

The Pennsylvania State University

The Graduate School

**A LATENT CLASS ANALYSIS APPROACH TO THE IDENTIFICATION OF
DOCTORAL STUDENTS AT RISK OF ATTRITION**

A Thesis in

Psychology

by

Samantha M. Stevens

© 2021 Samantha M. Stevens

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2021

The thesis of Samantha M. Stevens was reviewed and approved by the following:

Jonathan E. Cook
Associate Professor of Psychology
Thesis Adviser

Jes L. Matsick
Assistant Professor of Psychology and Women's, Gender, & Sexuality Studies

Dawn P. Witherspoon
Associate Professor of Psychology

Kristin A. Buss
Professor of Psychology and Human Development and Family Studies
Head of the Department of Psychology

ABSTRACT

To advance our understanding of doctoral student experiences and the high attrition rates among STEM doctoral students, we examined the psychological profiles of different types of doctoral students and conducted a preliminary examination of the effects of two social psychological interventions. We used latent class analysis (LCA) on self-reported psychological threat data from 1081 incoming doctoral students across three universities and found that the best-fitting model delineated four threat profiles: *Engaged and Confident*, *Relaxed and Nonchalant*, *Engaged but Worried*, and *Disengaged and Worried*. These profiles predicted outcomes measured at the beginning and end of students' first semester of graduate school that may influence attrition risk, including differences in academic preparation (e.g., prior attainment of a master's degree), self-evaluations (e.g., academic self-control), attitudes towards graduate school and academia (e.g., burnout), and interpersonal relations and perceived fit (e.g., sense of belonging). The demographic distributions of profiles also differed, with groups more likely to face social identity threat (e.g., women) being overrepresented in higher threat profiles (i.e., *Engaged but Worried* and *Disengaged and Worried* students) and underrepresented in lower threat profiles (i.e., *Engaged and Confident* and *Relaxed and Nonchalant* students). Moreover, students completed randomly assigned belonging intervention, affirmation intervention, or control writing exercises early in their first semester and we preliminarily examined effects of each intervention by threat profile on outcomes at the end of the first semester. We did not find robust evidence for intervention efficacy overall or difference in efficacy by class. We conclude that LCA may be useful to identify students at high risk of negative outcomes and that future work should further investigate the appropriateness of tailoring interventions by student threat profile to ultimately promote retention.

TABLE OF CONTENTS

List of Tables.....	v
List of Figures.....	vi
Acknowledgements.....	vii
Chapter 1. INTRODUCTION.....	1
Chapter 2. METHOD.....	14
Chapter 3. RESULTS.....	41
Chapter 4. DISCUSSION.....	85
Chapter 5. CONCLUSION.....	96
References.....	97

LIST OF TABLES

Table 1. Demographic Information.....	16
Table 2. Descriptive Statistics on Concurrent and Distal Outcomes.....	33
Table 3. LCA Fit Indices.....	45
Table 4. Classification Probabilities for the Most Likely Latent Class Membership (Column) by Latent Class (Row).....	46
Table 5. Parameter Estimates from the Selected 4-Class Model.....	48
Table 6. Proportions of Demographic Groups by Latent Class	52
Table 7. How Class Membership Predicts Descriptive Outcomes.....	55
Table 8. End-of-Semester Differences by Latent Class as a Function of Intervention Condition, Controlling for Baseline Levels.....	67

LIST OF FIGURES

Figure 1. Distribution of Indicator Items by Condition.....	25
Figure 2. Distribution of Indicators by Class.....	49
Figure 3. Academic Preparation and Context Outcomes by Class.....	58
Figure 4. Attitudes towards Graduate School and Academia Outcomes #1 by Class ...	59
Figure 5. Attitudes towards Graduate School and Academia Outcomes #2 by Class ...	60
Figure 6. Interpersonal Relations and Perceived Fit Outcomes by Class.....	61
Figure 7. Self-Evaluations by Class.....	62
Figure 8. Mental Health, Social Identity Threat, and Research Effort by Class	63

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Award No. 1661214.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

CHAPTER 1. INTRODUCTION

Nearly half of doctoral students in Science, Technology, Engineering, and Mathematics (STEM) fields leave school before graduating (National Research Council, 2011). While there are various possible explanations for any given individual deciding to terminate their studies before finishing (e.g., financial barriers), one possibility lies in the experiences of psychological threat that students may face, perhaps some more so than others. Given that these students are poised to contribute significantly to advancing knowledge as the next generation of scholars and technological innovators and the uncertainty about why so many prematurely terminate their studies, we need basic research on doctoral student experiences. Moreover, there is a need for research that identifies boundary conditions for belonging and values-affirmation interventions, including when, how, and for whom they work (Walton & Yeager, 2020); this is especially pressing given such interventions have been shown to affect students differently, in some cases backfiring on certain groups (Binning & Browman, 2019). Identifying boundary conditions can contribute to theory, conserve resources, and help to address issues of replicability. To address these gaps in the literature, this paper outlines the use of a person-centered approach (latent class analysis) to identify how incoming doctoral students vary by psychological threat, how membership in different threat subgroups predicts various concurrent and distal outcomes that may reflect attrition risk, and for whom belonging and values-affirmation interventions work by the end of the first semester.

Psychological Threat

Identified structural factors (e.g., financial support) have been shown to account for relatively little of the variance in Ph.D. student attrition rates (Groen et al., 2008). One explanation for the fact that the rate of STEM doctoral student attrition is disproportionately high for women and members of underrepresented racial/ethnic minority (URM) groups (National

Research Council, 2011) may lie within the differences in levels or types of psychological threat, that is, perceptions of a context's demands as beyond one's resources or abilities (Blascovich & Mendes, 2001). For instance, while many doctoral students might have some amount of worry that they will be perceived as incompetent or unintelligent by their peers, some may worry that they will be perceived as unintelligent due to their race or gender—in this case, we see the distinction between general psychological threat (i.e., not based on group membership) and social identity threat (i.e., based on group membership). General psychological threat and the additional psychological threat that women and URM students face, that is, social identity threat, may help to explain the high rates of attrition from STEM Ph.D. programs in general, as well as the particularly high rates for women and URM students.

High rates of in attrition in general, and the higher attrition rates among women and members of underrepresented racial/ethnic groups (National Research Council, 2011), may be attributed in part to the fact that doctoral education challenges students with many new demands and opportunities for failure, and that adverse psychological experiences (e.g., stress, doubt) are common in Ph.D. students. Research at earlier levels of education demonstrates that many students experience academic environments as psychologically threatening and that this psychological threat can account for academic achievement disparities between groups (Cohen et al., 2006; Sherman et al., 2013; Walton et al., 2015; Walton & Cohen, 2007, 2011; Walton & Yeager, 2020; Yeager et al., 2014b). Intuitively, it would make sense for doctoral education to be threat-provoking as well, given that doctoral students face a barrage of novel challenges with few clear milestones and far less reassurance than is received at lower levels of education that they are on the right track to succeed. Students at this level of education face more failure, more rejection, and more critical feedback alongside new professional and social norms. Overall,

doctoral students face much ambiguity that may be perceived as threatening, including the uncertainty associated with attaining long-term goals. However, there is little research on the psychological experiences of doctoral students, particularly with an eye towards understanding why the attrition rate is so high. The extant literature includes mostly correlational work on structural factors, and although these significant barriers require redress, they are not sufficient explanation by themselves for doctoral student attrition (Ehrenberg et al., 2007; Groen et al., 2008; Lovitts & Nelson, 2000). As such, there is a gap in our understanding of the psychological experiences of doctoral students and how they relate to attrition, which is notable given doctoral students are poised to become leaders in innovation and can contribute to a highly skilled workforce, particularly those in STEM fields.

While doctoral education is likely threatening for a range of individuals, it may be particularly threatening for students who contend with social identity threat, which could help explain the particularly high attrition rates among women and members of historically underrepresented racial and ethnic groups. Broadly defined, social identity threat (SIT) refers to worry that one will be treated differently due to one's social group membership (Steele et al., 2002). SIT can be triggered by various environmental cues and situations, such as the absence of a large number of people who share a social identity or group membership (e.g., perceiving oneself to be in a numerical minority in a certain setting); this example, of course, is relevant in higher education given the social group disparities in academia (e.g., the underrepresentation of certain racial/ethnic groups and of women in certain fields).

One way that SIT affects performance is through processes of vigilance and appraisal. First, individuals become vigilant to their surroundings, looking for cues that their identity is tied to how they will be evaluated, that is, threat cues. If these individuals perceive, or appraise, such

threat cues, their behavior and cognition may be negatively affected (Cohen et al., 2012; Cohen & Garcia, 2008; Inzlicht et al., 2011; Schmader et al., 2008; Spencer et al., 2016; Walton & Spencer, 2009). An illustrative example of SIT that has received ample attention from researchers over the years is stereotype threat. Stereotype threat is a specific form of SIT that can cause underperformance when highly motivated people are concerned that their performance in a context will confirm negative stereotypes about an important group to which they belong (Steele, 1997). Put another way, stereotype threat is underperformance rooted in the fear of confirming negative group stereotypes (Steele et al., 2002). For instance, Steele and Aronson (1995) found that Black undergraduates asked to indicate their race before taking a test (i.e., a threat cue) perform worse on the test than those who are not asked to indicate race beforehand because of the prevalent societal stereotype that Black people are unintelligent. Furthermore, stereotype threat is heightened in contexts where people are highly identified with a domain (e.g., women studying mathematics), and students who most want to succeed tend to be most affected (Aronson et al., 1999). Given the motivation required to be accepted into a doctoral program, most Ph.D. students are likely highly identified, high achieving, and eager for success in their program and field and may therefore be particularly at risk of stereotype threat. However, Ph.D. students may face SIT in broader contexts as well.

The processes of vigilance and appraisal that SIT entails can become recursive over time, feeding an aversive feedback loop that certain social psychological interventions are intended to interrupt (see Walton, 2014). Given these recursive processes, SIT affects not only immediate performance (as is the case with stereotype threat) but can influence important downstream outcomes. For example, a Hispanic student who has perceived over time multiple cues that they are less intelligent than their White and Asian peers may become hypervigilant to potential threat

cues and start to perceive constructive feedback from teachers as an indicator that the student is incompetent. This may lead to disengagement with the class that results in a worse grade, which then feeds into the notion of incompetence, and ultimately, particularly if the cycle repeats across classes, can lead to quitting a field of study or even school altogether. As this example illustrates, SIT can have significant consequences for people's life experiences (e.g., Major & Schmader, 2018; Steele et al., 2002).

Overall, given the general potential for psychological threat in doctoral education as well as for SIT to affect certain groups of students in particular, it would be helpful to be able to identify subgroups of students who differ in their patterns of psychological threat experiences. This identification is especially important because differences in psychological threat experiences may underlie differences in student outcomes like attrition. If we can pinpoint who is most likely to drop out of their program from the beginning, we can more precisely target interventions in the future. To identify students at potentially heightened risk of attrition, we can use a technique called latent class analysis.

Latent Class Analysis

Past researchers have tried to pinpoint what constructs explain achievement gaps and attrition rates at pre-doctoral levels of education. Many variables related to psychological threat that influence academic achievement have been identified, such as growth mindset (e.g., Blackwell et al., 2007; Yeager, Walton, et al., 2016), grit (e.g., Bowman et al., 2015; Duckworth et al., 2007; Strayhorn, 2014), academic identification (e.g., Finn, 1989; Fredricks et al., 2004), stereotype threat (e.g., Aronson et al., 1998; Steele et al., 2002), self-efficacy (e.g., Choi, 2005; Lent et al., 1984), and sense of belonging (e.g., Anderman, 2003; Cohen & Garcia, 2008; Good et al., 2012; Nieuwenhuis et al., 2019; Rosenthal et al., 2013). One approach to explore what

kinds of students may be more at risk of attrition would be to add every one of these potentially influential variables as predictors in a standard regression equation. However, this kind of variable-centered approach does not account for nuanced patterns among individuals and would become increasingly unwieldy as the list of potentially informative variables grew. An alternative approach is to use latent class analysis (LCA), which is a person-centered approach that takes into account the patterns of responses evident across individuals and also serves as a data reduction tool by transforming a list of variables into a smaller number of subgroups that are differentiated on some latent construct. Factor analysis could similarly be useful for data reduction; however, as it is a variable-centered approach, it would not take into account any patterns among individuals in the sample and would thereby not answer our question about what kinds of students are at-risk. LCA, on the other hand, can simplify how we approach the question of what at-risk students look like by allowing us to discover whether there are subtypes of people who vary in their risk of attrition from doctoral education, potentially for different reasons.

Like other mixture models, LCA unearths hidden subgroups from observed data, which enables one to assign each person to a subgroup (see Oberski, 2016). Specifically, LCA is a mixture model that is used when one suspects that an unobserved categorical variable separates a population into mutually exclusive and exhaustive subgroups, or latent classes (Goodman, 1974; Lazarsfeld & Henry, 1968). This technique is useful for reducing multiple variables into a smaller number of more interpretable groups as well as for examining whether a model works differently for different people (Oberski, 2016); in this case, we can use LCA to ask if intervention effects are different for different types of students. As a first step in LCA, however, we have to identify the subgroups in our sample, which is based on how multiple observed individual traits intersect. For example, one subgroup of students might tend to report low levels

of sense of belonging, grit, and academic identification, a constellation of traits that could portend suboptimal outcomes; a different subgroup might tend to report low belonging and academic identification, but high grit, a constellation of traits that could portend potentially mixed outcomes. We can choose these observed traits, also known as indicators, from psychological characteristics and experiences that have been previously identified as important to academic outcomes, such as those aforementioned (e.g., sense of belonging, grit, and academic identification). By using LCA, we can take into account multiple indicators of threat simultaneously to explore patterns of threat within a population and ultimately ascertain whether there are certain subtypes of doctoral students who are at higher risk of attrition (based on their psychological threat). Moreover, we can explore how the subgroups identified through LCA fare over time by exploring distal outcomes within the LCA framework. For instance, we can examine if subgroups characterized by ostensibly more threatening patterns of traits have worse psychological outcomes at the end of the semester than lower threat subgroups. Similarly, we can leverage this technique to explore potential boundary conditions of commonly used social-psychological interventions—namely, for whom do these interventions work?

Interventions

Students vary in their appraisals of different events, like receiving critical feedback, being rejected by a journal, or failing to win a fellowship. For instance, some view such negative events as reflective of their self-worth or as indicative that they do not belong in graduate school. Self-affirmation and belonging interventions are meant to uncouple negative external events from such negative appraisals. Social-psychological interventions such as these can reduce the threat-provoking potential of academic environments and thereby promote persistence among students who might otherwise disengage (e.g., Cook et al., 2012; Shnabel et al., 2013). These

brief interventions exert influence by shifting the way vulnerable students frame their experiences, particularly during sensitive transitions (e.g., beginning graduate school; Walton & Wilson, 2018). While these kinds of interventions have been tested at pre-doctoral levels of education to improve student outcomes and bridge achievement gaps between social groups (e.g., based on race/ethnicity; see Walton & Wilson, 2018), they have not yet been examined among doctoral students. Moreover, given this gap, it is unclear what kinds of doctoral students may benefit most from intervention. One possibility is that students who do not face substantial psychological threat do not benefit from intervention. It could also be the case that only students facing SIT benefit. We can explore these and other possibilities in the current research as we examine how different kinds of Ph.D. students react to intervention.

