control condition were not immersed in an academic
community of peers who shared their social class background. Furthermore, FG
participants in the LLC and FG participants in the control group had the same faculty
instructor for introductory biology; thus, the course content, lectures, assignments, and exams were identical across conditions.

We also compared the FG LLC students to another comparison group—a LLC of students in the honors college. Honors college students make up approximately 13% of the incoming first-year class at this university and are admitted based on high achievement in high school and a holistic review of their college application. Honors college students interested in the biological sciences were invited to apply to an honors biology LLC. Among those who applied, approximately 30% of applicants were accepted into the program. Similar to the FG biology LLC, students in the honors LLC lived in a common residence hall and also took honors introductory biology and a seminar class as a cohort. Most of them came from college-educated families (88-90%).

In sum, our studies comprised three groups of participants: 1) an intervention group comprised of FG students in a FG biology LLC; 2) a control group comprised of FG college students not in an LLC; and 3) a comparison group with honors students in an honors biology LLC. This design allowed two comparisons. First, by comparing the two groups of FG students (intervention vs. control group), we sought to determine whether immersion in a LLC would significantly enhance the academic outcomes of FG students. Second, by comparing the two LLC groups (intervention vs. honors), we sought to determine whether participation in a LLC would close the academic gap between FG students and high-performing honors students from mostly college-educated families.

Unexpectedly, the COVID-19 pandemic disrupted the experience of multiple cohorts of participants. The residential component of the LLCs was disrupted for Cohort 2 (spring 2020) and Cohort 3 (fall 2020) and resumed in spring 2021. Moreover, the
previously face-to-face biology courses were taught remotely for Cohort 2 (spring 2020) and Cohort 3 (academic year 2020-2021). Despite these disruptions, the other academic and social features of the LLCs continued unabated, including small class size (i.e., a small biology class of FG students only or honors students only), peer mentoring, and community-bonding activities.

Our measures included students’ psychological experiences in biology (i.e., feelings of belonging, motivation, confidence, and anxiety), perceived relevance of biology in the real-world, final grades in biology, and retention in biological science majors up to a year after the conclusion of the LLC. Participants were surveyed twice, during the middle and end of each semester, to assess the consistency of their responses throughout the semester. Three research questions guided our work:

1) Does participation in a biology LLC for first-generation students enhance their academic experiences in biology (belonging, anxiety, motivation, and confidence), perceptions of the relevance of biology in the real-world, and final grades as compared to first-generation students who are not in a LLC?

2) Are first-generation students who participated in a biology LLC more likely to persist in biological science majors beyond their first year of college as compared to first-generation students who are not in a LLC?

3) Does participation in a LLC close the achievement gap between first-generation students and honors college students in terms of academic experiences in biology, perceptions of biology, and retention in biological science majors?

**Study 1 Method**
Study 1 was a longitudinal quasi-experiment testing the impact of a non-traditional living learning intellectual community of first-generation college students in biology set within a naturally-existing university environment. Students were followed longitudinally from the beginning to end of the semester. Data were collected on their academic experiences in class and end-of-semester grades. This intervention group was compared to two comparison groups: first-generation students who were eligible for the LLC but were not in it, and honors students in biology in an honors LLC, who were primarily CG students.

**Participants**

Two hundred and forty-three students ($M_{\text{age}} = 18.09$ years, $SD = 0.40$) participated in our study during the fall semester of their first year in college. The sample comprised 162 females, 77 males, and 4 nonbinary students. In terms of race/ethnicity, 42% were White, 22% Asian, 16% Black or African American, 9.5% Multiracial, 7% Hispanic, and 3% other ethnicity. All students were first-years, enrolled in the first introductory biology course required for all biological science majors (Biology 151 or Biology 161H for honors students). Data collection occurred during the fall semesters of 2018, 2019, and 2020. The intervention group ($n = 86$) included FG students in a biology LLC; the control group ($n = 63$) included first-generation students not in a LLC but who were eligible to participate; and the second comparison group ($n = 94$) included first-year students in the honors program who were in a biology LLC for honors students. As an incentive, participants who completed the surveys were entered into a drawing for college memorabilia if they were in the 2018 cohort. Participants in the 2019 and 2020 cohorts
received $10 for survey participation and were also entered into a drawing for one of three $100 cash prizes.

Procedure

*Academic Experience Surveys*

All participants completed two surveys during the fall semester, once in the middle of the semester (4-6 weeks into the semester; Time 1) and again at the end of the semester (10-12 weeks into the semester; Time 2).

*Intervention Group: First-Generation LLC*

In the summer preceding their first year of college, all first-generation students who were planning to major in the biological sciences and had preregistered for introductory biology were invited to apply to take part in a LLC designed to be an affinity community of peers from similar identity backgrounds interested in the biological sciences. Students who applied for the LLC were randomly assigned to the LLC or the control condition with one constraint. The LLC condition oversampled students from groups that are historically underserved in STEM: low-income students who were eligible for Pell grants (an indicator of family poverty), and students who identified as Black or Hispanic. Students from these two identity groups were assigned to the LLC at a higher rate than the control condition (e.g., 41% of the intervention group were Black or Hispanic as compared to 26% of the control group).

There were 5 distinct features of the FG LLC. First, students took introductory biology together as a small class (i.e., 19-38 students). Second, students had a weekly seminar in which they learned about research opportunities and heard from biological sciences faculty who were also FG. Third, students had peer mentors—sophomores or
juniors—who were FG students in biology. Fourth, once a semester, LLC students gathered for a community-building social activity. Fifth, students lived together in the same residential hall and had a roommate who was also in the LLC. The pandemic compromised the residential feature of the LLC in some cohorts: while those in 2018 and 2019 (60% of the sample) experienced this residential component, those in the 2020 cohort were fully remote due to COVID-19 pandemic (40% of the sample).

Control Group

Instructors who were teaching Biology 151 to first-generation LLC students were also teaching other sections of the same course to non-LLC students. These sections were larger (350-400 students) and included both continuing-generation and some first-generation students. From these sections, we recruited first-generation students to be a part of the control group. In addition to having the same instructor, the control group encountered the same course material, syllabus, lectures, assignments, exams, and grading as the FG LLC condition. Students in the control condition attended weekly first-year seminars with other students, the majority of whom were not FG students. They did not have an assigned peer mentor and did not live with a roommate who matched their FG status and major.

Honors LLC

Participants in the honors LLC were members of the honors college, who entered the university with high school records that exceeded the entering class (i.e., the top ~13% of the incoming class). The vast majority of participants in the honors LLC (90%) were from families with college-educated parents (i.e., continuing-generation students). Similar to the FG LLC, students in the honors LLC took introductory biology (Biology
and attended a weekly first-year seminar in which faculty in the biological sciences presented their ongoing research (cohort size of 43-48 students). In 2018 and 2019, honors LLC students lived together in the same residence hall and also had roommates from the LLC (66% of the sample), while the 2020 cohort went remote during the pandemic (34%), and took their courses remotely, synchronously via Zoom.

Measures

Survey Measures

At Time 1 (middle of the semester) and Time 2 (end of the semester), participants reported their psychological experiences in their fall semester biology class. All measures utilized 7-point response scales ranging from 1 (not at all) to 7 (very much).

Belonging. Four items were used to measure participants’ feelings of belonging in Biology 151 or 161H. These items were: “I feel connected to my classmates in BIO151/BIO161H,” “I feel accepted by my classmates in BIO151/BIO161H,” “I feel like an outsider among my classmates in BIO151/BIO161H (reverse-coded),” and “I feel invisible in BIO151/BIO161H (reverse-coded).” These four items had good reliability (Time 1: \( \alpha = .81 \), Time 2: \( \alpha = .83 \)) and were averaged to create a construct for perceived belonging in biology for each timepoint.