Students who do not feel like they belong tend to disengage and withdraw (e.g., Woodcock et al., 2012). Belonging interventions enable students to appraise negative events in a more adaptive light than they might otherwise (Walton & Cohen, 2011; Wilson & Linville, 1985). When students transition to a new school (e.g., college) they can expect to face many challenges. Belonging interventions can capitalize on this for student gain by relaying to students that challenges and feelings of uncertainty or being out of place are normal. Through this normalization, students' sense of belonging is buffered and negative external events that might otherwise cause their feelings of belonging to decrease are tempered. For instance, a new doctoral student might feel uncertain about how to spend their time and how to juggle multiple new responsibilities, and therefore worry that they may not be competent enough for graduate school. A well-timed intervention that normalizes these challenges and uncertainty may be enough to reassure this student and safeguard them from appraising future negative events as confirmation that they do not belong. Such interventions also serve to provide an explanation for

lack of felt fit in an academic context other than internal attributes associated with stereotypes about one's group (Walton & Brady, 2020). Students who are underrepresented or face negative stereotypes about their group membership—and therefore may be more likely to question their belonging as a broader concern that is related to stereotype threat—are meant to have their sense of belonging insulated by such interventions and have been shown to be most positively impacted by belonging interventions (Walton & Brady, 2020; Walton & Cohen, 2007).

Belonging interventions have been tested in various contexts with different levels of success. From these efforts, we can begin to understand the boundary conditions of the intervention; however, many questions remain. Borman et al. (2019) implemented a social belonging intervention for sixth grade students in several public middle schools across a Midwestern U.S. district. Benefits for these students, over the course of the academic year, included fewer discipline referrals and absences, better grades, and better attitudes about school. The benefits were not different by demographic characteristics (i.e., gender and race/ethnicity). In contrast, Yeager et al. (2016) found that, in an undergraduate context, a social belonging intervention specifically helped disadvantaged students, namely racial minority and first-generation students (i.e., students whose parents did not attend college). In addition to mixed evidence regarding for whom the belonging intervention works and in what contexts, it is unclear why the intervention sometimes fails to work at all. Although at times belonging interventions have been helpful to promote student outcomes—including grades, sense of belonging, and psychological well-being (e.g., Walton et al., 2015; Walton & Cohen, 2007, 2011)—at other times there is no discernible effect of the intervention (e.g., Broda et al., 2018). Past research has pointed to several important contextual moderators, including how the intervention may only work for students facing SIT (e.g., Yeager et al., 2016) or in certain school contexts (Hanselman

et al., 2014), but there is not yet an accepted, unified theory or a single explanatory construct that explains the variations in intervention effects. Testing the intervention in different contexts and examining who it works for is one way to contribute towards a greater understanding of both the intervention itself and scholastic achievement and psychological threat at a more basic level.

Another social-psychological intervention that can change how students react to potential threats, values-affirmation, is based on self-affirmation theory. According to this theory, people want to perceive themselves as having self-integrity, which includes being capable and moral (Steele, 1988). When people face threats that undermine these perceptions, they may respond adaptively or maladaptively. For instance, one student may respond to a low grade on a test by disengaging from the class (maladaptive) whereas another student may respond to the same grade by increasing efforts (adaptive). Values-affirmation interventions allow people facing threats to their self-integrity to respond by focusing on values unrelated to the threat and thereby reassert their worth as individuals (Cohen & Sherman, 2014). For instance, a new doctoral student may feel threatened by critical feedback on a paper because they are used to feeling—and likely value being—highly competent in an academic context. A values-affirmation intervention would remind this student about other values they hold, like being a good sibling or a devout person, and thereby bolster their self-integrity. Thus, this student could then respond to critical feedback and other potential threats adaptively (e.g., revising their work to improve it) rather than maladaptively (e.g., avoiding revisions due to anxiety or defensiveness). People who face SIT should particularly benefit from values-affirmation interventions, as they face specific kinds of challenges to their self-worth based on their social identity, be it race, gender, class, or something else.

Values-affirmation interventions have also been tested in various contexts to varying levels of success, suggesting potential boundary conditions. Cohen et al. (2006) showed the efficacy of this intervention in reducing the racial achievement gap between Black and White middle-school students, as the intervention positively affected Black students and did not affect White students. Later studies found similar benefits for Latino (compared to White) middle-school and undergraduate students (Brady et al., 2016; Goyer et al., 2017; Sherman et al., 2013). These interventions have also been effective in other contexts, like with first-generation college students (Harackiewicz et al., 2014), women in graduate business schools (Kinias & Sim, 2016), and students in massive open online courses (Kizilcec et al., 2017). However, other studies have not demonstrated an effect of a values affirmation intervention (e.g., Bratter et al., 2016; Hanselman et al., 2017; Hayes et al., 2019; Jones & Huey, 2020; Protzko & Aronson, 2016) and several have even revealed backfire effects for certain groups of students (Bayly & Bumpus, 2019; Brady et al., 2016; Dee, 2015; Liu & Huang, 2019; Miyake et al., 2010). As with the belonging intervention, it is unclear why values affirmation interventions do not always work.

However, one theory that can help us think about mixed results for various kinds of social-psychological interventions is that interventions fail to work when the context they are delivered in does not provide the necessary psychological affordances; in other words, the context does not permit students to adopt the shift in perspective that the intervention suggests (Walton & Yeager, 2020). For instance, we can imagine a university that provides very few opportunities for Black students to belong; that context does not support the idea suggested in a belonging intervention that concerns about fitting in will diminish over time and may thus render the intervention ineffective for Black students. We take the perspective that not only does the psychological context matter in terms of the environment permitting individuals to adopt

perspective shifts, but also the way we examine intervention efficacy differences among population subgroups matters. Thus, in the present research, we examine if intervention efficacy depends on the type of student in a different way than past research. Rather than using social group membership as a proxy for psychological threat and estimated risk level as many prior studies have done (to mixed results), we shift the focus back to the underlying psychology of students, that is, their psychological threat experiences.

Overview of Current Research

The present study uses data from the *Study for the Advancement of Graduate Education and Scholarship*, or SAGES, a randomized controlled longitudinal field experiment that examines the psychological experiences of doctoral students that predict attrition and retention, particularly in STEM fields. SAGES participants, drawn from Stanford, Columbia, and Pennsylvania State Universities, were randomly assigned to one of three conditions: social belonging intervention, values-affirmation intervention, and control. A primary goal of SAGES is to understand the paths that lead students to leave before earning their degree; an additional goal is to examine to what extent interventions can mitigate psychological threat and improve retention outcomes for doctoral students.

In the present research, we aim to increase our understanding of the psychological experiences of doctoral students and conduct a preliminary exploration of the effects of two social psychological interventions in this under-examined but important population; we do this with an eye towards understanding who is at risk of attrition and who may be best targeted by interventions. We describe the present study in four phases. Phase 1 is the conduction of the initial unrestricted LCA, or the delineation of threat profiles (AKA classes) using data from a baseline survey that incoming Ph.D. students completed. In Phase 1, we leveraged LCA to

identify different subgroups of doctoral students based on varying experiences of psychological threat, including but not limited to self-report of social identity threat (SIT); we suggest these threat profiles reflect differences in potential risk of attrition. In Phase 2, we examined how the selected model maps onto student demographics and baseline characteristics to get a better sense of the students in the various classes. We examined expected demographic differences in threat profiles based on past literature. For instance, we expected that female Ph.D. students in many STEM fields would experience greater identity threat than men due to prevalent intellectual stereotypes about gender and women's historic underrepresentation in STEM fields. Thus, we expected women to be over-represented in high-threat profiles. In Phase 3, we examined whether threat profiles predict distal outcomes. We tested prospectively for profile differences at the end of the first semester in key psychological experiences identified from previous research. We did this to help determine the predictive value of the identified threat profiles in identifying students at more or less risk of attrition. In Phase 3, we examined how threat profiles predict distal outcomes for students in the control condition specifically—we focus on the control condition students alone in order to home in on the experiences of typical Ph.D. students who do not receive intervention. Finally, in Phase 4, we examined how social belonging and values affirmation interventions might impact Ph.D. students with different threat profiles differently with goals of (1) identifying whether these interventions are effective in this population, and (2) identifying whether the efficacy of the interventions depends on a student's threat profile. Given past research with these interventions at lower levels of education, we expected interventions to be more helpful for students characterized by relatively higher threat.

CHAPTER 2. METHOD

Participants

Participants were two cohorts of first-semester doctoral students at Stanford, Columbia, and Pennsylvania State universities (PSU) who participated in a baseline survey in the fall of 2018 or fall of 2019 for the *Study for the Advancement of Graduate Education and Scholarship* (SAGES; pre-registered on the Open Science Framework at osf.io/f7s6r). All first-year STEM Ph.D. students at these universities were eligible to participate in SAGES in its entirety, and a smaller number of non-STEM students were also specifically targeted (at PSU). We categorized fields as STEM or non-STEM based primarily on the National Science Foundation (NSF) classifications of STEM fields; we also made use of available information (e.g., from university websites) to match ambiguous fields to the fields represented in the NSF classifications. Students from non-STEM fields were also welcomed to participate if they became aware of the study from recruitment at general orientation events for Ph.D. students (at PSU and Stanford). The goal of the project was to recruit at least 540 students in total to complete the longitudinal aspect of the study (i.e., to continue participating after the baseline survey). We discuss further details about recruitment of participants in the following section. Ultimately, 1115 students completed at least part of the baseline survey, with 1085 finishing it and potentially being invited to the longitudinal study. Of those who completed the baseline survey, 656 continued onto the longitudinal part of SAGES.

Phase 1 and 2 analyses were based on the 1085 students who completed the baseline survey at the beginning of the semester. Participants who did not finish the baseline survey, and were therefore not invited to the longitudinal study, were not included in analyses.¹ Four students

¹ Although not presented in this paper, we did also conduct the unrestricted 4-class LCA for the ultimately selected model using all available data (i.e., including students who did not finish the baseline survey, but who may have

did not have data for any of the indicators, reducing the sample to 1081. Three students with missing responses on one or more indicators were retained because LCA uses maximum likelihood estimation, which allows partial data on indicators. For Phases 3 and 4, which focused on distal outcomes, the sample size was lower because there were fewer students participating in the longitudinal aspect of the study. Moreover, each Phase 3 and 4 analysis varies slightly in sample size because not all students in the longitudinal study completed each survey wave or each question on a given survey. The demographic characteristics of the final sample for the current study ($n = 1081$), are in Table 1. Students in the final sample were nearly evenly female and male (522 and 546, respectively, with 20 specifying another gender), and 24 years-old on average ($M = 24.17$, range 19 to 55).

Procedures

We recruited students via email across all campuses as well as at in-person orientation sessions for new doctoral students at Stanford and PSU. Recruitment occurred in the summer before students had started their programs whenever possible and otherwise in the first week or two of the semester.² We told participants that the study was focused on “understanding the experiences of Ph.D. students like you throughout graduate school.” Recruitment materials included a link to a baseline survey intended to unearth how students thought and behaved before beginning graduate school; the research team paid participants \$15 to complete this initial 45-minute web-based survey and invited them to participate in the longitudinal aspect of SAGES. The research team informed students that they could complete only the baseline survey if they did not want to participate in the longitudinal study. We excluded students from the longitudinal

responded to some or all of the measures in the LCA). Adding these 44 participants (total $N = 1125$) did not change the latent classes that emerged, and the percentage of people in each class was similar (see Chapter 3. Results)..

² Stanford uses a quarter system rather than a semester system but herein we use the term semester for all campuses.

Table 1*Demographic Information*

	<i>N</i> (proportion)	Mean (<i>SD</i>)	Range
Age	1069	24.17 (3.36)	19-55
Socioeconomic status	1074	6.07 (1.80)	1-10
Gender			
Genderqueer	17		
Female	522		
Male	546		
Trans male	3		
Race/ethnicity			
Asian	124 (.115)		
Black	24 (.022)		
Hispanic	27 (.025)		
Multiracial	42 (.039)		
Asian/White	14 (.013)		
Other multiracial identity	28 (.026)		
Native American	2 (.002)		
White	410 (.379)		
American, race unknown	3 (.003)		
International	449 (.415)		
Sexual orientation			
Asexual	16 (.015)		
Bisexual	111 (.103)		
Gay/Lesbian	56 (.052)		
Straight	864 (.803)		
Other	29 (.027)		
First-generation status			
Yes	218 (.203)		
No	857 (.797)		
International student			
Yes	449 (.415)		
No	632 (.585)		
Region of birth for international students			
Africa	12 (.027)		
Arab States	3 (.007)		
Asia & Pacific	322 (.719)		
Europe	49 (.109)		
Middle east	29 (.065)		
North America	8 (.018)		
South/Latin America	25 (.056)		

Note. The maximum *n* is 1081 (number of students used in final analytic sample for unrestricted LCA).

study if they did not have an Android or iPhone smartphone or were PSU medical students matriculating at a geographically distinct campus. We also excluded students from the study if they were under 18 years of age. We next describe procedures for each campus.

At PSU, the Graduate School agreed to send a recruitment email to incoming Ph.D. students. The research team wrote this email, with an introduction from the Graduate School Dean. The dean's message noted the Graduate School's support for the study but emphasized the independence of the Graduate School from the study and assured students that their participation would not be known by the University or affect their graduate career in any way. The recruitment email was sent to every new Ph.D. student in 7 STEM-focused colleges (the College of Agricultural Sciences, the College of Earth and Mineral Sciences, the College of Engineering, the College of Information Sciences and Technology, the Eberly College of Science, the Intercollege Graduate Programs, and the College of Medicine). For the non-STEM comparison group, the email was also sent to students in the College of Arts and Architecture and the College of the Liberal Arts and to students in the Departments of Biobehavioral Health and Human Development and Family Studies (within the College of Health and Human Development). The recruitment email was sent to students a month before the Fall semester began, with two reminder emails in the following week. To maintain the confidentiality of students, the emails were all sent through the Graduate School. Thus, the research team was not aware of any details of individual non-participating students, and the Graduate School was not aware of whether students had opted to participate. The Graduate School sent three emails at the request of the research team and provided the research team limited demographic data. If students outside of the targeted fields became aware of the study from recruitment at the in-person orientation

sessions, where the research team handed out fliers with links to the survey and reminded students about the study, they were allowed to participate.

At Columbia, with assistance from their respective deans, recruitment emails were sent to all new doctoral students in the Graduate School of Arts and Sciences and the School of Engineering and Applied Science. Emails were not substantively different from those used at PSU, but edited minimally to be appropriate for each campus. Initial emails were sent during the second week of the Fall semester with two reminders sent the following week. At Stanford, recruitment emails were sent to all new engineering Ph.D. students, the first email being sent during the second week of the Fall quarter and two reminders sent the week following. Because at Stanford we did not have direct access to incoming STEM Ph.D. students outside of engineering, the research team sent the recruitment email to directors of graduate studies and department chairs in STEM departments in the School of Humanities and Sciences with a request to forward the message to incoming doctoral students. The SAGES team also recruited students at Stanford at an in-person orientation event for new graduate students where the research team handed out fliers and reminded students about the study.

Ultimately, 1085 students fully completed the baseline survey across the campuses and were invited via email to continue with the longitudinal smartphone study by attending an orientation and training session. We excluded, and thus did not invite to the orientation session, students from the longitudinal study if they did not have an Android or iPhone smartphone, which was required for the survey app, MetricWire. We also excluded students from PSU's College of Medicine because they attended a campus in a different location than the other students, which would have made it difficult to attend the in-person, required orientation and training sessions. In the email invitation to the orientation session, students were told briefly

about the study and compensation, which follows a general rate of \$25/hour based on number of surveys completed, with bonuses to incentivize higher completion rates. In this invitation email, students were also given a link to sign up for an orientation session. A day before their session, participants received a reminder email detailing how to install the MetricWire app required for smartphone surveys on their phones.

Orientation sessions were held over the course of one early week in the semester and lasted approximately one hour in sufficiently small groups that all attendees could be seated comfortably with space between them in a single room. At the orientation sessions, the goals were to provide people information about the study, encourage their participation and motivate their commitment, and conduct the intervention. The same member of the research team led all orientation sessions using a scripted presentation to describe study goals, procedures, and compensation rates, as well as administer the intervention or control exercises. He handed out folders labeled with each participant's SAGES ID and email address that contained, in order, a cover page with the study title and logo (first cohort only due to a logistical issue), an informational handout about the study, two copies of the consent form (one for the research team, one for the participant), one transcript release form, an envelope labeled with the participant's SAGES ID that contained the participant's assigned intervention activity, and a packet labeled "Activity 2" with post-intervention questionnaires. Participants also received a note on their folders instructing them to not open the folders until told to do so. The researcher began with a brief overview of the study and what would be involved and then asked participants to review and sign consent and transcript release forms, answering any questions that might arise. Next, the researcher introduced the 20-25-minute intervention/control exercise, which he described as "a brief activity that involves reading some passages and doing some informal

writing to get a better sense of students' values and perspectives at the beginning of graduate school." He told students that the exercise was part of the university's "interest in supporting doctoral students and promoting student success." To avoid suspicion, the researcher also told students that "the activities may differ slightly from person to person, because...we're spreading out different activities across different people." If any students were still completing the activity after 20 minutes, the researcher gave them a notification that 5 minutes were remaining and then again when 1 minute was remaining, so that all students finished the activity within 25 minutes. Students either completed the belonging, values-affirmation, or control condition exercise after random assignment to condition via the R package blockTools (Moore & Schnakenberg, 2012). Assignment to condition, done beforehand, was blocked by gender, race, first-generation college student status, and international student status in order to balance these covariates across conditions and improve the precision of estimating treatment effects (Gerber & Green, 2017).

Intervention procedures followed established protocols and were designed to look similar across conditions and to take the same amount of time. Briefly, affirmation condition participants chose from a list of 11 values the 2 or 3 values that were most important to them, and then were prompted to think about their chosen values and write about why they valued them (Miyake et al., 2010). Additional prompts had participants write the top two reasons they held their chosen values and rate how much they lived up to and cared about these values, as well as the extent to which the selected values were important aspects of themselves.

The materials for the social belonging exercise, adapted from previous protocols (e.g., Walton et al., 2017), began with a summary of findings from a survey that found that many doctoral students were initially concerned about fitting in in graduate school overcame their worries over time. The summary was followed by quotes from Ph.D. students that followed the

same theme. After reading the summary and quotes, participants were asked to write about why Ph.D. students might initially worry about belonging but over time feel at home and then to write a brief letter to an incoming graduate student about how people's graduate school experiences change over time—specifically, how they go from uncertainty about belonging to feeling at home in graduate school.

The control condition exercise was similar to the social belonging exercise but based on study skills; this exercise was adapted from Walton et al. (2015). The control condition exercise was designed to be similarly cognitively tasking and equivalent in length to the intervention exercises, as well as emotionally non-reactive. Control condition participants read about survey findings that doctoral students initially were concerned about how to study in graduate school but learned over time. After, participants read quotes from Ph.D. students that followed these themes and then wrote about why Ph.D. students might initially worry about study techniques but over time develop good study habits. Lastly, participants wrote a brief letter to an incoming graduate student about how people's graduate school experiences change over time—specifically, how they go from uncertainty about the challenging workload and deadline tracking to developing good study skills.

After completing the intervention or control exercise, students responded to a short set of measures to examine if there were any immediate intervention effects on psychological outcomes—these are not a focus of the present research. Then, the presenting researcher described how to use the smartphone application, invited questions, and concluded the orientation. Participants were provided \$25 for their time.

Over the next ten months of graduate school during eight non-consecutive weeks, participating students completed the smartphone study for SAGES. At PSU and Columbia, three

of these survey weeks, or assessment bursts, occurred during the Fall semester, one during winter break, three during the Winter/Spring semester, and one at the start of summer break. At Stanford, which uses a quarter-based academic calendar, two assessment bursts occurred in the Fall quarter, one during winter break, four during the Winter and Spring quarters, and one at the start of summer break. At the end of each assessment burst, participants completed a web-based end-of-week (EOW) survey; for all but one assessment burst, this EOW survey followed a week of smartphone surveys, but herein we focus on the EOW surveys rather than the smartphone surveys. The EOW surveys, which took 15-20 minutes, examined recent academic activity as well as psychological and behavioral changes from baseline. Participants were asked to take the EOW surveys on a laptop or desktop computer (rather than their smartphone) so that all questions would display properly; the survey was administered through Qualtrics. Participants were compensated four times throughout the year at a fixed rate for each completed survey; bonus payments were given twice a year for high rates of survey completion. In addition, participants in the smartphone study received 3 booster interventions throughout their first year; these requested that participants expand on whichever themes were present in their initial intervention condition via their responses to a condition-specific prompt that changed slightly between boosters. These booster interventions are not a focus for the current study.

The current study only uses data from the baseline survey, the EOW3 survey (administered in the latter half of the semester), and the EOW4 survey (administered at the end of the first semester). We were primarily interested in outcomes at the end of the first semester; however, given logistical issues regarding length of surveys, one outcome of interest was not measured at the EOW4 and thus we examined it at the preceding timepoint.

Measures

Demographic characteristics.

Participants reported their gender, sexual orientation, race/ethnicity, socioeconomic status (SES), first-generation status, and international student status in the baseline survey (see Table 1). Participants reported their current gender identity at the beginning of the survey because the wording of a few measures in the survey changed slightly based on selected gender identity. Participants reported other demographic information at the end of the survey. Participants reported gender identity by selecting one or more of the following categories: female, male, trans female, trans male, genderqueer or non-conforming, or they could specify a different identity in an open-text response. For sexual orientation, participants could select among heterosexual or straight, gay or lesbian, asexual, bisexual, or they could specify other in an open-text response. For race/ethnicity, participants could both self-report in an open-text response and select any of several categories that applied (see Table 1). For first-generation status, participants reported whether they considered themselves a first-generation college (undergraduate) student. For SES, using one of the MacArthur Scales of Subjective Social Status, participants placed themselves on a ladder with rungs labeled bottom-up from 1 to 10 in relation to others in the United States (“Where would you place yourself on this ladder? Please select the rung where you think you stand at this time in your life, relative to other people in the United States;” Adler et al., 2000). We classified students as non-international if they reported being either U.S. citizens or U.S. permanent residents; otherwise, we classified students as international.

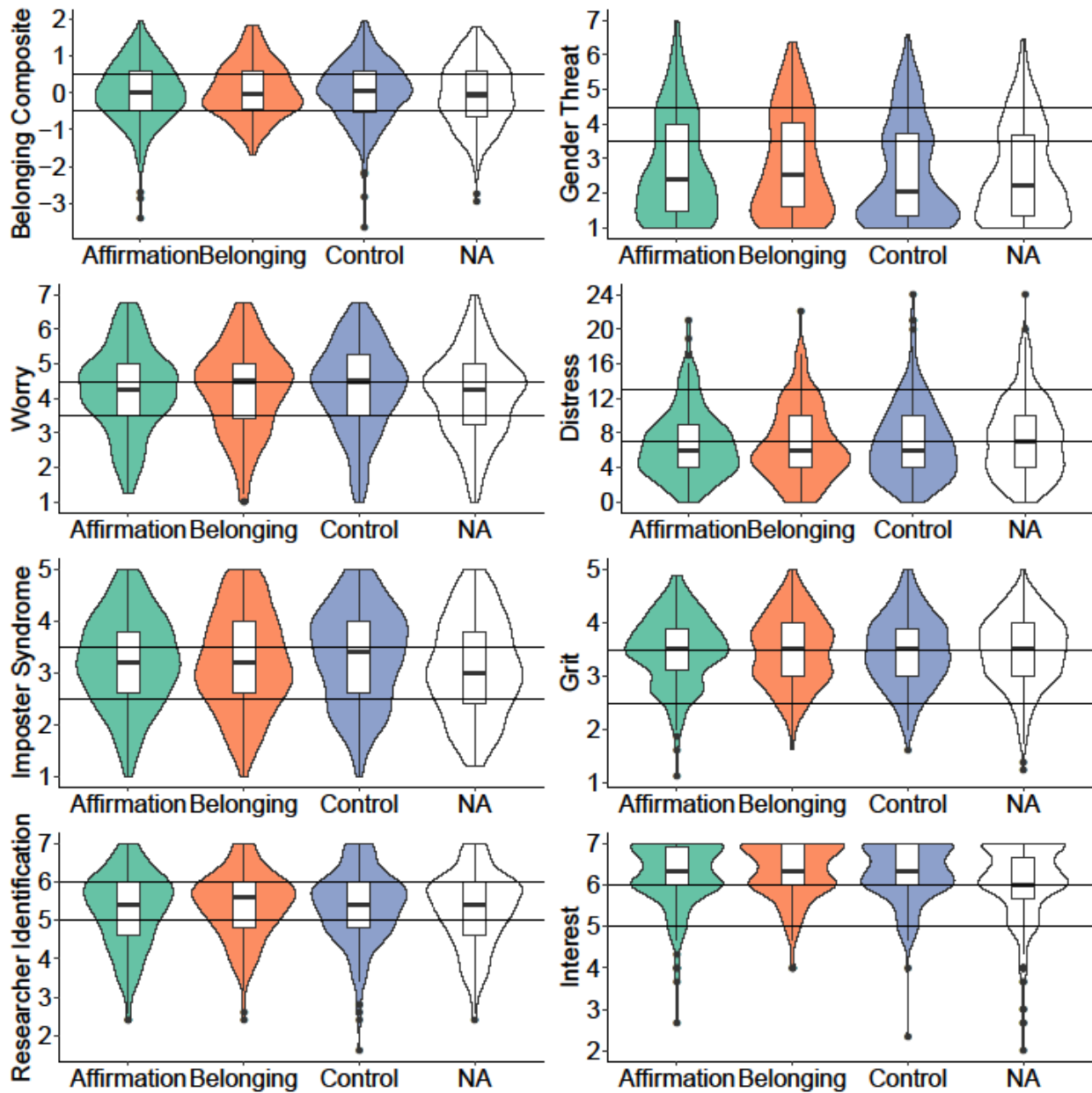
Phase 1 psychological threat (indicators).

The Phase 1 psychological threat measures, measured at baseline, are the indicators for the LCA, or the manifest variables that we entered into the LCA to detect a latent construct. In

general, we chose the indicators based on previous literature suggesting their relevance to academic achievement, which we briefly note alongside the description of each measure below. The distributions of the indicators ultimately used for Phase 1's selected model are displayed in Figure 1. We derived the indicators from continuous variables that we ultimately trichotomized because LCA requires categorical variables. We did first attempt to run a similar analysis that uses continuous indicators, latent profile analysis, but the models were unstable, possibly due to violations of strict normality assumptions (Oberski, 2016).

Academic Belonging. We used items originally from the Social and Academic Fit Scale by Walton and Cohen (2007), abbreviated by Cook et al. (2012). This abbreviated scale consists of two subscales, social belonging (5 items, e.g., "People in my program accept me") and potential to succeed (4 items, e.g., "I know what I need to do to succeed in grad school"). In the current study, students responded using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*), which, to be consistent with other scales in the current study, is a slight modification from the original abbreviated scale using a 6-point scale. Wording was also slightly modified to reflect a graduate student context (e.g., using phrases like "in grad school"). The 9 items together have high reliability in the current sample, $\alpha = .81$, and given the original conception of the scale as a single construct and the need in the current study to narrow down indicator variables for the latent class analysis, we averaged all 9 items to create one scale score ($M = 5.14$, $SD = 0.79$). Sense of belonging, a target for the social belonging intervention, is associated with various academic performance outcomes (e.g., Anderman, 2003; Cohen & Garcia, 2008; Good et al., 2012; Nieuwenhuis et al., 2019; Rosenthal et al., 2013).

Graduate School Self-Efficacy. Three items adapted from Shryock and Froyd's (2017) 8-item engineering self-efficacy scale assessed graduate school self-efficacy (e.g., "I expect to do

Figure 1*Distribution of Indicator Items by Condition*

Note. Horizontal lines indicate cut-offs used to categorize these continuous variables into trichotomous variables.

Affirmation, belonging, and control refer to intervention conditions—NA refers to students who completed only the baseline survey and did not participate in the longitudinal study (and therefore were not assigned a condition).

well in graduate school”). Student responses using a 4-point scale from 1 (*not at all true of me*) to 4 (*very true of me*) were averaged to create one scale score ($M = 3.08$, $SD = 0.62$, $\alpha = .79$). Self-efficacy predicts college performance, including grades, persistence, and perceived career options (e.g., Choi, 2005; Lent et al., 1984).

Psychological Need Satisfaction. This scale reflects three subscales based on self-determination theory: the need for competence (e.g., “I successfully complete difficult tasks and projects”), connection (e.g., “I feel close and connected with other people who are important to me”), and autonomy (e.g., “I am free to do things my own way”). Students responded using a 7-point scale ranging from 1 (*not at all true*) to 7 (*extremely true*). The 18 items, 6 for each subscale, were taken from Sheldon and Hilpert (2012), although we changed the items and instructions from past tense to present tense because we wanted to measure students’ current psychological experiences. Because we removed two items from each subscale for the second cohort of participants (because of constraints due to survey length), we averaged only the remaining 12 items that participants from both cohorts completed for a scale score ($M = 4.94$, $SD = 0.77$, $\alpha = .78$). Greater psychological need satisfaction predicts greater academic motivations, effort, and persistence (e.g., Gnambs & Hanfstingl, 2016; La Guardia & Ryan, 2002; Liu & Flick, 2019).