Anxiety. Four items were used to measure participants’ feelings of anxiety regarding Biology 151 or 161H. These items were: “I feel worried about BIO151/BIO161H,” “I feel stressed about BIO151/BIO161H,” “I feel unsure about BIO151/BIO161H,” and “BIO151/BIO161H is difficult this semester.” These four items

---

6 The course content and teaching structure of Biology 161H differed from Biology 151, which was taken by the FG participants. Though Biology 151 was a lecture-based course, Biology 161H was a team-based learning course.
(Time 1: $\alpha = .84$, Time 2: $\alpha = .89$) were averaged to create an anxiety index for each timepoint.

**Motivation.** Four items measured participants’ motivation to persist in Biology 151 or 161H. These items were: “I have the skills and abilities to successful in BIO151/BIO161H this semester,” “I will be able to overcome any challenges I experience in BIO151/BIO161H this semester,” “I have what it takes to handle BIO151/BIO161H,” and “I feel confident about BIO151/BIO161H.” These items were averaged to create a motivation index at each timepoint (Time 1: $\alpha = .93$, Time 2: $\alpha = .94$).

**Confidence.** Two items measured participants’ confidence in their ability in biology. These items were: “In general, how confident do you feel about your ability in biology?” and “Do you think you have a knack or talent for biology?” These items were averaged to create an index of confidence at each timepoint (Time 1: $\alpha = .83$, Time 2: $\alpha = .84$).

**Real-World Relevance of Biology.** Two items measured how important participants found biology to be. These items were: “To what extent does biology solve important real-world problems that help people?” and “Is biology important in everyday life?” These items were averaged to create an index of biology importance at each timepoint (Time 1: $\alpha = .69$, Time 2: $\alpha = .73$).

**Fall Semester Grades**

Participants’ written consent was requested to access their transcript from the university registrar. If permission was granted, we recorded students’ final letter grade in biology and converted it using the conventional 4.0 grading scale (i.e., A = 4.0, A- = 3.7,
B+ = 3.3, B = 3.0 etc.). Students who took the class pass/fail were not included in the present analyses. Because honors biology (Biology 161H) had slightly different content and a different teaching method, their grades were not comparable to the biology course (Biology 151) taken by first-generation students. Thus, student grades in the honors LLC condition were not compared to the FG LLC group. We only compared condition differences in grades between the control group and the first-generation LLC group because they experienced the same course content (lectures, assignments, exams) with the same instructor.

Study 1 Results

Manipulation Check

Because student assignment to condition was quasi-random in the sense that Black, Hispanic, and low-income students were oversampled into the FG LLC, we conducted a manipulation check to compare student perceptions of college preparation across conditions. The goal was to ensure that those in the FG LLC did not have any baseline benefits compared to the control group. The manipulation check questions also allowed us to verify whether, as expected, FG students felt less prepared for college compared to those in the honors LLC, consistent with previous research finding gaps between FG and CG students.

We asked participants three questions that were used as manipulation checks: 1) how much their family prepared them for college, 2) how much their high school prepared them for college-level biology, and 3) the extent to which they had assumed they would go to college while in high school. One-way Analysis of Variance (ANOVAs) were conducted with Bonferroni corrections to compare participant responses.

Participants reported these measures during their first timepoint in the study.
between conditions. For family preparation, a significant main effect of condition, $F(2, 290) = 20.02, p < .001, d = 0.74$, revealed that students in the honors LLC reported that their family prepared them for college significantly more ($M = 5.88$) than either of the FG student groups (control: $M = 5.27, p = .027$; intervention group: $M = 4.36, p < .001$).

First-generation students in the control group felt significantly more prepared for college by their family than those in the FG LLC ($p < .001$). For high school preparation, a significant effect of condition, $F(2, 290) = 3.86, p = .022, d = 0.32$, showed that students in the honors LLC reported that their high school prepared them for college ($M = 5.25$) more than students in the FG LLC ($M = 4.61, p = .020$). The control group fell in-between ($M = 4.85$) and did not significantly differ from either group ($ps \geq .237$). Finally, a significant condition difference emerged in students’ high school assumptions of whether they would go to college, $F(2, 289) = 5.83, p = .003, d = 0.40$. Those in the honors LLC were significantly more likely to assume that they were college-bound ($M = 6.97$) than those in the FG LLC ($M = 6.71, p = .002$). Participants in the control group ($M = 6.87$) fell in-between and did not significantly differ from either group ($ps \geq .096$). The results of these manipulation check confirm that: 1) those in our FG LLC group were similar to the control group in terms of high school preparation, and 2) students in the honors LLC felt more prepared for college than the two FG groups.

**Data Analytic Strategy**

We conducted 3 Condition (control, FG LLC, honors LLC) x 2 Time ANOVAs for each of the dependent measures. If a significant effect of condition, time, or a condition x time interaction emerged, follow-up post hoc tests were conducted using Bonferroni corrections to avoid Type I errors (false positives). Sensitivity power analyses
(GPower 3.1; Faul et al., 2009) revealed 95% power (3 groups, 2 measurements across time) for an effect size as small as $d = 0.22$ (243 participants).

**Belonging**

As hypothesized, a significant main effect of condition, $F(2, 240) = 8.15, p < .001$, $d = 0.52$, revealed that across the semester, FG students in the LLC ($M = 5.51$, $SE = 0.12$) developed a stronger sense of belonging in biology than FG students in the control condition ($M = 4.90$, $SE = 0.14$, $p = .003$, $d = 0.31$). The FG LLC was statistically similar to the honors LLC condition in terms of belonging ($M = 5.57$, $SE = 0.11$, $p > .99$), but the FG controls felt significantly less belonging in biology compared to the honors LLC ($p = .001$, $d = 0.34$). While the main effect of time was nonsignificant, $F(1, 240) = 0.52, p = .471$, a significant time x condition interaction, $F(2, 240) = 3.36, p = .036, d = 0.33$, showed that the magnitude of difference in feelings of belonging between the FG LLC and the FG control conditions was larger at Time 1 ($p < .001$, $d = 0.36$) than at Time 2 ($p = .075$, $d = 0.20$). At both timepoints of the semester, the difference between the FG LLC and the honors LLC was nonsignificant (Time 1: $p > .99$; Time 2: $p = .615$). Means for each timepoint are graphed in Figure 15.
Anxiety

A significant main effect of condition, $F(2, 240) = 10.28, p < .001, d = 0.58$, revealed that FG students in the LLC ($M = 3.44, SE = 0.12$) reported less anxiety than FG students in the control group ($M = 4.07, SE = 0.15, p = .005, d = 0.32$) and those in the honors LLC ($M = 4.22, SE = 0.12, p < .001, d = 0.40$). There was no difference in anxiety between the control and the honors LLC conditions ($p > .99$). Additionally, a significant main effect of time, $F(1, 240) = 5.31, p = .022, d = 0.30$, showed an overall increase in anxiety across the semester from Time 1 ($M = 3.82, SE = .09$) to Time 2 ($M = 4.00, SE = 0.09$), for all conditions. There was no time x condition interaction, $F(2, 240) = 0.04, p = .960$. Means for each timepoint are graphed in Figure 16.
Figure 16. Participants’ feelings of anxiety in Biology 151 in Study 1.

Motivation

We found a significant main effect of condition for motivation, \( F(2, 240) = 3.40, p = .035, d = 0.34 \), such that FG students in the LLC (\( M = 5.70, SE = 0.11 \)) reported significantly more motivation in biology than FG students in the control condition (\( M = 5.26, SE = 0.13, p = .033, d = 0.23 \)). The honors LLC condition did not significantly differ from the two FG conditions (\( ps \geq .324 \)). Furthermore, the main effect of time was nonsignificant, \( F(1, 240) = 0.003, p = .958 \), as was the interaction between time x condition, \( F(2, 240) = 2.91, p = .057 \). Means for each timepoint are graphed in Figure 17.
Figure 17. Participants’ feelings of motivation in Biology 151 in Study 1.

Confidence

Students' confidence in biology did not vary significantly by condition, $F(2, 238) = 1.99$, $p = .139$, or by time, $F(1, 238) = 0.80$, $p = .372$. However, a significant time x condition interaction, $F(2, 238) = 4.43$, $p = .013$, $d = 0.38$, indicated that students in the honors LLC increased in confidence over time (Time 1: $M = 5.05$, $SE = 0.12$; Time 2: $M = 5.34$, $SE = 0.12$, $p = .003$, $d = 0.16$), while students in the FG LLC and control conditions did not significantly change in their confidence over time ($ps \geq .218$.)

Perceived Relevance of Biology

Participants’ perceptions of the relevance of biology did not vary by condition, $F(2, 237) = 1.82$, $p = .165$, time of semester, $F(1, 237) = 2.87$, $p = .092$, or time x condition, $F(2, 237) = 0.90$, $p = .408$.

Fall Semester Grades
We tested whether students’ final grades in biology varied systematically by condition. As a reminder, FG students in both conditions took the same course with the same instructor (i.e., same lecture content, assignments, exams), though those in the LLC took it in a small section with others in the FG LLC and those in the control group took it in a large lecture class. We did not compare grades with the honors LLC as their biology course differed in content and type (i.e., team-based learning course as opposed to lecture-style). As predicted, FG students in the LLC earned significantly higher grades in biology ($M = 3.32$ or B+, $SD = 0.63$) than FG students in the control group ($M = 2.93$ or B, $SD = 0.67$), $t(134) = 3.56, p = .001, d = 0.61$. These values are graphed in Figure 18.

Figure 18. Participants’ final grades in Biology 151 in Study 1.

**Study 1 Discussion**

In Study 1, which focused on the first semester of students’ college experience, immersion in a LLC with other FG students had benefits for FG students’ belonging, motivation, anxiety, and grades in introductory biology compared to FG students in the control group. In addition, students in the FG LLC were less anxious about biology than
students in the honors LLC\textsuperscript{8}, and had similar levels of belonging, motivation, and confidence as the honors LLC group. These findings suggest that a LLC significantly enhances FG students’ experiences and performance in academic science compared to other FG students from a very similar background who did not participate in a LLC. Moreover, participating in the LLC appears to erase the gap between FG students and CG students accepted into a selective honors program. In sum, this first study provides strong preliminary evidence that for the first semester of college, a biological sciences LLC consistently enhances the academic and psychological experiences of first-generation college students in the biological sciences.

\textbf{Study 2 Method}

To see whether the benefits observed in the first semester would continue through the second semester, we conducted a second study examining the impacts of a LLC for first-generation students in their second semester of college while they enrolled in a higher-level course in biology.

\textbf{Participants}

One hundred and ninety-nine students ($M_{\text{age}} = 18.35$ years, $SD = 0.54$) participated in our study during the spring semester of their first year in college. The sample was comprised of 133 females, 62 males, and 4 nonbinary participants, and was 41\% White, 26\% Asian, 13\% Black or African American, 10\% Hispanic, 8\% Multiracial, and 2.5\% other ethnicity. All students were enrolled in their first-year of college taking a second biology course that was part of the introductory sequence (Biology 152 or Biology 162H for honors students). Data collection occurred during spring 2019, 2020, 2021.

\footnote{\textsuperscript{8} We did not predict that those in our honors LLC would report greater anxiety than those in the FG LLC. We speculate that this may have occurred as honors students are likely to express perfectionistic tendencies, which may lead to increased stress (Rice et al., 2006; Wimberley & Stasio, 2012).}
and 2021. The intervention group \( (n = 76; 75 \text{ were the same from Study 1}) \) included FG students in a biological sciences LLC; the control group \( (n = 50; 10 \text{ students were the same from Study 1 while 40 students were new}) \) included FG students not in a LLC but were eligible to participate; and the second comparison group \( (n = 73; 65 \text{ were the same from Study 1}) \) included first-year students in the honors program who were in a biological sciences LLC for honors students. Mirroring Study 1, participants were entered into a drawing for college memorabilia for the 2019 cohort for completing the survey, while participants in the 2020 and 2021 cohorts received $10 and were also entered into a drawing for one of 3 $100 cash prizes.

**Procedure**

**Survey Timeline**

Similar to Study 1, all participants completed two surveys, once in the middle of the semester (4-6 weeks into the semester) and the other at the end of the semester (10-12 weeks into the semester). Participants in the control group and the honors LLC completed their survey outside of class time whereas those in the first-generation LLC completed their survey during their weekly seminar time.

** Intervention Group: First-Generation LLC**

All elements of the LLC were the same as to Study 1 (taking introductory biology and a weekly seminar together, peer advisors, bonding activity, residential component). The FG LLC also took Biology 152 (the second class in the introductory biology sequence) with the same instructor as those in the control group (i.e., same content). Due to the COVID-19 pandemic, the residential component was disrupted for the 2020 cohort (34% of the sample). The 2021 cohort (41% of the sample) lived on campus with other
students in the LLC but took Biology 152 remotely due to the COVID-19 pandemic. In contrast to Biology 151 (Study 1), Biology 152 did not have a discussion section where key concepts were reviewed. Furthermore, students took Biology 152 asynchronously during the second half of 2020. In 2021, one-third of class lectures each week were taught asynchronously and two-thirds were taught synchronously over Zoom.

**Control Group**

Students in the control group were FG students taking Biology 152 with the same instructor who also taught the FG LLC. The course structure, assignments, exams, and grading rubric used by this instructor were identical in both conditions. Thus, students in the control group also took 152 asynchronously in the second half of 2020 and in a hybrid format for 2021. Similar to the fall semester and similar to other students who take introductory biology, they took this introductory biology course in a large lecture class (i.e., 210 to 350 students).

**Honors LLC**

Similar to the FG LLC, the elements of the honors LLC remained the same as Study 1 (taking introductory biology and a weekly seminar together, residential component). Students in the honors LLC took a separate introductory biology course (Biology 162H), which used a team-based learning format and only enrolled honors LLC students. The 2020 cohort’s living community (38% of the sample) was partially disrupted; and the 2021 cohort (34% of the sample) lived on campus with the other members of the LLC, but took their introductory biology course in a synchronous remote class.

**Measures**
Measures were identical to Study 1, with the exception of belonging, anxiety, and motivation, which students answered about Biology 152/162H, rather than Biology 151/161H. All measures had adequate to high reliability (belonging: $\alpha_s \geq .74$, anxiety: $\alpha_s \geq .88$, motivation: $\alpha_s \geq .89$, confidence: $\alpha_s \geq .80$, real-world relevance of biology: $\alpha_s \geq .80$). We also examined participants’ grades by accessing their transcripts.