Belonging Composite. The scale scores for academic belonging, graduate school self-efficacy, and psychological need satisfaction loaded onto the same factor in a principal components analysis, which we ran for several combinations of the indicator variables we considered because we wanted to reduce the number of potential indicator variables. The resulting composite variable, arrived at after standardizing each of the three measures and then averaging, showed good reliability ($M = 0.00$, $SD = 0.83$, $\alpha = .76$). Given our desire to reduce the

number of indicator variables for a latent class analysis because of fit difficulties with models that use too many indicators, we used the composite variable. As noted previously, latent class analysis requires categorical indicator variables; thus, we trichotomized this continuous composite variable. Because this variable was based on a composite, we could not choose cutoffs based directly on the scale response options, as we did with other variables; instead, we chose cutoffs based on the distribution of responses (i.e., based on the standard deviations and the percentage of people in each cell). We categorized respondents as “low” if their score on the composite was below -0.5 ($n = 283$), “medium” if between -0.5 and 0.5 ($n = 489$), and “high” if above 0.5 ($n = 308$). Because we transformed the data, we could not choose cutoffs based directly on scale meaning; instead, we chose cutoffs based on the distribution of responses (i.e., based on the standard deviations and the percentage of people in each cell).

Stereotype Threat. Adapted from Cohen and Garcia (2005), the stereotype scale has six items that assess stereotype threat. We included measures of stereotype threat based on race and gender, both of which we modified to reflect a graduate student context (e.g., “I worry that people in my graduate program will judge me based on what they think of my racial group [people of my gender]”). Students responded using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Separate average scores were created for gender ($M = 2.69$, $SD = 1.36$, $\alpha = .91$) and race-based stereotype threat ($M = 2.65$, $SD = 1.43$, $\alpha = .91$). Herein we focus on gender-based stereotype threat because race-based stereotype threat was not ultimately used as an indicator (due to its failure to distinguish well between classes). Higher levels of stereotype threat predict worse academic performance, and in particular the underperformance of women and URM students (e.g., Aronson et al., 1998; Steele et al., 2002).

Identity Interference. Four items taken from Settles (2004), who originally used 17 items, assessed identity interference, which happens when the pressures of one identity conflicts with the performance of another (e.g., “I feel that other [scientists/researchers] do not take me seriously because of my gender”). Students responded using a 7-point scale from 1 (*strongly disagree*) to 7 (*strongly agree*). We modified the phrasing of the items to be about gender in general rather than women specifically so that men and non-binary participants could complete the scale as well, and we changed the word “scientist” to “researcher” in the second cohort given not all participants were in a science-focused program of study. Higher levels of identity interference predict not only negative psychological outcomes like worse self-esteem, but also lower perceived academic performance in science (Settles, 2004; Settles et al., 2009).

Gender Threat. Principal components analysis revealed that the gender-based stereotype threat subscale score loads onto a gender threat factor along with two survey items from the identity interference scale, “I feel that other [scientists/researchers] do not take me seriously because of my gender” and “It is hard to be my gender and a [scientist/researcher] in my field at the same time.” The resulting composite variable, hereafter called gender threat, had good reliability ($M = 2.68$, $SD = 1.42$, $\alpha = .85$). We then trichotomized this variable for use in LCA based on scale meaning such that responses 3.49 or below (*somewhat to strongly disagree*) were considered low ($n = 758$), between 3.49 and 4.49 medium (*neither agree nor disagree*, $n = 171$), and above 4.49 high (*somewhat to strongly agree*, $n = 151$).

Worry. These items also come from Cohen and Garcia’s (2005) stereotype threat scale; however, these 4 items do not contain reference to race, gender, or any other social identity, and as such can be thought of as reflecting general worry about being seen negatively rather than a form of social identity threat (e.g., “I worry that people in my graduate program will think I’m

dumb if I do badly”). These items were administered with the stereotype threat scale and participants answered them on the same 7-point scale. Again, we rephrased these items for a graduate student context. We averaged the items to create a scale score of general worry ($M = 4.20$, $SD = 1.24$, $\alpha = .77$). We then trichotomized this variable based on scale meaning such that responses 3.49 or below (*somewhat to strongly disagree*) were considered low ($n = 270$), between 3.49 and 4.49 medium (*neither agree nor disagree*, $n = 295$), and above 4.49 high (*somewhat to strongly agree*, $n = 515$). Tendencies to worry, measured in various ways, predict academic performance, although the relationship is notable in that especially high levels can worsen performance, but moderate levels can enhance it (e.g., Davey et al., 1992; Seipp, 1991).

Imposter Syndrome. Our imposter syndrome scale took 5 of the 20 items from the Clance Impostor Phenomenon Scale (Clance, 1985) and then added two items and slightly modified response option anchors for a resulting 5-point scale ranging from 1 (*not at all like me*) to 5 (*very much like me*). Because we excluded one item in the second cohort’s survey due to survey length concerns, we ultimately used only the 6 other items (e.g., “Sometimes I feel like a fraud”) to create an average scale score ($M = 3.20$, $SD = 0.93$, $\alpha = .79$). We then trichotomized this variable based on scale meaning such that responses 2.49 or below (*not at all like me to not much like me*) were considered low ($n = 267$), between 2.49 and 3.49 medium (*somewhat like me*, $n = 394$), and above 3.49 high (*mostly like me to very much like me*, $n = 419$). Impostor syndrome, although it is a characteristic of many successful people and does not necessarily preclude achievement, does predict worse psychological well-being, including higher burnout and anxiety (e.g., Bravata et al., 2019).

Grit. Eight items adapted from Duckworth and Quinn (2009) assessed grit (e.g., “I finish whatever I begin”) on a 5-point scale ranging from 1 (*not like me at all*) to 5 (*very much like*

me). We averaged the items to create a scale score ($M = 3.48$, $SD = 0.66$, $\alpha = .77$) and then trichotomized this variable based on scale meaning such that responses 2.49 or below (*not at all like me to not much like me*) were considered low ($n = 69$), between 2.49 and 3.49 medium (*somewhat like me*, $n = 433$), and above 3.49 high (*mostly like me to very much like me*, $n = 579$). Higher levels of grit predict several relevant outcomes, including higher grades, a greater sense of belonging, and more college satisfaction (e.g., Bowman et al., 2015; Duckworth et al., 2007; Strayhorn, 2014).

Psychological Distress. Our measure of psychological distress, adapted from Kessler et al. (2002), asked participants how frequently they felt nervous, hopeless, restless or fidgety, so depressed that nothing could cheer them up, that everything was an effort, and worthless during the past 30 days. Participants responded using a 5-point scale ranging from 0 (*none of the time*) to 4 (*all of the time*). We summed ($M = 7.04$, $SD = 4.26$, $\alpha = .85$) and then trichotomized items based in part on the cutoff for clinical concern, with students categorized as low in distress if they scored 7 or lower ($n = 627$), medium if they scored between 7 and 13 (the clinical cutoff, $n = 374$), and high if they scored above the clinical concern cutoff of 13 ($n = 78$). Higher levels of psychological distress predict lower academic achievement (e.g., Stallman, 2010).

Interest. We assessed interest in research (e.g., “I am interested in my research topic”) and interest in field (e.g., “I am interested in learning more about my field of study”) using three of four items from Choe et al. (2017). Participants responded using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*); the original scale was only 5 points. Items were also rephrased from engineering-specific to be more general. We averaged the three items to create one composite interest score ($M = 6.18$, $SD = 0.77$, $\alpha = .84$), which we then trichotomized. To avoid problems with small cell size given the skew of this variable (see Figure 1), we

trichotomized interest such that responses of 5 or below were considered low (*somewhat agree to strongly disagree*, $n = 116$), between 5 and 6 medium (*somewhat agree to agree*, $n = 392$), and above 6 high (*agree to strongly agree*, $n = 573$). Greater interest predicts retention in STEM fields (e.g., Patrick et al., 2018).

Researcher/Academic Identification. We assessed identification as a researcher and academic (e.g., “In general, being a researcher is an important part of my self-image”) using three items from Sellers et al. (1997) and two items from Choe et al. (2017). Participants responded using a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Items from Sellers et al.'s (1997) centrality scale were adapted to focus on research and graduate school. We averaged the five items to create one scale score ($M = 5.31$, $SD = 0.95$, $\alpha = .80$), which we then trichotomized. To avoid problems with small cell size given the skew of this variable (see Figure 1), we trichotomized it such that responses of 5 or below were considered low (*somewhat agree to strongly disagree*, $n = 393$), between 5 and 6 medium (*somewhat agree to agree*, $n = 506$), and above 6 high (*agree to strongly agree*, $n = 182$). Higher levels of academic identification are associated with greater academic persistence (e.g., Jones et al., 2013).

Outcomes.

We assessed several outcomes at the beginning and end of the first semester for Phases 2 and 3, respectively. We selected outcomes that may be related to attrition risk (e.g., SIT), as well as ones that can help us more fully understand what the students in different classes are like (e.g., outcomes that provide contextualizing information, like amount of academic preparation for graduate school). We averaged the items from scales with decent reliability (determined holistically, but with an emphasis on Cronbach's alpha and average inter-item correlation) to create scale scores, while we examined the items from scales with poor reliability individually.

We categorize outcomes into seven categories, described below: academic preparation and context, self-evaluations, attitudes towards graduate school and academia, interpersonal relations and perceived fit, mental health, social identity threat, and research effort. See Table 2 for descriptive statistics.

Academic Preparation and Context. For Phase 2, we examined contextualizing variables: the percentage of women in each student's field and students' previous research and educational experiences. We approximated the percentage of women in each student's field using data from the most recently available NSF Survey of Earned Doctorates (SED; National Science Foundation, 2018). Some of the fields reported by SAGES participants were not exact matches for the fields listed by the NSF. Thus, we examined the stimuli that were presented to participants of the SED (i.e., a detailed listing of specific fields of study categorized into larger areas of study) to extract information on how different fields were categorized. We used this information and other materials (e.g., a graduate program's website) to match the fields with SED data on percentage of women to the list of fields reported by SAGES participants. For instance, we used this approach to classify Acoustics, which was not a named field in the SED data, under Physics and Astronomy because that was how it was categorized in the SED material presented to participants. For fields where it was more ambiguous as to how they should be categorized, we searched the relevant university's website for information on how they described and/or classified their own programs of study.

Five author-generated items assessed participants' previous educational/research experiences, which may be helpful in promoting persistence in doctoral education. One dichotomous, yes/no item was simply a measure of whether or not participants had received a terminal master's degree prior to starting their Ph.D., which may be an indicator of greater

Table 2*Descriptive Statistics on Concurrent and Distal Outcomes*

Variable	Baseline			Distal		
	<i>n</i>	Mean (<i>SD</i>)	Range	<i>n</i>	Mean (<i>SD</i>)	Range
Academic Preparation & Context						
Undergrad research preparation	996	3.16 (1.06)	1 – 5			
Years undergrad research	1077	1.87 (1.09)	0 – 6			
Postgrad research preparation	545	3.74 (1.02)	1 – 5			
Years postgrad research	1074	1.08 (1.63)	0 – 20			
Proportion of women in field	1080	0.38 (0.17)	0.13 - 0.73			
Attitudes towards Grad School/Academia						
Interest composite	1081	6.18 (0.77)	2 – 7	580	6.01 (0.81)	1.67 – 7
Right decision	1081	3.46 (0.55)	1 – 4	579	3.39 (0.59)	1 – 4
Doubt decision	1081	2.19 (0.83)	1 – 4	580	2.27 (0.83)	1 – 4
Determination						
Item 1	1081	5.03 (1.56)	1 – 7	580	4.83 (1.59)	1 – 7
Item 2	1081	5.36 (1.44)	1 – 7	580	5.03 (1.53)	1 – 7
Item 3	1081	4.66 (1.66)	1 – 7	580	4.45 (1.62)	1 – 7
Item 4	1080	4.45 (1.47)	1 – 7	580	4.41 (1.33)	1 – 7
Preference for academic career	1078	1.46 (2.88)	-5 – 5	577	0.75 (3.17)	-5 – 5
Grad school evaluation				580	67.64 (19.22)	0 – 100
Grad school fun				580	61.97 (22.79)	0 – 100
Institutional Trust				557	5.85 (0.95)	1 – 7
Burnout				580	3.04 (1.37)	1 – 7
Interpersonal Relations and Perceived Fit						
Perceived social support	1079	4.03 (0.90)	1 – 5	580	4.10 (0.92)	1 – 5

Similarity to colleagues	1079	3.77 (1.33)	1 – 6	579	55.79 (24.56)	0 – 100
Isolation				580	2.74 (0.73)	1 – 5
Belonging	1081	5.14 (0.79)	2 – 7	580	5.20 (0.91)	2.11 – 7
Belonging Uncertainty						
Item 1	1080	4.49 (1.64)	1 – 7	580	4.88 (1.67)	1 – 7
Item 2	1079	5.67 (1.10)	1 – 7	580	5.74 (1.11)	1 – 7
Item 3	1078	3.63 (1.54)	1 – 7	580	3.28 (1.66)	1 – 7
Self-Evaluations						
Neuroticism	1081	4.09 (1.36)	1 – 7			
Self-efficacy	1080	3.08 (0.62)	1 – 4	580	3.17 (0.62)	1 – 4
Self-esteem	1080	3.17 (1.18)	1 – 5	580	3.30 (1.12)	1 – 5
Academic Self-Control						
Item 1	1081	3.29 (0.94)	1 - 5			
Item 2	1080	3.05 (1.10)	1 - 5			
Imposter Syndrome	1080	3.20 (0.93)	1 – 5	580	3.29 (0.91)	1 – 5
Sense of Accomplishment				579	2.98 (1.18)	1 – 5
Mental Health						
Distress	1079	7.04 (4.26)	0 – 24	580	6.96 (4.76)	0 - 24
Social Identity Threat						
Stereotype threat-race	1079	2.65 (1.43)	1 – 7	579	2.53 (1.4)	1 – 7
Stereotype threat- gender	1079	2.69 (1.36)	1 – 7	579	2.83 (1.41)	1 – 7
Research Effort						
Research hours				584	10.88 (14.91)	0 – 100

	<i>N</i> (proportion) yes	<i>N</i> (proportion) no
Has Master's	316 (.292)	765 (.708)

Note. All distal outcomes measured at end of the semester (EOW4), except for institutional trust, which we measured the time point before the end of the semester (EOW3).

preparedness for a Ph.D. program. Students also reported the number of years of research experience they had (1) during their undergraduate education and (2) between college and beginning their doctoral program (including research as a master's student, if applicable). Finally, students reported on two items about how well their previous research experience, during their undergraduate education and after, had prepared them for their doctoral program, from 1 (*not well at all*) to 5 (*extremely well*).

Self-Evaluations. Neuroticism, self-esteem, and self-efficacy are theorized to share the same underlying construct, which can be called *core self-evaluations* (Judge et al., 2002). We were interested in these three measures as well as other forms of self-evaluation: academic self-control, imposter syndrome, and sense of accomplishment. We examined neuroticism and self-control as part of Phase 2 to further understand more stable characteristics of our classes and the other outcomes primarily as part of Phase 3 to examine distal links between classes and self-evaluations. More positive self-evaluations are linked to greater task persistence and performance (e.g., Erez & Judge, 2001; Shrauger & Sorman, 1977; Woodcock & Bairaktarova, 2015).