### Study 2 Results

**Data Analytic Strategy and Power Analysis**

Similar to Study 1, we conducted 3 Condition (control, FG LLC, honors LLC) x 2 Time ANOVAs for each of the dependent measures. If a significant main effect or interaction emerged, follow-up post hoc tests were conducted using Bonferroni corrections. Sensitivity power analyses (GPower 3.1; Faul et al., 2009) revealed 95% power (3 groups, 2 measurements across time) for an effect size as small as $d = 0.24$ (199 participants).

**Belonging**

Replicating Study 1, a significant main effect of condition, $F(2, 196) = 8.62, p < .001, d = 0.59$, showed that overall, FG students in the LLC ($M = 5.26, SE = 0.13$) felt a stronger sense of belonging in biology compared to FG students in the control group ($M = 4.64, SE = 0.16, p = .007, d = 0.31$), but were statistically similar to students in the honors LLC ($M = 5.46, SE = 0.13, p = .795$). Students in the honors LLC also reported significantly stronger belonging in biology than those in the control group ($p < .001, d = 0.41$). A significant main effect of time, $F(1, 196) = 5.68, p = .018, d = 0.33$, revealed that students’ feelings of belonging decreased from Time 1 ($M = 5.21, SE = 0.08$) to Time 2 ($M = 5.04, SE = 0.09$) across conditions. There was no significant time x condition
interaction, $F(2, 196) = 1.85, p = .160$. Means for each timepoint are graphed in Figure 19.

![Graph of Study 2 Belonging](image)

*Figure 19. Participants’ feelings of belonging in Biology 152 in Study 2.*

**Confidence**

A significant main effect of condition, $F(2, 195) = 8.74, p < .001, d = 0.61$, showed that the FG LLC ($M = 5.08, SE = 0.12$) reported more confidence in biology than the control group ($M = 4.60, SE = 0.15; p = .037, d = 0.25$). There was no difference in confidence between the FG LLC versus the honors LLC ($M = 5.40, SE = 0.12, p = .182$). The honors LLC also reported more confidence than the control group ($p < .001, d = 0.42$). The effect of time, $F(1, 195) = 0.01, p = .909$, and time x condition interaction were nonsignificant, $F(2, 195) = 2.50, p = .084$. Means for each timepoint are graphed in Figure 20.
Perceived Relevance of Biology in the Real World

A significant main effect of condition, $F(2, 191) = 8.00, p < .001, d = 0.58$, showed that students in the FG LLC ($M = 6.40, SE = 0.08$) perceived biology to have greater relevance in everyday life than the control group ($M = 6.01, SE = 0.10, p = .010, d = 0.30$). Perceptions were no different between the two LLC groups: FG LLC vs. honors LLC ($M = 6.53, SE = 0.09, p = .800$), though students in the honors LLC felt that biology was more relevant to everyday life than the control group ($p < .001, d = 0.39$). The effect of time, $F(1, 191) = 0.99, p = .320$, and time x condition interaction, $F(2, 191) = 0.48, p = .620$, were nonsignificant. Means for each timepoint are graphed in Figure 21.
Anxiety

Unlike Study 1, we did not find a significant main effect of condition, $F(2, 196) = 2.07, p = .129$. Instead, a significant main effect of time, $F(1, 196) = 20.08, p < .001, d = 0.65$, showed an overall increase in anxiety from Time 1 ($M = 3.65, SE = 0.09$) to Time 2 ($M = 4.08, SE = 0.10$), across conditions. Also, a significant time x condition interaction, $F(2, 196) = 9.60, p < .001, d = 0.62$, showed that both the FG LLC group (Time 1: $M = 3.27, SE = 0.15$; Time 2: $M = 4.24, SE = 0.17; p < .001, d = 0.44$) and the honors LLC group (Time 1: $M = 3.95, SE = 0.15$; Time 2: $M = 4.26, SE = 0.17; p = .045, d = 0.14$) increased in anxiety over time, while the control group did not significantly change over time ($p > .99$).

Motivation

Figure 21. Participants’ perceived relevance of biology in the real-world in Study 2.
We did not find a significant effect of condition for motivation, $F(2, 196) = 0.94, p = .394$. A significant main effect of time, $F(1, 195) = 20.60, p < .001, d = 0.64$, showed that student motivation decreased from Time 1 ($M = 5.68, SE = 0.06$) to Time 2 ($M = 5.34, SE = 0.09$) across conditions. A significant time x condition interaction, $F(2, 195) = 6.60, p = .002, d = 0.51$, revealed that FG LLC showed decreased motivation over time (Time 1: $M = 5.82, SE = 0.10$; Time 2: $M = 5.10, SE = 0.14, p < .001, d = 0.42$), while participants in the control group and the honors LLC did not change over time ($ps \geq .149$).

**Spring Semester Grades**

Similar to Study 1, we had planned to compare participants’ final grades in their second semester of biology between the control and the FG LLC conditions. However, due to the unprecedented circumstances of the COVID-19 pandemic and switch to remote learning, all students were given the option to change their letter grades to pass/fail grading to preserve their cumulative grade point average. This led 32% of the sample to choose the pass/fail option, reducing analyzable sample of letter grades in the spring semester (control: $N = 26$, FG LLC: $N = 44$). The reduced sample of grades and the self-selection of students opting out of letter grades, reduces the validity of grades. Thus, grades in biology were not analyzed for the spring semester.

**Persistence in Biological Science Majors Beyond the First Year of College**

At the conclusion of participants’ first year, the living learning community ended, but we continued to track students’ persistence in biological science majors in their sophomore year by accessing their college transcripts with their permission. Our goal was to determine whether participation in a FG LLC in the first year of college would
increase student retention in biological science majors relative to controls and the honors LLC up to one year later. For this analysis, we combined data from Studies 1 and 2, resulting in a total sample of 317 students (control: \( N = 109 \), FG LLC: \( N = 80 \), honors LLC: \( N = 128 \)). Student transcripts were coded; those who remained enrolled at the university and continued in a biological sciences major were coded as 1, while those who switched out of a biological science major or left the university were coded with a 0. A significant chi-square test, \( \chi^2(2) = 29.50, p < .001, \phi_c = 0.31 \), revealed significant differences across conditions. A larger percentage of students in the FG LLC (88%, 70 out of 80) remained biological science majors compared to the control condition (67%, 73 out of 109; \( \chi^2(1) = 10.56, p = .001, \phi_c = 0.24 \)). Student retention in the biological sciences was similar across the two LLC groups: FG LLC versus honors LLC (93%, 119 out of 128), \( (\chi^2(1) = 1.77, p = .183) \). Finally, retention was higher in the honors LLC compared to the control group (\( \chi^2(1) = 25.86, p < .001, \phi_c = 0.33 \)). Retention percentages are graphed in Figure 22.
Discussion

This present investigation is the first to examine whether a non-traditional living learning community designed specifically for first-generation undergraduate students sparks and sustains their engagement in the biological sciences, enhances performance, and persistence in biological science majors beyond their first year of college after the intervention has ended. We compared first-year students who took part in an affinity-based cohort model that combined intellectual community and residential life with peers who share a similar identity with other students who took the same classes in large lecture sections and lived in general residential housing.

Our present studies expand upon prior research in four important ways. First, unlike previous work on FG students in STEM, we employed an intervention in which the institution restructured FG students’ learning environment, rather than expecting FG students to change their mindset. Second, a vast majority of learning environment interventions in the past focused on women in STEM, while our present focus is on FG students. Third, these were the first studies to investigate the efficacy of an identity-based peer cohort for only FG students (as opposed to a mix of FG and CG students) in STEM. Finally, we examined students’ subjective academic experiences as well as their objective performance and retention in STEM, while prior work on FG students and LLCs only addressed performance and/or retention.

In Study 1, the first semester of college, immersion in a LLC with other FG peers significantly increased students’ feelings of belonging in biology, motivation, reduced their anxiety, and improved end-of-semester grades in biology by one-third of a letter
grade, from a B to a B+, compared to FG peers in the control group. By the end of the second semester of college, Study 2 continued to show substantial benefits of the LLC for FG students despite the significant disruption in residential life and teaching created by the COVID-19 pandemic. First-generation students in the LLC continued to feel a greater sense of belonging in biology, more confidence, and believed biology to be more socially important in everyday life in comparison to the control group. A substantially larger percentage of FG LLC students persisted in biological science majors in their sophomore year compared to FG students who were not in a LLC.\(^9\) Importantly, both studies converge to show that immersing FG students in a LLC closes the gap between them and honors students who, on average, started college with more social capital in terms of family resources and high school preparation.\(^{10}\) These findings suggest that a LLC can minimize and perhaps erase achievement gaps between FG and CG students in college.

Though the LLC revealed clear benefits for FG students in both studies, a few inconsistencies emerged between the two studies. First, the effect of the LLC on reducing student anxiety and increasing motivation found in Study 1 did not replicate in Study 2. We speculate this may have occurred because of the disruption created by the COVID-19 pandemic during spring 2020, as suggested by the overall increase in anxiety and decrease in motivation revealed in our data during that semester. Recall that due to the COVID-19 pandemic, students had to leave their on-campus residential learning communities at short notice and their biology course migrated online to an asynchronous

---

\(^9\) The percentage of FG students in the control condition who remained in the biological sciences during their second year (67%) was comparable to the second-year retention rate for FG students in biological sciences who entered college during the years of 2013-2017 (69%).

\(^{10}\) For specificity, all analyses with the honors LLC condition (psychological experiences and retention) were also conducted with only the continuing-generation students in the honors college LLC. Results remained the same, regardless of whether or not the first-generation students in the honors college LLC (n = 13) were included in the analyses.
format. Given the pre-pandemic emphasis on communal living and in-person peer community in the LLC, its abrupt interruption is likely to have contributed to increased anxiety and decreased motivation especially in this group of first-generation students.

A second inconsistency was that in Study 2, the FG and honors LLC reported significantly greater confidence and recognized the real-world relevance of biology compared to the control group, though these groups did not vary in Study 1. Our speculation is that LLC students’ confidence about their ability in biology may have grown over time as they became more comfortable with college-level classes, hence the delayed effect. Similarly, the weekly LLC seminars in the fall semester provided exposure to faculty research and the social impact of that research, which may have helped protect their beliefs in the relevance of biology in the real-world.

Future research is necessary to expand on our present findings. Our LLC intervention combined several elements that cannot be teased apart in our studies: small class size, learning in an identity-based peer community, residential community with first-generation peers, peer mentors, and community-building social events. An important direction for future research is to elucidate which elements of an identity-based living learning community is most important for minority students and which other elements are discretionary.

Another important future direction is to examine the generalizability of these findings to other identity groups that are marginalized in STEM (e.g., Black, Hispanic, and Indigenous students). It is also important to test the long-term effects of identity-based LLCs beyond the first two years of college. In our studies, we find that the benefits of a one-year identity-based LLC program linger throughout students’ first year and
beyond, up to one year after the program has ended. However, longer-term benefits that span the entire college experience of 4-6 years are yet unknown. This is an important avenue for future research.

In sum, it is of great societal importance to design and rigorously test research-informed solutions to promote the success of underrepresented students in STEM and mitigate disparities between them and more advantaged students. Our investigation reveals empirical evidence of the benefits of a peer-based cohort for underrepresented minorities in STEM, which universities should continue to test and implement. To truly promote diversity and diminish social disparities in STEM, it is the responsibility of institutions to create welcome learning environments that signal to minorities that they belong and that their success is valued, rather than expecting them to adjust on their own. Thus, implementing identity-based LLCs and similar interventions that expand the presence of ingroup peers for minorities in STEM are crucial for future progress towards equity in STEM.
### Supplemental Analyses

Table 3.  
*Descriptive Statistics of Survey Variables*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Control</th>
<th>FG LLC</th>
<th>Honors LLC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time 1</td>
<td>Time 2</td>
<td>Time 1</td>
</tr>
<tr>
<td><strong>Study 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>4.83 (1.11)</td>
<td>4.98 (1.29)</td>
<td>5.60 (1.23)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4.00 (1.34)</td>
<td>4.14 (1.35)</td>
<td>3.34 (1.41)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.36 (1.08)</td>
<td>5.17 (1.25)</td>
<td>5.69 (1.11)</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.94 (1.17)</td>
<td>4.80 (1.37)</td>
<td>4.98 (1.18)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>6.21 (0.92)</td>
<td>6.38 (0.79)</td>
<td>6.37 (0.77)</td>
</tr>
<tr>
<td><strong>Study 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>4.78 (1.05)</td>
<td>4.51 (1.24)</td>
<td>5.38 (1.14)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>3.74 (1.24)</td>
<td>3.74 (1.33)</td>
<td>3.27 (1.36)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.55 (0.95)</td>
<td>5.33 (1.12)</td>
<td>5.82 (0.92)</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.62 (1.24)</td>
<td>4.58 (1.28)</td>
<td>5.15 (1.18)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>6.04 (0.93)</td>
<td>5.97 (1.03)</td>
<td>6.45 (0.70)</td>
</tr>
</tbody>
</table>
Table 4.
*Condition Differences of Survey Variables Separated by Cohort (Having In-Person or Virtual Courses)*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Descriptive Statistics by Condition</th>
<th>Effect Size of Condition Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (M(SD))</td>
<td>FG LLC (M(SD))</td>
</tr>
<tr>
<td><strong>Study 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Fall 2018 &amp; Fall 2019 (N = 159; In-Person)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>5.16 (1.10)</td>
<td>5.61 (1.10)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4.03 (1.24)</td>
<td>3.52 (1.24)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.36 (1.06)</td>
<td>5.72 (1.06)</td>
</tr>
<tr>
<td>Confidence</td>
<td>5.03 (1.04)</td>
<td>5.05 (1.04)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>6.33 (0.66)</td>
<td>6.34 (0.66)</td>
</tr>
<tr>
<td><strong>Fall 2020 (N = 84; Virtual)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>4.26 (0.94)</td>
<td>5.36 (0.94)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4.17 (1.15)</td>
<td>3.31 (1.15)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.02 (0.95)</td>
<td>5.67 (0.95)</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.49 (1.06)</td>
<td>4.88 (1.06)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>6.19 (0.66)</td>
<td>6.45 (0.66)</td>
</tr>
<tr>
<td><strong>Study 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Spring 2019 &amp; Spring 2020</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>4.84 (1.13)</td>
<td>5.34 (1.13)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>3.46 (1.23)</td>
<td>3.85 (1.23)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.56 (0.83)</td>
<td>5.51 (0.83)</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.89 (0.99)</td>
<td>5.22 (0.99)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>6.04 (0.67)</td>
<td>6.46 (0.68)</td>
</tr>
<tr>
<td><em>Spring 2021 (N = 73; Virtual)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belonging</td>
<td>4.27 (1.02)</td>
<td>5.15 (1.02)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4.27 (1.12)</td>
<td>3.61 (1.12)</td>
</tr>
<tr>
<td>Motivation</td>
<td>5.21 (1.00)</td>
<td>5.39 (1.00)</td>
</tr>
<tr>
<td>Confidence</td>
<td>4.02 (1.04)</td>
<td>4.88 (1.04)</td>
</tr>
<tr>
<td>Real-World Relevance</td>
<td>5.94 (0.77)</td>
<td>6.31 (0.77)</td>
</tr>
</tbody>
</table>