Three neuroticism items were taken from John et al.'s measure of Big Five personality (2008) (e.g., "I am someone who worries a lot") and rated on a Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). We examined this outcome at baseline only, as neuroticism is considered a fairly stable trait characteristic (e.g., Magnus et al., 1993), and we explored it to get a better idea of the students in each class, a Phase 2 goal. We averaged the three items to create one scale score with higher scores reflecting greater neuroticism ($M = 4.09$, $SD = 1.36$, $\alpha = .76$).

Academic self-control was measured with two items taken from Yeager et al. (2014a), who originally had four items. These items were rephrased for a graduate student population.

Item 1 was “I pay attention and resist distraction in my work.” Item 2 was “I get to work right away rather than procrastinating.” Participants responded using a Likert scale from 1 (*not at all like me*) to 5 (*very much like me*). We assessed items separately instead of averaging them because the correlation was relatively low ($r = .60$). Self-control is also relatively stable over time, particularly in the short-term, and we examine it at baseline alone as part of Phase 2 (e.g., Hay & Forrest, 2006).

A well-established single-item scale assessed self-esteem at baseline and the end of the semester, “[Right now,] I have high self-esteem” (Robins et al., 2001). “Right now” was only included at the end of the semester, as the baseline measure was meant to capture more of a trait-like, stable self-esteem and later measures were meant to capture more state-like self-esteem. Responses ranged from 1 (*not at all true*) to 5 (*very true*).

We previously described self-efficacy because it is also a component of an indicator variable (i.e., belonging composite). We measured self-efficacy at baseline and the end of the semester ($\alpha = .84$). We also previously described imposter syndrome because it is an indicator but we also assessed this at both baseline and end of semester ($\alpha = .80$).

For sense of accomplishment at the end of the semester, students responded to a single, author-created item, “I have accomplished meaningful things so far in graduate school,” on a scale from 1 (*Not at all true*) to 5 (*Extremely true*).

Attitudes towards Graduate School and Academia. For Phase 3, we wanted to examine how class membership predicts attitudes towards graduate school and academia after a semester. We explore general interest, confidence in the decision to pursue a Ph.D., determination to finish the Ph.D., preference for an academic career, evaluations of graduate

school, institutional trust, and burnout. Such attitudes predict academic persistence (e.g., Ramage, 1990; Tao & Gloria, 2019).

We previously described interest because it is an indicator but we assessed it at both baseline and end of semester ($\alpha = .83$).

Using author-created items, students reported on how often they felt they made the right choice in pursuing a Ph.D. from 1 (*never*) to 4 (*always*) and if they ever doubted this decision from 1 (*I never doubt my decision*) to 4 (*I frequently doubt my decision*) at baseline and end of the semester, or end of week 4 (EOW4). We examined these items separately given the relatively low correlation ($r = -0.65$).

Four items, taken from a 16-item scale (Nieuwhof et al., 2004) but reworded for a graduate student context, captured determination, or strength of motivation to continue the Ph.D. at baseline and EOW4. Students rated these items on a Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*). We examined these items separately given the relatively low reliability ($\alpha = .55, .58$ for baseline and end of semester, respectively). Item 1 was “Even if I could hardly maintain my social life, I would still continue graduate school,” item 2 “It wouldn’t really bother me too much if I could no longer study my field,” item 3 “I would quit studying if it became apparent that there were no jobs after graduation.,” and item 4 “I would like to study my field, even if I have to spend a lot of time on topics that later turn out to be a waste of time.” Items 2 and 3 were reverse coded, so that higher scores indicate higher motivation.

Using an author-created item, participants at baseline and the end of the semester indicated how strong their preference was for an academic (university) job after getting their Ph.D. on a sliding scale, anchored at -5 (*strongly prefer non-academic*), 0 (*equal preference for academic and non-academic*), and 5 (*strongly prefer academic*).

At the end of the semester, using author-created items, students evaluated overall how graduate school was going for them so far on a 0 to 100 slider anchored at *Not at all well*, *A little well*, *Moderately well*, *Very well*, and *Extremely well*. They also evaluated how much fun they were having in graduate school so far on a 0 to 100 slider anchored at *None or very little*, *A little*, *A moderate amount*, and *A lot*. Given the correlation ($r = .71$) was not as high as desired to average two items, we examined these items separately.

We assessed institutional trust/fairness at the end of survey week 3 (EOW3), the survey week before the end of the first semester (i.e., in the latter half of the semester). Participants reported how much four groups of people – faculty, administrators, other students at their university, and other people in their field – treated them fairly. We adapted this from one of the items from Yeager et al. (2014b) that was used to measure school trust but split it into four items to better reflect the environment of graduate school; the scale was also widened from a 6-point to a 7-point Likert scale (to be consistent with other measures) with responses ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). We used the average of the four items, with higher scores indicating greater trust ($\alpha = .88$).

We measured six items related to burnout at the end of the semester, taking the two items with the highest factor loadings from each of the three original subscales of the Maslach Burnout Inventory (Maslach & Jackson, 1981). The items were rephrased to better suit a graduate school context. Students responded to items assessing potentially negative feelings about their work in graduate school (e.g., “I feel emotionally drained from my work”) on a Likert scale from 1 (*never*) to 7 (*every day*). Analysis revealed improved reliability by removing the two items from the accomplishment subscale (e.g., “I feel I'm positively influencing other people's lives through my work in grad school”) ($\alpha = .83$), so we removed these items.

Interpersonal Relations and Perceived Fit. As part of Phase 3, we also explored outcomes related to student perceptions of their fit in graduate school and their interpersonal lives (e.g., sense of belonging, perceived social support), given social support and integration predicts academic persistence at the undergraduate level and in doctoral programs (e.g., Nicpon et al., 2006; Rockinson-Szapkiw et al., 2016).

We previously described belonging because it is also an indicator, but we assessed it at both baseline and end of semester ($\alpha = .85$).

At the end of the semester, we assessed level of isolation with three author-created items (e.g., “To what extent are you isolated in your day to day activities?”) from 1 (*not at all*) to 5 (*extremely*). We also assessed level of loneliness at the end of semester with three separate items (e.g., “How often do you feel left out?”) taken from Hughes et al. (2004). Responses ranged from 1 (*never*) to 5 (*almost always*). Given that the items from the isolation and loneliness scales were reliable in a scale when combined ($\alpha = .81$) and our desire to reduce the number of indicators, we averaged the six items together.

Perceived social support was measured using an author-created item at baseline and end of the semester. Participants were asked how true it was that “There are people I can count on to support me” from 1 (*not at all true*) to 5 (*extremely true*).

Three items adapted from Walton and Cohen (2007) captured belonging uncertainty at baseline and the end of the semester from 1 (*strongly disagree*) to 7 (*strongly agree*): “Sometimes I feel like I belong in grad school and sometimes I feel like I don't belong,” “When something good happens, I feel like I really belong in grad school,” and “When something bad happens, I feel like maybe I don't belong in grad school.” Students were instructed to answer

based on their experiences in graduate school so far. Reliability was low ($\alpha = .61, .69$ for baseline and end of semester, respectively), so we assessed the three items separately.

A single author-created item captured perceived similarity to colleagues at baseline and end of the semester, “How similar or different to other people in your department do you see yourself?” At the end of the semester, responses were on a slider from 0 to 100, with anchors *Very different, Moderately different, Slightly different, Slightly similar, Moderately similar, and Very similar*. However, due to a logistical issue, at baseline responses were on a 6-point Likert scale from 1 (*Very different*) to 6 (*Very similar*).

Mental Health. We previously described distress because it is also an indicator, but we assessed it at baseline and end of semester ($\alpha = .88$).

Social Identity Threat. We also previously described stereotype threat based on gender and race and measured them at both baseline and end of semester ($\alpha = .92, .93$, for gender and race at semester’s end, respectively).

Research Effort. At the end of the semester (EOW4), participants were asked, “How many hours did you spend working on research activities over the past week? Give your best estimate.” We included this outcome primarily as an exploration of how research effort might be related to psychological threat (i.e., class membership). Presumably, doctoral students need to spend sufficient time on research to succeed given the nature of doctoral education, but students who are feeling disengaged might not exert as much effort into their research as engaged students, which could ultimately hinder success.

CHAPTER 3. RESULTS

Analytic Strategy

For Phase 1, the delineation of classes, we considered several indicator variables initially, and the process to select a useful set of indicator variables was iterative given the goal to develop useful models. We conducted multiple LCAs with different combinations of potential indicator variables (all related to psychological threat, with many indicated by past literature as relevant to academic outcomes). We looked for indicators that distinguished among the classes (i.e., threat profiles) well and for a model that was stable and theoretically interpretable. For instance, we eliminated growth mindset and race-based stereotype threat as potential indicators because they did not distinguish between classes well. Ultimately, we selected one final model with eight indicators and four classes.

As part of Phase 2, we examined how the classes differed by demographic variables often used as proxies for psychologically threatened groups (e.g., gender, SES). One of our goals in Phase 2 was to test whether our interpretations of the risk level of the classes mapped onto these demographic variables in expected ways. We chose the following demographic variables to align with this goal: gender, sexual orientation, race/ethnicity, socioeconomic status (SES), first-generation status, and international student status. To explore these using the proper LCA technique (BCH; Bakk & Vermunt, 2016; Dziak et al., 2016), we dummy coded them. We collapsed the gender categories into male and not male (which was primarily female students) given the low number of students who identified as outside of the female/male binary ($n = 20$) and the potential for both female and genderqueer identities to suffer heightened psychological threat (e.g., Bower-Brown et al., 2021); we coded male as zero to be the reference group. Results did not meaningfully differ when only male and female identified students were included. We

collapsed the categories for sexual orientation into heterosexual and queer—used here to denote non-heterosexual identities—with heterosexual as the reference group. First-generation students were compared to students who were not the first in their family to attend college, the reference group. We dichotomized SES, such that students who identified themselves as < 5 were compared to the reference group of students who identified themselves as being ≥ 5 on the ladder. This cutoff allows us to compare students who see themselves as having a lower than average (<5) socioeconomic status with those who see themselves as having average (5) or above average (>5) SES. Race/ethnicity was coded such that students who were international, only White, only Asian, or only White and Asian were the reference group and non-international students who identified as Native American, Hispanic, or Black in any combination with other identities were coded as 1 in order to get the proportions of under-represented minorities by class. We coded race/ethnicity in this way because Native American, Hispanic, and Black students are underrepresented among people with STEM doctoral degrees whereas White and Asian students are not underrepresented (National Research Council, 2011; National Science Foundation, 2018). International students were not grouped with under-represented minorities regardless of race or ethnicity given the unique racial context of the United States, which might not meaningfully apply to students from other countries. Of course, international students may have a different experience in graduate school compared to non-international students, which we test by including a separate variable, where non-international was the reference group compared to international students.

For Phase 2, we not only aimed to examine the demographic distribution of classes, but we also wanted to examine how class membership predicts baseline outcomes in order to validate and further develop class descriptions. To accomplish both of these goals, we used the

BCH procedure, which is the currently recommended approach to predict binary and continuous outcomes from class membership (Bakk & Vermunt, 2016; Dziak et al., 2016). The BCH procedure entails linear and logistic regression to predict outcomes from class membership, and bases analyses on modal class assignment (i.e., assignment of individuals to their most likely class) with measurement-error weighting. Alternative approaches include classify-analyze, which is contraindicated because it does not take into account the uncertainty associated with assigning individuals to classes (Nylund-Gibson et al., 2019).

For Phase 3, we predicted distal outcomes in the control condition using the entire sample. For Phase 4, we examined how class membership moderated intervention efficacy on these distal outcomes. We accomplished both of these goals in the same set of analyses, which again required the BCH procedure. Each analysis included two dummy coded intervention variables such that the control condition was the reference group. Whenever possible, we included the baseline measure of the outcome as a covariate. Baseline covariates were only available for some measures due to logistical concerns regarding baseline survey length; moreover, some outcomes were only relevant after time in graduate school (e.g., burnout). In all analyses with baseline covariates, covariates were grand-mean centered to facilitate interpretation of lower-order coefficients, which is particularly important because the intercepts are how we examine control condition students for Phase 3.

Model Selection

For Phase 1's initial step of selecting a model, we compared models with varying number of classes, beginning with two classes and continuing on until fit indices revealed more classes would not improve the model. We did this with many unique combinations of indicator

variables. This iterative process ultimately resulted in one interpretable, robust model with eight indicators and four classes.

We conducted latent class analyses in Mplus version 8.4 (Muthén & Muthén, 1998-2017). We based the number of classes tested for each unique set of indicators and the ultimate model selected on fit indices, which included Bayesian information criterion (BIC; Schwarz, 1978), sample-size adjusted BIC (aBIC; Sclove, 1987), Akaike information criterion (AIC; Akaike, 1974), the bootstrapped likelihood ratio test (BLRT; see Nylund et al., 2007), and the Vuong-Lo-Mendell-Rubin adjusted likelihood ratio test (VLMR-LRT; see Nylund et al., 2007). Fit indices for the model with the ultimately selected sets of indicators are shown in Table 3. Lower values for BIC, aBIC, and AIC indicate relatively better balance between parsimony and model fit. We emphasized the BIC and BLRT in particular given evidence showing their unique strength in identifying the ideal number of classes (Nylund et al., 2007). Log-likelihood values inform AIC, BIC, aBIC, BLRT, and VLMR-LRT indices. We also present entropy, a measure of class separation for which higher values indicate greater separability and higher classification utility, but this was not used as a model selection statistic (Masyn, 2013). We emphasized theoretical model interpretability in optimal model selection (e.g., Collins & Lanza, 2010). Model identification for all models was evaluated using 1000 sets of random initial stage starting values and 500 final stage starts.

Posterior probabilities, a product of LCA, are each individual's probabilities of membership in each class; they are used for the measurement-weighting of the BCH procedure. The average posterior probabilities, which reflect the overall certainty of classification of individuals in each class, are presented in Table 4. Higher average posterior probabilities reflect greater certainty that members of a class are assigned correctly.

Table 3*LCA Fit Indices*

K	LL	BIC	aBIC	AIC	BLRT p	VLMR-LRT p	Entropy
1	-8354.490	16820.750	16769.931	16740.980	--	--	--
2	-7872.915	15976.355	15871.540	15811.829	<.0001	<.0001	.714
3	-7731.617	15812.515	15653.705	15563.233	<.0001	.0066	.716
4	-7638.886	15745.810	15533.004	15411.772	<.0001	.0001	.702
5	-7592.008	15770.810	15504.009	15352.016	<.0001	.5328	.700
6	-7564.315	15834.179	15513.382	15330.629	<.0001	.0203	.691
7	-7538.127	15900.560	15525.767	15312.254	.0128	.2993	.703
8	-7519.591	15982.244	15553.456	15309.182	1.0000	.7636	.724

Note. K = number of classes; LL = log-likelihood; BIC = Bayesian Information Criterion; aBIC = Sample-size adjusted BIC; AIC = Akaike Information Criterion; BLRT = bootstrapped likelihood ratio test; VLMR-LRT = Vuong-Lo-Mendell-Rubin adjusted likelihood ratio test; p = p -value; **Bolded** values indicate “best” fit for each respective statistic. Entropy is included in the table for brevity but should not be used as a model selection statistic (Masyn, 2013).

Table 4

Classification Probabilities for the Most Likely Latent Class Membership (Column) by Latent Class (Row)

Class	1	2	3	4
1	.834	.060	.086	.020
2	.085	.853	.062	.000
3	.095	.046	.797	.063
4	.035	.000	.104	.861

Note. Bolded values indicate average posterior probabilities.

The two kinds of parameters estimated in LCAs are latent class prevalences (i.e., the proportion of the sample in each class) and item response probabilities, which represent the probability of answering a certain way given membership in each class. These parameters are presented in Table 5. They can also be seen in Figure 2 to facilitate comparison across indicators and classes.

Model Description

In this section, we present the description of each of the four class in the selected model in order from lowest to highest threat based on our interpretations of the classes.

Engaged and Confident class.

Class 4, about 21% of the sample, is the group with the least amount of psychological threat. The vast majority in Class 4 are in the highest level of the belonging composite, which reflects academic belonging, graduate school self-efficacy, and psychological need satisfaction. The vast majority in Class 4 also have low gender threat, low psychological distress, low worry, high grit, and high interest. Compared to the other classes, these students also have the least amount of imposter syndrome, with over half (59%) having low imposter syndrome. These students vary in their researcher/academic identification, but notably only 21% of students reported, on average, less than “somewhat agree” on the items assessing academic identification. Hence, most of these students do feel at least moderately identified with academia and as a researcher, and in relative terms, they feel more strongly identified than Classes 2 or 3. Overall, this class appears psychologically prepared for and engaged with graduate school with few concerns. We label Class 4 as *Engaged and Confident*.

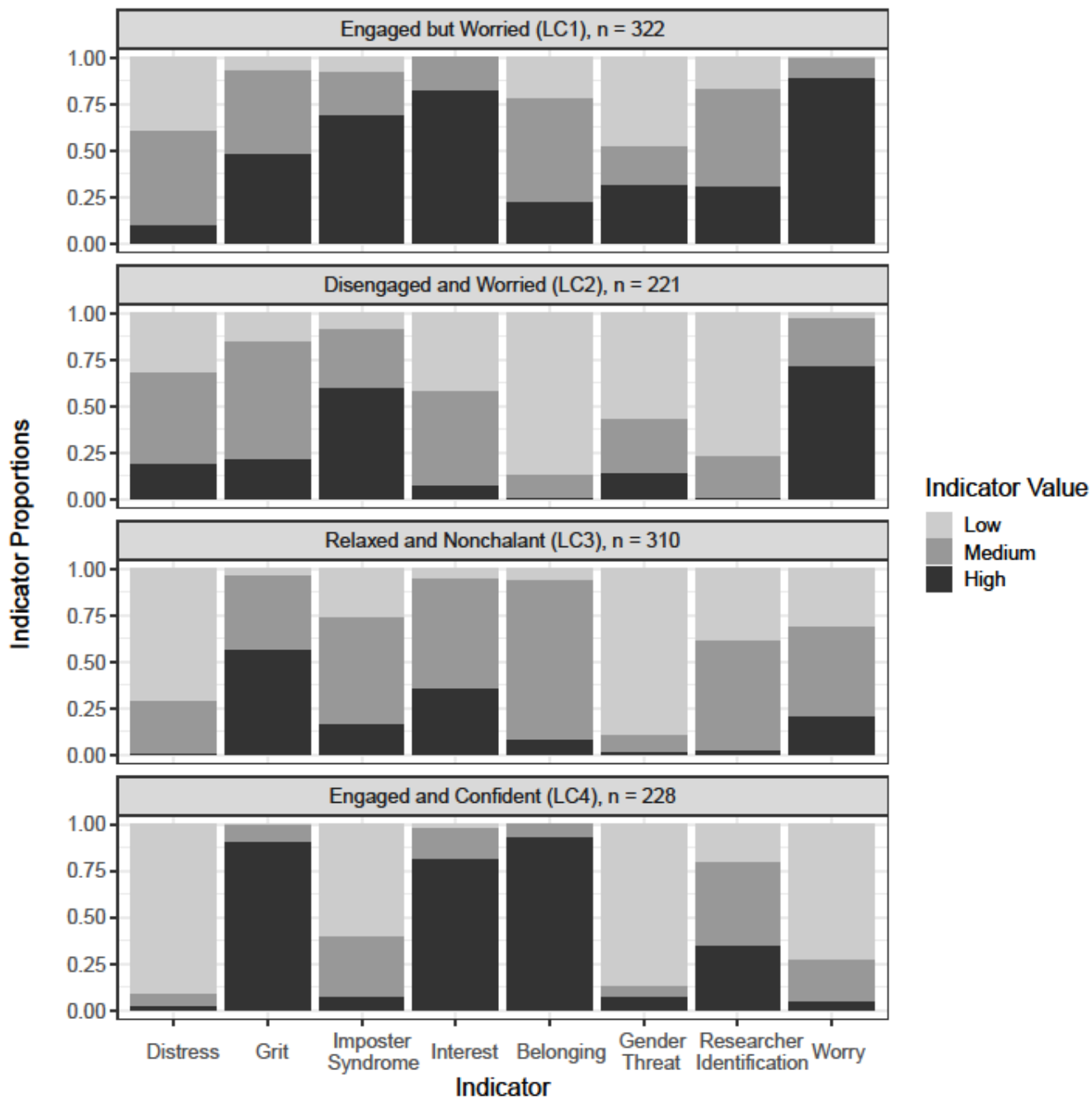
Table 5*Parameter Estimates from the Selected 4-Class Model*

	Class 1 (Engaged but Worried)	Class 2 (Disengaged and Worried)	Class 3 (Relaxed and Nonchalant)	Class 4 (Engaged and Confident)
Latent class membership probabilities <i>N</i> (%)	322 (0.298)	221 (0.204)	310 (0.287)	228 (0.211)
Indicators	Category 1 (Low) Category 2 (Medium) Category 3 (High)			
Belonging Composite	.236 .551 .213	.823 .177 .000	.088 .793 .119	.000 .133 .867
Gender Threat	.500 .201 .299	.570 .280 .150	.884 .098 .018	.863 .064 .073
Worry	.003 .138 .858	.036 .265 .699	.303 .450 .247	.715 .234 .052
Distress	.421 .486 .093	.322 .485 .193	.693 .300 .007	.896 .089 .015
Imposter Syndrome	.078 .262 .660	.086 .310 .604	.274 .538 .187	.592 .330 .078
Grit	.067 .447 .486	.163 .624 .213	.034 .423 .543	.004 .099 .896
Researcher Identification	.185 .518 .297	.763 .237 .000	.382 .592 .026	.208 .455 .337
Interest	.000 .227 .773	.415 .481 .104	.058 .556 .386	.029 .183 .788

Note. Bold indicates high values ($\geq .66$) to facilitate interpretation.

Figure 2

Distribution of Indicators by Class



Note. LC = Latent Class.

Relaxed and Nonchalant class.

Class 3, about 29% of the sample, is distinguished by the vast majority of students scoring (1) in the middle level on the belonging composite, (2) low on gender threat, and (3) low on distress. These students appear to not be facing social identity threat or mental health issues; however, they face mixed amounts of other psychological threat. Their levels of general worry vary, with nearly half (45%) not feeling strongly either way about whether or not they are worried about others perceiving them negatively. Similarly, over half (54%) have medium levels of imposter syndrome. Class 3 students either have medium or high grit, medium or low researcher identification, and medium or high interest. Taken together, there is some variation among the students in Class 3, but overall, these students do not seem to be particularly high threat based on their response patterns; however, they also are not quite as sure of themselves as the *Engaged and Confident* class. Given this and the general lack of strong positive or negative feelings, we label Class 3 as *Relaxed and Nonchalant*.

Engaged but Worried class.

Class 1, the largest class at 30%, is distinguished by having the vast majority of members score high on (1) general worry, (2) imposter syndrome, and (3) interest. There is more variation on the belonging composite than in the other classes, although the majority (55%) of members fall in the middle. Although about half of the students in Class 1 report low gender threat, there is more gender threat in this class than in the others, with 30% reporting relatively high levels of gender threat. Students are either low or medium on distress for the most part, and medium or high on grit. They are most similar to the *Engaged and Confident* class on researcher/academic identification. Overall, Class 1 appears to face some psychological threat, including SIT, but also seems to be highly engaged in their paths. We label Class 1 as *Engaged but Worried*.

Disengaged and Worried class.

Class 2, the smallest class at about 20% of the sample, consists of students who, for the most part, are low on the belonging composite, high in general worry about how others will perceive them, and low in researcher/academic identification. While the majority (57%) experience minimal gender threat, more students in this class than those in the *Relaxed and Nonchalant* or *Engaged and Confident* classes experience moderate to high levels of gender threat. Students in this class have varying levels of psychological distress, with almost half having a moderate amount and 19% having clinically concerning levels of distress, which is more than any of the other classes. The majority of students also have high levels of imposter syndrome, although a sizable proportion have medium levels. A higher proportion of students in Class 2 than in any other class scores low on grit and a lower proportion scores high on grit, although most in Class 2 have a medium level of grit. The vast majority of students in Class 2 score low or medium on interest, which is also worse than the other classes. Relatively speaking, this class has the most psychological threat. We label Class 2 as *Disengaged and Worried*.

Demographic Distribution of Classes (Phase 2)

As the first part of Phase 2, to see how classes differ based on demographic variables (e.g., race/ethnicity), we dummy-coded demographic variables (as previously described) and used the automated BCH procedure (see Table 6). The output is Chi-square results using estimated means of the outcome by class; in this case, these estimated means correspond to proportions given the variable coding. The overall Chi-square significance test thus answers the question of whether the classes overall are different in the proportions of students by the given demographic variable. We can also see all the significance tests for pairwise comparisons

Table 6*Proportions of Demographic Groups by Latent Class*

Group	Engaged but Worried	Disengaged and Worried	Relaxed and Nonchalant	Engaged and Confident	Chi-Square Results <i>n</i> = 1081
Gender: Female/ Genderqueer	.69 ^{a, c, d}	.49 ^{b, c}	.36 ^{a, b}	.41 ^d	$\chi^2(3) = 48.74$ $p < .001$
Race/ethnicity: Underrepresented Minority	.10	.05	.06	.06	$\chi^2(3) = 3.87$ $p = .276$
Sexual orientation: Queer	.34 ^{a, b, c}	.15 ^b	.12 ^a	.14 ^c	$\chi^2(3) = 30.47$ $p < .001$
First-generation student: Yes	.23	.20	.15	.23	$\chi^2(3) = 4.08$ $p = .254$
Socioeconomic status: Low	.22 ^b	.26 ^c	.23 ^a	.13 ^{a, b, c}	$\chi^2(3) = 12.57$ $p = .006$
International student: Yes	.29 ^{a, c}	.46 ^c	.55 ^{a, b}	.36 ^b	$\chi^2(3) = 25.26$ $p < .001$

Note. Matching letters denote significant pairwise differences at $p < .05$.

between classes (e.g., does the *Engaged but Worried* class have a significantly higher proportion of woman than the *Engaged and Confident* class?).

Overall, the classes significantly differed by gender, sexual orientation, socioeconomic status, and international student status, but not by first-generation status or race/ethnicity. We expected relative overrepresentation of potential at-risk social categories (e.g., women, first-generation students) in the higher risk classes and relative underrepresentation in the lower risk classes. In alignment with this expectation, we found that the largest proportion of female and genderqueer (i.e., non-male) identities, 69% of the members, was among the *Engaged but Worried* students. Similarly, the largest proportion of queer identities, 34%, was found among the *Engaged but Worried* students. Also aligning with expectations, the lowest proportion of low-socioeconomic students (13%) was among the *Engaged and Confident* class. In terms of international students, for whom we did not have clear expectations, the *Relaxed and Nonchalant* class had the highest proportion at 55%. The test for race/ethnicity was likely underpowered given the overall low prevalence of URM students in the sample, and although we do see that the highest percentage of URM students is in the *Engaged but Worried* class, the percentage is still relatively low (10%).

In sum, the *Engaged but Worried* class has overrepresentation of non-male and queer identities, as well as slight underrepresentation of international students compared to other classes. This class also has the highest percentage of URM students; however, the race/ethnicity test did not reach statistical significance. These results suggest that the *Engaged but Worried* class is characterized by relatively high SIT, which coincides with their initial description based on indicators (i.e., they had relatively high gender threat compared to other classes). The *Engaged and Confident* class is underrepresented in terms of low-SES students and somewhat

underrepresented in terms of international students, as well. We expected this lowest threat class to have fewer students who are in at-risk social categories, so the low-SES students being underrepresented is unsurprising. The *Relaxed and Nonchalant* class, and to a lesser extent the *Disengaged and Worried* class, has slight overrepresentation of international students. We did not know what to expect with international students, but these results may indicate that at least some international students are at heightened risk compared to non-international students.

Descriptive Information by Class (Phase 2)

For the other aspect of Phase 2, we examined how class membership was related to concurrent baseline outcomes besides demographic variables to get a better sense for the individual student differences by class. For these analyses, we used the same approach as we did for the demographic variables—the automated BCH procedure that produces Chi-square results (see Table 7). The patterns of means by class for Table 7 results are presented visually in Figures 3 through 8.

Class membership significantly predicted baseline self-evaluation outcomes (neuroticism and academic self-control) and the academic preparation and context outcomes. We present these results by class in order of threat level, starting with the highest threat class.

Disengaged and Worried class.