---

11 Values are averaged across both timepoints in each study.

12 The Spring 2020 semester was in-person for the first half of the semester, and virtual for the second half.
Race

Though the target of our intervention was first-generation (FG) students, it is also important to address intersecting social identities. At the national level, 25% of White and Asian American students are FG students, while 41% of Black and 61% of Hispanic students are first-generation (NCES, 2014). Since Black and Hispanic individuals have historically been underrepresented in STEM (National Academies of Science, Engineering, and Medicine, 2019; National Center for Science and Engineering Statistics, 2021), it is important to address whether interventions that target students’ class identity benefit underrepresented racial minority (URM) students as well. Thus, we tested whether our LLC would reveal benefits based on students’ intersectional identity—FG students who were also from underrepresented racial minority groups (i.e., Black or Hispanic) as compared to FG URM students who were not in the LLC.\(^{13}\)

In Study 1, mirroring our results with the full sample, FG URM students in the LLC reported greater belonging (FG LLC: \(M = 5.58\), Control: \(M = 4.84\); \(F(1, 48) = 4.63, p = .037\)), lower anxiety (FG LLC: \(M = 3.01\); Control: \(M = 4.11\); \(F(1, 48) = 7.58, p = .008\)), and greater motivation (FG LLC: \(M = 5.88\); Control: \(M = 5.18\); \(F(1, 48) = 4.15, p = .047\)) than FG URM students who were not in a LLC. For Study 2, while the results were in the same direction (i.e., FG URM students in the LLC reporting better academic experiences than those not in the LLC), these condition differences did not reach significance, \(ps \geq .154\). One reason for this discrepancy may be due to low statistical power. In the control condition, while 16 out of 63 (25%) participants were URM students in Study 1, only 10 out of 50 (20%) were URM students in Study 2. In comparison, because we

\(^{13}\) We did not include URM students from the honors LLC in this analysis as there were too few URM students in this condition (Study 1: 6 out of 94; Study 2: 5 out of 73).
oversampled URMs for the FG LLC, 40-41% of the FG LLC condition were URMs (Study 1: 34 out of 86; Study 2: 31 out of 76). Furthermore, for retention in the biological sciences during their sophomore year, 78% (25 out of 32) of URMs in the FG LLC remained in their biological sciences major, compared to 50% (14 out of 28) of FG URMs in the control condition, $\chi^2(1) = 5.19, p = .023$. In sum, we offer preliminary evidence that a LLC designed for FG students may have benefits for FG students who are also members of URM groups, though future research should expand on these findings with larger samples.

**Gender**

Along with race, we also examined whether women FG students also benefitted from a LLC designed for FG students. Women have also historically been underrepresented in STEM fields, though they currently are no longer underrepresented in the biological sciences, now making up 60% of Bachelor’s and master’s degrees in the biological sciences (NSF, 2018). This was mirrored in our sample, as women made up 67% of our total sample in both studies (62-63% in FG LLC condition; 80% in control condition; 62-63% in honors LLC condition). While women are no longer underrepresented, they may still face negative stereotypes regarding their abilities and intelligence in biology (e.g., Kerkhoven et al., 2016; Lane et al., 2012). Therefore, it remains essential to address whether our intervention for FG students also has benefits for women FG students.

Similar to Study 1, a significant effect of condition of belonging emerged, $F(2, 159) = 6.33, p = .002$, as women FG students in the LLC ($M = 5.44$) reported greater belonging than women FG students who were not in the LLC ($M = 4.77; p = .007$).
Women students in the honors LLC ($M = 5.43$) reported similar belonging to those in the FG LLC ($p > .99$) and had significantly greater belonging than FG students not in the LLC ($p = .006$). There was also a significant condition effect for anxiety, $F(2, 159) = 6.83$, $p = .001$, such that women FG students in the LLC ($M = 3.61$) reported lower anxiety than those in the control condition ($M = 4.20$, $p = .043$) and the honors LLC condition ($M = 4.44$, $p = .001$). There were no differences between the control and the honors LLC condition ($p = .943$). Finally, the condition effect for motivation did not reach significance ($p = .107$), though women in the FG LLC descriptively reported greater motivation ($M = 5.56$) than those in the control ($M = 5.12$) and honors LLC ($M = 5.29$) conditions.

In Study 2, there was also a significant effect of condition for belonging, $F(2, 130) = 3.96$, $p = .021$. Though women in the FG LLC ($M = 5.20$) descriptively reported greater belonging than those in the control condition ($M = 4.73$), this difference did not reach significance ($p = .135$). Women in the honors LLC condition ($M = 5.37$) reported greater belonging than those in the control condition ($p = .021$) but were statistically similar to those in the FG LLC ($p > .99$). For confidence, a significant overall effect of condition, $F(2, 129) = 7.94$, $p = .001$, revealed that women in the FG LLC ($M = 5.18$) reported greater confidence than those in the control condition ($M = 4.37$, $p = .002$). Women in the honors LLC ($M = 5.19$) also reported greater confidence than those in the control condition ($p = .002$) and were statistically similar to those in the FG LLC ($p > .99$). Finally, there was also a significant effect of condition for perceived real-world relevance of biology, $F(2, 127) = 4.83$, $p = .010$. Women in the FG LLC ($M = 6.53$) reported greater perceptions of relevance than those in the control condition ($M = 6.05$; $p$
= .013). Women in the honors LLC \((M = 6.46)\) also reported greater perceptions than the control condition \((p = .042)\) and were statistically similar to those in the FG LLC \((p > .99)\).

For retention in the biological sciences, a significant chi-square test revealed differences across conditions, \(\chi^2(2) = 22.77, p < .001\). A greater percentage of women in the FG LLC (90%; 46 out of 51) remained in their biological sciences major in comparison to FG women in the control condition (66%; 54 out of 82; \(\chi^2(1) = 9.99, p = .002\)). Women in the honors LLC (93%; 69 out of 74) also remained in the biological sciences to a greater degree than women in the control condition, \(\chi^2(1) = 17.50, p < .001\). Retention rate did not differ between the two LLC conditions, \(p = .537\).

In sum, we found very similar effects for women, such that a LLC designed for first-generation students had benefits for women students, despite not specifically focusing on their gender. Future research should attempt to replicate whether designing LLCs or other learning communities specifically for members of one stigmatized social identity can also benefit members who share multiple stigmatized social identities.
CHAPTER 5

GENERAL DISCUSSION

Summary of Key Findings

This dissertation presents work that focuses on the experiences of underrepresented college students in STEM, specifically examining the impact of negative stereotypes on their experiences of social identity threat (SIT), and whether this threat can be mitigated through learning environment interventions. Across three manuscripts, this dissertation examined these concepts on a micro and macro level. On a micro level, the first manuscript studied how SIT manifests in real-time, finding that a subtle gender stereotypic cue elicited increased neurological vigilance on a second-by-second basis among in women invested in STEM. On a macro level, I examined how yearlong interventions mitigated the negative effects of SIT on minorities’ real-world experiences and outcomes in STEM over time. Specifically, the second manuscript investigated how female peer mentors have benefits for women’s academic experiences and outcomes in engineering throughout college, while the third manuscript tested how a living learning community for first-generation students has benefits for first-generation students’ academic experiences and outcomes in biology in their first two years of college. In the section below, I briefly summarize the main findings from each manuscript.