The *Disengaged and Worried* students were least academically prepared: they had the fewest years of postgraduate prior research experience, they felt the least prepared by prior undergraduate or postgraduate experience, and their class had the smallest proportion of students who previously earned a master's degree. These students were on average pursuing their Ph.D.'s in fields that are ~36% female, which is notable because the class is fairly evenly split by gender (51% male). Thus, female students in this class may tend to be underrepresented in their fields,

Table 7*How Class Membership Predicts Descriptive Outcomes*

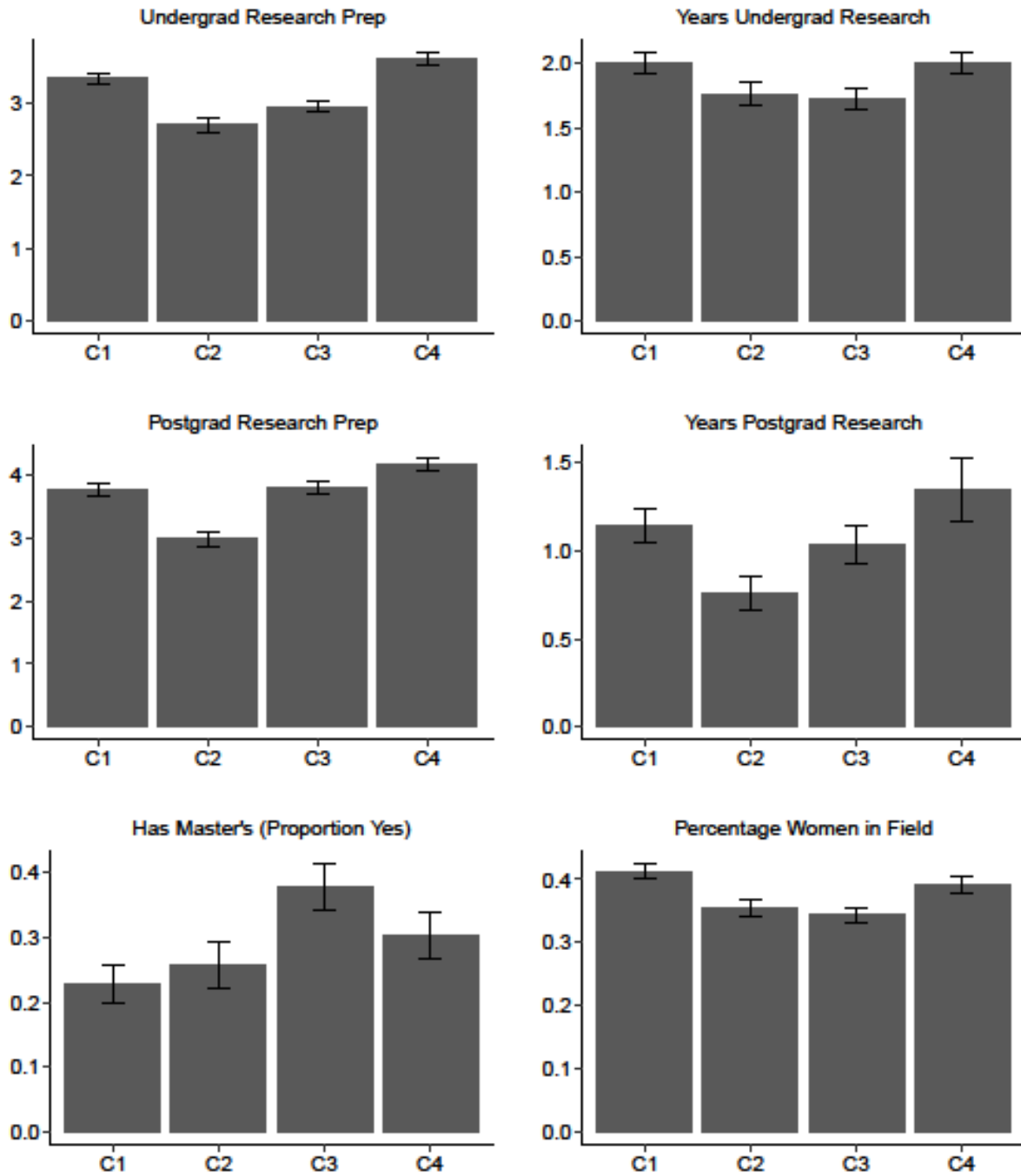
Variable	Engaged but Worried Mean (SE)	Disengaged and Worried Mean (SE)	Relaxed and Nonchalant Mean (SE)	Engaged and Confident Mean (SE)
Academic Preparation and Context				
Has master's (proportion yes)*	0.228 (0.030) ^a	0.257 (0.035) ^b	0.376 (0.036) ^{a, b}	0.303 (0.035)
Undergrad research prep*	3.343 (0.075) ^{a, b, c}	2.697 (0.089) ^{a, d, e}	2.949 (0.074) ^{b, d, f}	3.609 (0.085) ^{c, e, f}
Years undergrad research*	2.000 (0.075) ^{a, b}	1.760 (0.086) ^{a, c}	1.727 (0.083) ^{b, d}	1.995 (0.083) ^{c, d}
Postgrad research prep*	3.782 (0.094) ^{a, b}	2.990 (0.125) ^{a, c, d}	3.817 (0.097) ^{c, e}	4.174 (0.097) ^{b, d, e}
Years postgrad research*	1.146 (0.098) ^a	0.762 (0.093) ^{a, b}	1.034 (0.109)	1.347 (0.179) ^b
Percentage of women in field*	0.412 (.012) ^{a, b}	0.355 (.013) ^{a, c}	0.344 (0.012) ^{b, d}	0.391 (0.013) ^{c, d}
Attitudes towards Grad School/Academia				
Interest composite*	6.731 (0.033) ^{a, b}	5.078 (0.070) ^{a, c, d}	6.020 (0.042) ^{b, c, e}	6.652 (0.037) ^{d, e}
Right decision*	3.506 (0.036) ^{a, b}	2.997 (0.042) ^{a, c, d}	3.422 (0.040) ^{c, e}	3.863 (0.031) ^{b, d, e}
Doubt decision*	2.251 (0.055) ^{a, b}	2.753 (0.061) ^{a, c, d}	2.199 (0.058) ^{c, e}	1.558 (0.052) ^{b, d, e}
Determination #1*	5.299 (0.106) ^{a, b}	4.820 (0.118) ^{a, c}	4.650 (0.119) ^{b, d}	5.365 (0.120) ^{c, d}
Determination #2*	5.919 (0.092) ^{a, b}	4.587 (0.115) ^{a, c, d}	4.930 (0.110) ^{b, c, e}	5.887 (0.098) ^{d, e}
Determination #3*	4.972 (0.117) ^{a, b}	4.027 (0.128) ^{a, c, d}	4.386 (0.118) ^{b, c, e}	5.178 (0.123) ^{d, e}
Determination #4*	4.765 (0.100) ^{a, b}	4.019 (0.106) ^{a, c}	4.196 (0.104) ^{b, d}	4.777 (0.124) ^{c, d}
Preference for academic career*	1.978 (0.191) ^{a, b}	0.857 (0.228) ^{a, c}	0.934 (0.211) ^{b, d}	1.992 (0.229) ^{c, d}
Grad school evaluation EOW4*	67.193 (1.691) ^{a, b}	51.991 (2.460) ^{a, c, d}	69.658 (1.610) ^{c, e}	77.832 (1.886) ^{b, d, e}
Grad school fun EOW4*	62.545 (1.901) ^{a, b}	45.140 (3.016) ^{a, c, d}	64.104 (2.067) ^{c, e}	71.123 (2.353) ^{b, d, e}

Institutional trust EOW3*	5.834 (0.078) ^{a, b}	5.285 (0.151) ^{a, c, d}	5.941 (0.079) ^c	6.214 (0.104) ^{b, d}
Burnout EOW4*	3.471 (0.119) ^{a, b}	3.562 (0.177) ^{c, d}	2.754 (0.128) ^{a, c, e}	2.314 (0.127) ^{b, d, e}
Interpersonal Relations and Perceived Fit				
Perceived social support*	4.223 (0.059) ^{a, b}	3.434 (0.076) ^{a, c, d}	4.001 (0.063) ^{b, c, e}	4.378 (0.063) ^{d, e}
Similarity to colleagues*	3.632 (0.096) ^{a, b, c}	3.184 (0.099) ^{a, d, e}	4.024 (0.091) ^{b, d}	4.183 (0.102) ^{c, e}
Isolation EOW4*	2.831 (0.062) ^{a, b}	3.147 (0.093) ^{a, c, d}	2.667 (0.065) ^{c, e}	2.363 (0.081) ^{b, d, e}
Belonging*	5.152 (0.047) ^{a, b}	4.135 (0.054) ^{a, c, d}	5.176 (0.039) ^{c, e}	6.002 (0.040) ^{b, d, e}
Belonging uncertainty #1*	5.289 (0.106) ^{a, b}	5.083 (0.096) ^{c, d}	4.371 (0.109) ^{a, c, e}	3.015 (0.123) ^{b, d, e}
Belonging uncertainty #2*	6.075 (0.067) ^{a, b, c}	5.097 (0.087) ^{a, d, e}	5.589 (0.071) ^{b, d}	5.744 (0.101) ^{c, e}
Belonging uncertainty #3*	2.936 (0.107) ^{a, b}	2.990 (0.099) ^{c, d}	3.850 (0.101) ^{a, c, e}	4.892 (0.105) ^{b, d, e}
Self-Evaluations				
Neuroticism*	4.962 (0.078) ^{a, b}	4.860 (0.086) ^{c, d}	3.585 (0.091) ^{a, c, e}	2.846 (0.088) ^{b, d, e}
Self-efficacy*	3.118 (0.039) ^{a, b}	2.435 (0.048) ^{a, c, d}	3.027 (0.035) ^{c, e}	3.697 (0.029) ^{b, d, e}
Self-esteem*	2.791 (0.078) ^{a, b, c}	2.306 (0.094) ^{a, d, e}	3.470 (0.073) ^{b, d, f}	4.079 (0.069) ^{c, e, f}
Academic self-control #1*	3.239 (0.066) ^{a, b}	2.740 (0.073) ^{a, c, d}	3.376 (0.064) ^{c, e}	3.782 (0.065) ^{b, d, e}
Academic self-control #2*	2.977 (0.075) ^{a, b}	2.501 (0.081) ^{a, c, d}	3.084 (0.077) ^{c, e}	3.615 (0.081) ^{b, d, e}
Imposter syndrome*	3.885 (0.056) ^{a, b}	3.707 (0.065) ^{c, d}	2.822 (0.054) ^{a, c, e}	2.295 (0.055) ^{b, d, e}
Sense of accomplishment EOW4*	2.989 (0.104) ^{a, b}	2.121 (0.131) ^{a, c, d}	3.104 (0.105) ^{c, e}	3.484 (0.130) ^{b, d, e}
Mental Health				
Distress*	8.822 (0.269) ^{a, b, c}	10.434 (0.364) ^{a, d, e}	5.254 (0.238) ^{b, d, f}	3.768 (0.239) ^{c, e, f}
Social Identity Threat				
Stereotype threat- gender*	3.427 (0.103) ^{a, b}	3.186 (0.105) ^{c, d}	2.133 (0.079) ^{a, c}	1.965 (0.084) ^{b, d}
Stereotype threat- race*	2.902 (0.109) ^{a, b, c}	3.341 (0.111) ^{a, d, e}	2.466 (0.095) ^{b, d, f}	1.904 (0.087) ^{c, e, f}

Research Effort

Research Hours EOW4	11.417 (1.314)	12.494 (2.108)	9.818 (1.386)	10.275 (1.504)
---------------------	----------------	----------------	---------------	----------------

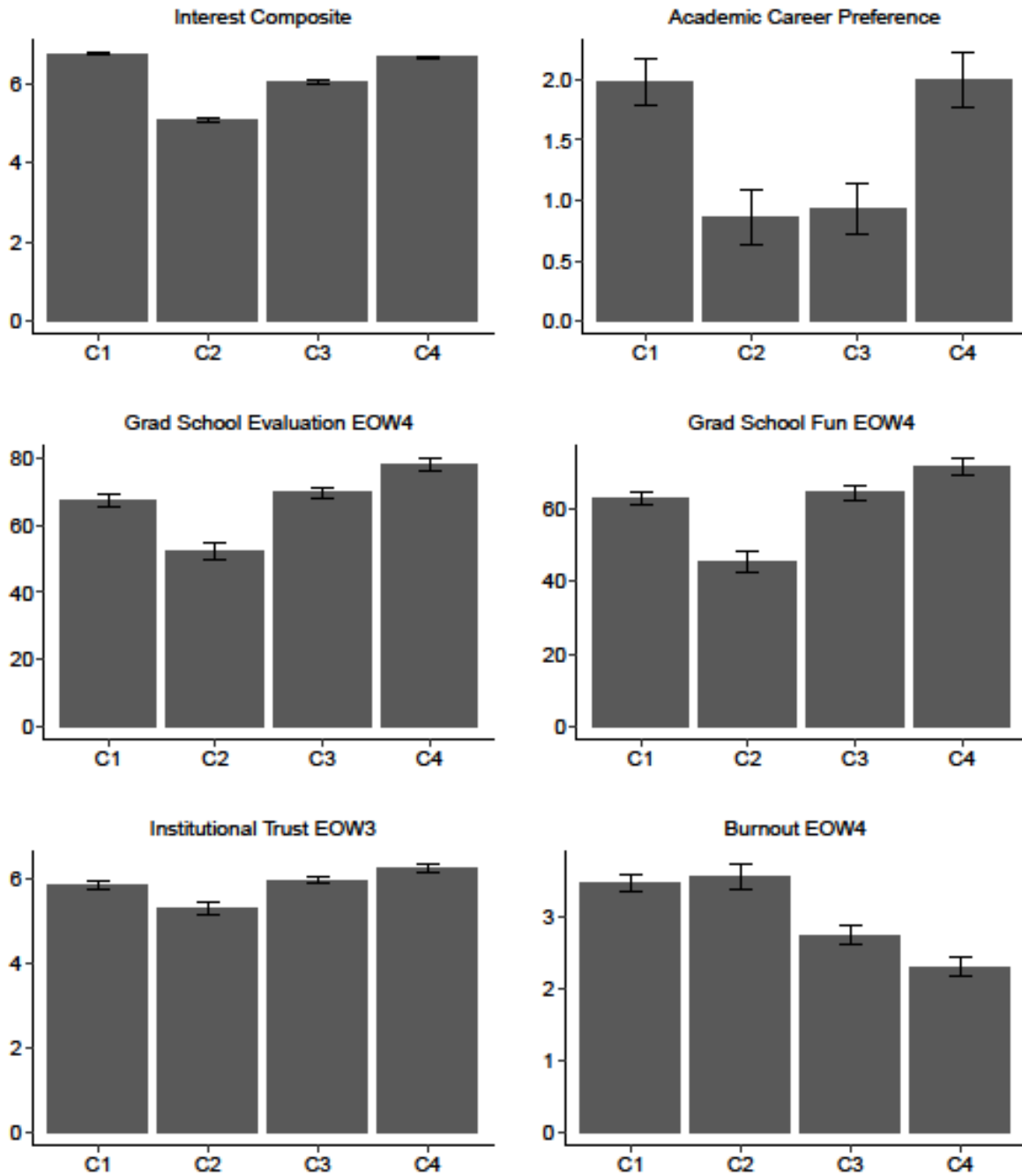
Note. This table contains Phase 2 results, as well as results for outcomes in Phase 3 (e.g., end-of-semester burnout). We do not focus on these Phase 3 results in the text because these analyses collapse across intervention condition, and the Phase 3 analyses focus on the normative experience in the control condition. We also include baseline covariates in this table that we do not discuss in the text because we focus instead on their distal counterpart measures in subsequent phases. Matching letters denote significant pairwise differences at $p < .05$. Asterisks beside outcome name indicate overall Chi-square test was significant at $p < .05$. Number of people in each group are approximated based on the most likely latent class membership. Outcomes are not distal unless denoted EOW3 (end of week 3) or EOW4 (end of week 4, or end of the semester).

Figure 3*Academic Preparation and Context Outcomes by Class*

Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline.