In Chapter 2, I presented a lab study that examined whether women would be more vigilant to gender stereotypic cues while completing a math-related task. In this study, men and women who were STEM majors were asked to complete a numerical Stroop task (framed as a math intelligence task) while their neural activity (measured via
electroencephalogram or EEG) was measured (framed as neural markers of math intelligence). In the task, before each trial, participants viewed either a picture of a male watching face or a control image for 90 milliseconds. The ERN (error-related negativity) component was the outcome of interest, which is a neurological component that captures attentional vigilance to errors 50 to 100 milliseconds after an error commission. While men displayed similar ERN responses to the male faces and the control images, women displayed larger ERN responses or increased attentional vigilance to the male faces in comparison to the control images. Furthermore, this priming difference among women was driven by women who were highly invested in pursuing STEM careers, as women who were less interested did not show an ERN difference between primes. Altogether, this first manuscript showed that subtle stereotypic cues can elicit preconscious vigilance on a momentary basis for members of stigmatized groups when in an environment that is highly relevant to their personal interests.

While Chapter 2 provided evidence that even subtle stereotypic environments can elicit increased vigilance on a second-by-second basis, Chapters 3 and 4 presented longitudinal field studies that rearranged environments to alleviate the negative effects of the vigilance or threat. In Chapter 3, I examined whether peer mentorship had long-lasting benefits for college women in engineering. In this field experiment, women in their first year of college were randomly assigned a female peer mentor, male peer mentor, or no mentor. Peer mentors were third or fourth years in engineering. Women’s academic experiences and outcomes in engineering were measured each following year until graduation. While women with a male mentor or no mentor increased in their feelings of anxiety relative to motivation and decreased in their aspirations to pursue
post-graduate engineering degrees and emotional well-being, women with a female mentor remained steady on these outcomes. Furthermore, a greater percentage of women with a female mentor obtained an engineering internship in college and graduated with a STEM degree, in comparison to women with a male mentor or no mentor. This intervention showed that developing an authentic relationship with a near peer mentor who shares one’s stigmatized identity during a key transitional period (the first year of college) can have long-lasting benefits for individuals’ academic experiences, health, and retention in an environment where they are underrepresented and negatively stereotyped.

In a similar fashion, Chapter 4 also examined whether a peer-based intervention during the first year of college helped underrepresented students in STEM. In contrast to Chapter 3, the pair of studies in Chapter 4 investigated whether an identity-based living learning community (LLC) for first-generation (FG) students had benefits for their academic experiences, performance, and retention in the biological sciences. In this field quasi-experiment, there were three groups: 1) an intervention group with FG students in a biological sciences LLC; 2) a control group with FG students who were not in the LLC, and 3) an additional comparison group with honors college students, mostly from college-educated families, who participated in a separate biological sciences LLC. In Study 1 (fall semester), FG students in the LLC reported greater belonging and motivation to persist, lower anxiety, and better grades than FG students who were not in the LLC. In Study 2 (spring semester), FG students in the LLC reported greater belonging, confidence, and perceptions of the real-world relevance of biology than those in the control group. Across both semesters, these academic experiences in biology were similar for FG and honors students in LLCs. In their second year, a larger percentage of
FG and honors students remained in biological sciences majors than FG students who were not in a LLC. In sum, these two studies provide evidence that learning and living with a cohort of peers who shared the same stigmatized identity can benefit individuals’ academic experiences, performance, and retention in fields where they face underrepresentation.

**Theoretical Implications and Applied Considerations**

All three investigations demonstrate that situational cues have profound impacts on the well-being and success of stigmatized individuals. In Chapter 2, a subtle stereotypic situational cue elicited increased vigilance in minority students. In Chapters 3 and 4, creating situations where beginner students who are underrepresented in STEM could develop authentic relationships and learn with others like themselves improved their experiences.

The power of situations is directly related to our observations of our environments. For example, we learn stereotypes by observing society. We notice that social groups are regarded in different ways, that some hold more power and status than others, and that specific groups are expected to fulfill certain roles and expectations (Dasgupta, 2013). Groups that are associated with increased power and status are expected to fulfill professional and technical roles while groups that are lower in power and status are often expected to hold domestic and communal roles (Banaji & Hardin, 1996; Blair & Banaji, 1996). Because professional and technical roles are also regarded more highly, jobs in this domain are better paid and more accessible to groups of higher status (Budig, 2002; Mitra, 2003). As groups of lower societal status are not expected or welcomed to hold professional and technical jobs, this leads to underrepresentation and
negative stereotypes that call their abilities into question when they attempt to enter these fields, as well as decreased wealth (e.g., Allen et al., 2015; Dasgupta & Stout, 2014; Leslie et al., 2015; Phelan et al., 2010; Steele & Aronson, 1995).

Underrepresentation, negative stereotypes, and wealth inequality not only lead to increased health and well-being inequality (Elgar, 2010) but also contribute to labor shortages for critical technical jobs (President’s Council of Advisors, 2012; Zaza et al., 2020). Thus, while STEM jobs have traditionally favored White and Asian men, it is essential to change societal norms to promote the inclusion of women, Black, Hispanic, and Indigenous Americans. In applying this to research, it is important to identify when situations signal to minorities that they do not belong (Chapter 2) and how to rearrange situations to actively promote the inclusion and success of minorities (Chapters 3 and 4).

Prior research and theory regarding stereotypes in STEM have discussed the near-automatic nature of these stereotypes (e.g., Lane et al., 2012; Smeding, 2012). Because of the continued societal observations, these associations (e.g., that STEM is associated with White and Asian males) can become ingrained, even for members of minority groups (Lane et al., 2012; Rogers et al., 2021). These implicit associations have been captured using Implicit Association Tests (IATs), a rapid reaction time task, in which even women pursuing STEM are faster to associate men (rather than women) with science and math (e.g., Steele & Ambady, 2006; Stout et al., 2011). While these behavioral tasks measure individuals’ automatic behavioral responses, neurological measures can capture individuals’ preconscious processing of cues. Within the context of minorities in STEM, understanding subtle situational cues that provoke a threat response from minorities in STEM can help us identify solutions to mitigate these threatening responses.
The investigation in Chapter 2 expands upon prior theory by exploring the boundary conditions necessary to elicit a vigilant response from minorities in STEM. Specifically, do women display more attentional vigilance on a speeded math task when a prime signaling gender is only shown for 90 milliseconds? Also, is the vigilance momentary – is it only present for the trials that have this stereotypic cue and absent on trials where there is a neutral cue? The results of Chapter 2 – that women dedicated to STEM do display more attentional vigilance only on trials with a subtle gender cue – reveal that people’s minds are able to quickly identify threats when they are in a context that they are dedicated to, even preconsciously. Beyond this theoretical contribution, Chapter 2’s investigation also has implications for applications. Because women were more vigilant to the presence of male faces, and as males are high-status within STEM fields, the presence of females may help assuage this vigilance. The presence of ingroup members who are also negatively stereotyped in STEM fields may function as a comforting or welcoming signal for underrepresented minorities in STEM.

Chapter 3 examines whether having a peer mentor sends a welcoming signal to beginners entering a STEM field, and if female or male peer mentors have different impacts on women’s academic experiences and success in STEM. In a similar vein, Chapter 4 investigates whether first-generation students who live and learn in an affinity-based community with other first-generation students have more positive experiences than first-generation students who do not have this community. Both of these interventions were informed by the Stereotype Inoculation Model (Dasgupta, 2011), a theoretical model which posits that situational factors such as exposure and relationships with ingroup peers and experts can protect minority members’ feelings of belonging,
confidence, motivation, and future aspirations in fields where they face negative stereotypes due to their social identities. Other than the Stereotype Inoculation Model, brief mental reappraisal interventions (e.g., value affirmation, Cohen et al., 2009; Harackiewicz et al., 2014; Hecht et al., 2021; Miyake et al., 2010; Shnabel et al., 2013; Walton et al., 2011; growth mindset, Paunesku et al., 2015; Yeager & Dweck, 2012; Yeager et al., 2016; 2019; utility value, Harackiewicz et al., 2016; and emotion regulation, Rozek et al., 2019) have been proven to decrease gender, race, and class disparities in academic performance in STEM fields.

The investigations in Chapters 3 and 4 build on the Stereotype Inoculation Model and these prior interventions for minorities in STEM in four important ways. First, though past work on ingroup role models have benefits for women’s experiences and performance in STEM, much of this research was conducted in artificial lab settings and did not create authentic relationships (e.g., Cheryan et al., 2011; 2013; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013). Chapters 3 and 4 both present field interventions, in which students developed authentic relationships with their peer mentors (Chapter 3) and their peer cohort (Chapter 4). Second, previous interventions addressing disparities in STEM assessed impacts at a single time-point or over a relatively short period of time (e.g., Cheryan et al., 2011; 2013; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2011; Ramsey et al., 2013; Shapiro et al., 2013); thus, Chapter 3 examines the impact of an intervention over the course of 4-6 years. Chapter 4 also investigates whether a yearlong intervention has benefits for STEM retention a year after the intervention ends. Third, many field interventions that do focus on mitigating STEM disparities only report measures of academic performance or
retention, without examining students’ actual academic experiences (e.g., Harackiewicz et al., 2016; Paunesku et al., 2015; Rozek et al., 2019; Walton et al., 2015; Yeager et al., 2016; 2019). The interventions in Chapters 3 and 4 investigate academic experiences in addition to performance and retention. Fourth, interventions that require students to mentally reappraise their situations places the burden on underrepresented individuals to adjust to academic institutions by changing their mindset rather than placing the burden on academic institutions to change learning environments to fit the needs of diverse students. Since universities possess more resources to create change, the responsibility for change should reside with institutions. Therefore, our interventions in Chapters 3 and 4 create welcoming learning environments for underrepresented minorities, rather than requiring students to mentally reappraise their circumstances.

Chapters 3 and 4 provide evidence that interventions fostering peer relationships and community among underrepresented students in STEM have long-lasting benefits. These types of interventions should be implemented in future research. Additionally, both of these interventions were conducted during a transitional period, specifically, the first year of college. During the first year of college, students are adjusting to a new environment and new styles of learning. In transitional periods, building communities and relationships for underrepresented minorities can help build a foundation that will set them up for long-term success.

**Directions for Future Research**

There remain unanswered questions that should be addressed in future research. First, though peer-based interventions (such as Chapters 3 and 4) that took place during transitional periods succeeded and had long-term effects, the mechanisms underlying the
success of interventions during transitional periods have not been explored. Two potential mechanisms that should be explored are informed by attachment theory (e.g., Bowlby, 1969) and social baseline theory (Beckes & Coan, 2011; Coan & Sbarra, 2015). The literature on attachment theory has mainly focused on the relationships between children and their parents and between romantic partners, with the former informing the latter (Bartholomew & Shaver, 1998; Berlin et al., 2008; Fraley & Davis, 1997; Rothbaum et al., 2002). Attachment theory posits that children develop secure or insecure attachments to their parents, which is often reflected in their adult romantic relationships (Berlin et al., 2008; Roisman et al., 2001). In other words, childhood is a key transformational period, which is why the attachment style developed during this period can be long-lasting. The first year of college may also serve as a key transformational period, during which individuals’ attachment style may change (Lopez & Gormley, 2002), as students are developing many new relationships in a new and unfamiliar environment. Peer-based interventions for underrepresented minorities may aid students in forming strong close relationships, helping them to either become securely attached (from insecurely attached) or to maintain their secure attachment style. Students’ secure attachment may have downstream effects on their college experience. For example, just as securely-attached children are more confident in exploring unfamiliar new environments than anxiously or avoidantly-attached children, securely-attached students may take risks and explore the unfamiliar university environment with greater confidence.

The formation of these close and secure relationships may also have a direct effect on student stress and well-being. Social baseline theory posits that the benefits that close relationships have on well-being and health are due to close proximity and social
interactions (Beckes & Coan, 2011; Coan & Sbarra, 2015). Specifically, the presence of others can decrease perceived and neurological stress and threat (e.g., Coan et al., 2006), helping individuals retain their ‘baseline’ calm state when placed in stressful situations. Applied to peer-based interventions during the first year of college, underrepresented students who are supported or surrounded by peers who share the same stigmatized identity may feel less stressed during their transition to college, helping them maintain their interests and goals to succeed in college.

Taking attachment theory and social baseline theory together, increasing the presence of ingroup peers for underrepresented minority students in STEM during their first year of college may create an environment where STEM and university stressors are more easily manageable because students form more close relationships in which they feel supported and understood. These close relationships in turn can increase feelings of security, confidence, belonging, as well as decrease stress and anxiety in their academic experiences during their first year of college. Having this stable foundation during their first year of college may then translate to long-term benefits on academic outcomes, similar to how childhood attachment style affects adult relationships.

Another avenue for future research is examining the efficacy of identity-based interventions for students with multiple underrepresented social identities. Almost all interventions that employ an ingroup peer or mentor for underrepresented minorities have focused on one specific identity. There are a few exceptions. Research with Black and Latina women in STEM found that when briefly exposed to hypothetical role models that were White men or women or Black/Latinx men or women, Black and Latina women felt more belonging and trust when exposed to the Black/Latinx man or woman (Johnson et
Thus, for women who are also underrepresented racial minorities (URMs), having a same-race mentor may be more effective than having a same-sex mentor. However, this does not necessarily mean that interventions based solely on gender or generation status may not be effective for individuals who are also URMs. For example, while the intervention in Chapter 5 focused exclusively on first-generation students, FG URMs in the LLC reported greater belonging and motivation, lower anxiety, and remained in biological sciences majors at a higher rate in comparison to FG URMs who were not in the LLC. However, the percentage of URMs within the FG LLC (40%) was also much higher than the university average (17%), which may have provided URMs in the LLC with increased opportunity to develop authentic relationships with other URM peers. It is possible that part of the LLC benefits for FG URM students in Chapter 4 might have occurred due to the critical mass of URMs in the LLC. In sum, it is imperative that future work on identity-based interventions should more actively address intersectional identities as well as explore which types of interventions have the greatest benefits for individuals with multiple minority identities.

Conclusion

The underrepresentation of women, ethnic racial minorities, and first-generation students in STEM in the United States as well as the consequences of this underrepresentation (i.e., income inequality, labor shortage) have been known for decades. While society has progressed, there is still much more to be done. A growing body of research that employs rigorous, theory-driven interventions has been shown to provide benefits for minorities’ experiences, performance, and retention in STEM. One particular effective intervention is increasing the presence of ingroup minority peers in
learning environments. The increased presence of ingroup peers leads to the development of authentic relationships that can protect minorities as they navigate fields where they are underrepresented and negatively stereotyped. Institutions and universities should seek to implement these interventions in order to contribute to the efforts for societal equality.


Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances, 5*(2), eaau4734.


Outcome expectations: The role of living-learning programs and other college environments. *Research in Higher Education, 54*(8), 851-873.


Pietri, E. S., Johnson, I. R., & Ozgumus, E. (2018). One size may not fit all: Exploring how the intersection of race and gender and stigma consciousness predict