Figure 4

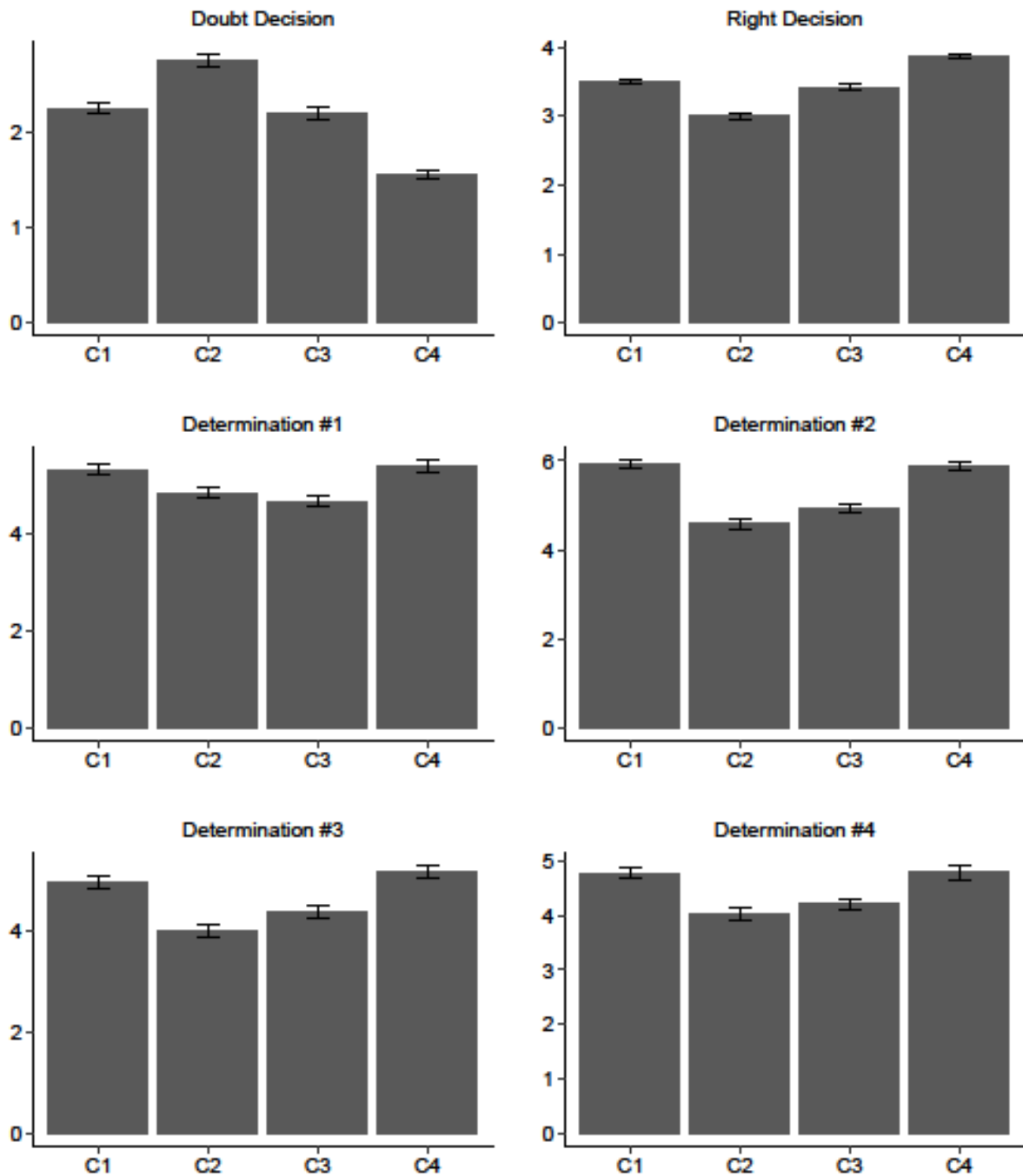
Attitudes towards Graduate School and Academia Outcomes #1 by Class



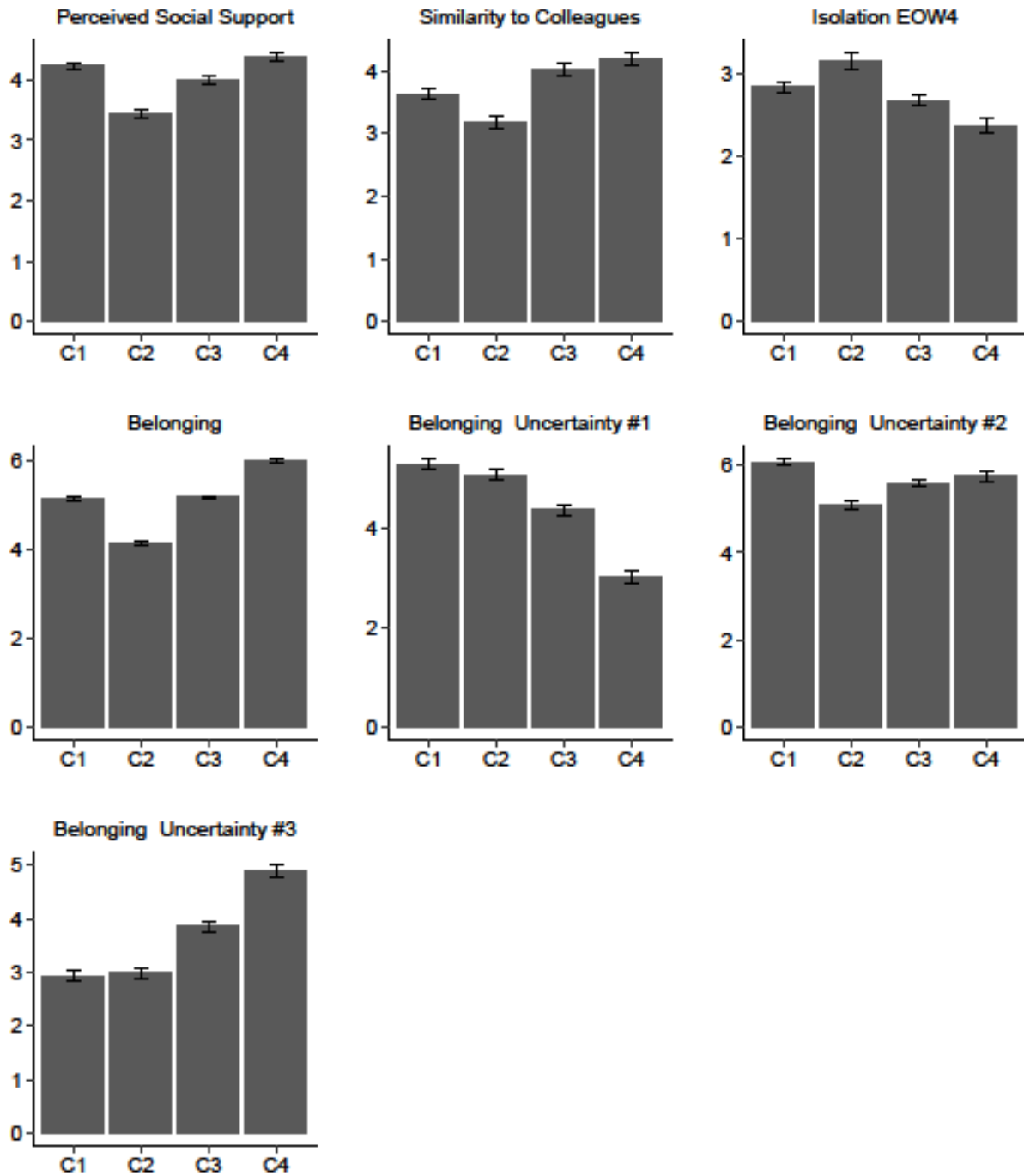
Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline, except if denoted EOW3 or EOW4 (End of Week 3, End of Week 4, respectively).

Figure 5

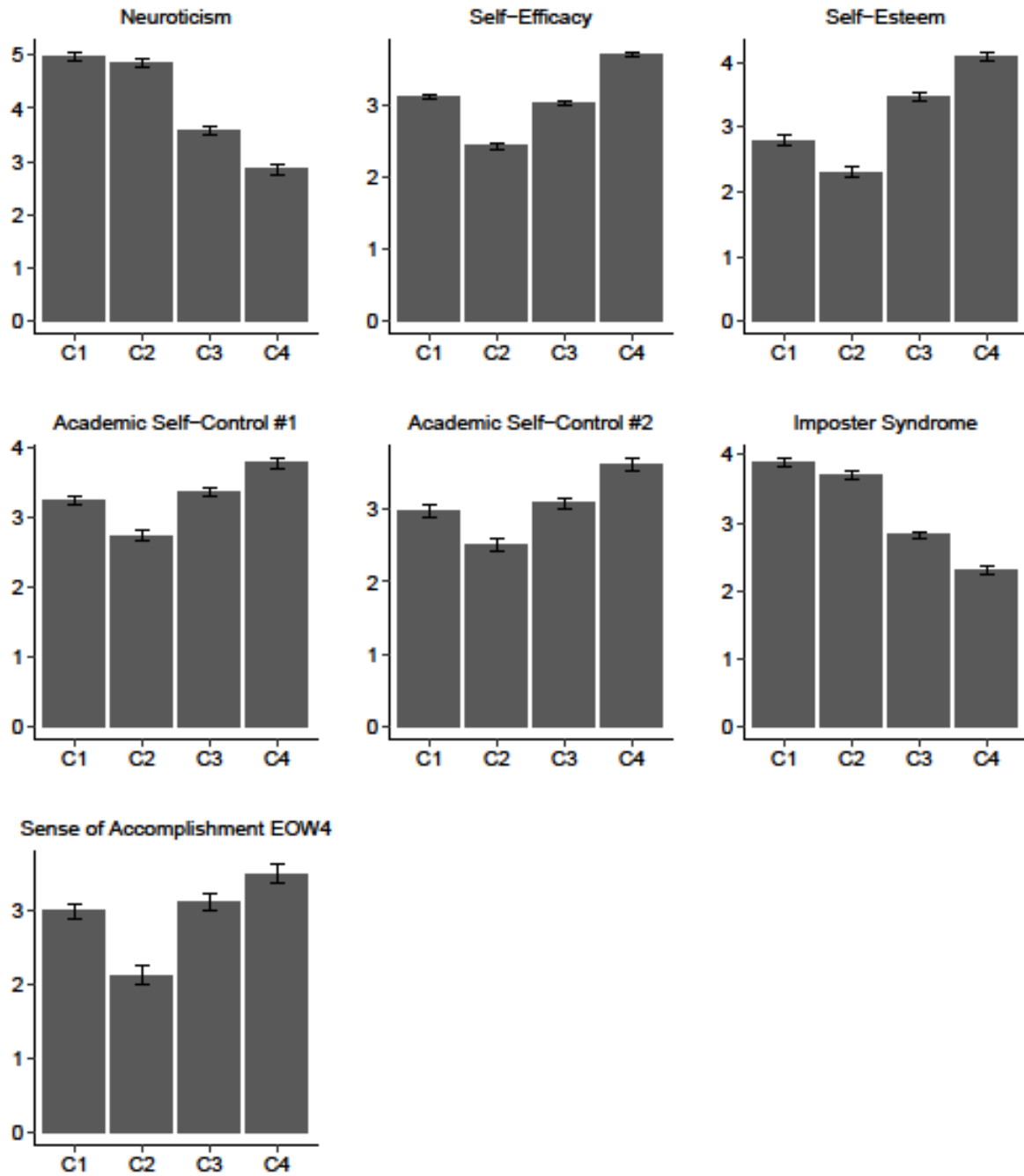
Attitudes towards Graduate School and Academia Outcomes #2 by Class



Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline.

Figure 6*Interpersonal Relations and Perceived Fit Outcomes by Class*

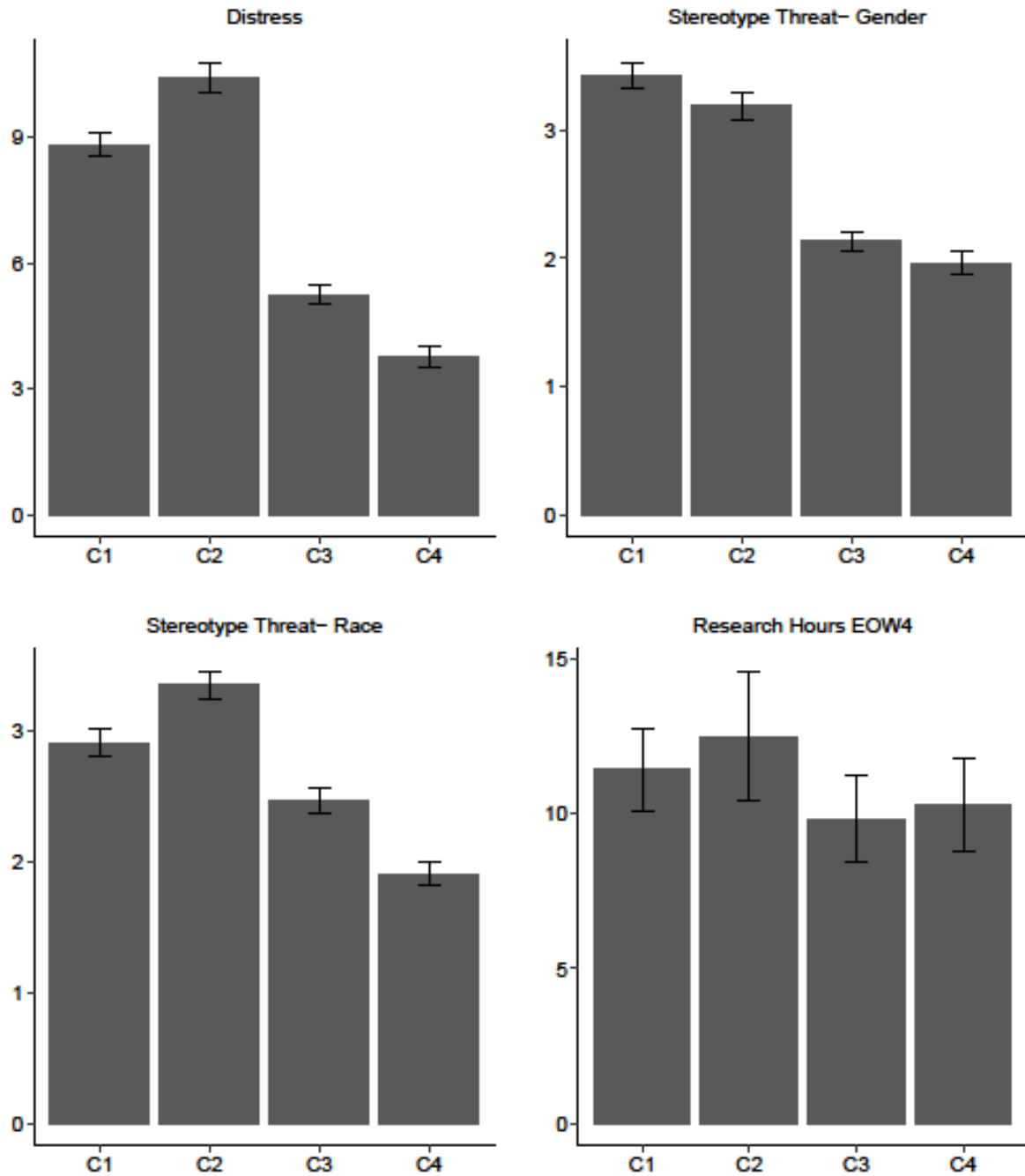
Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline, except if denoted EOW4 (End of Week 4).

Figure 7*Self-Evaluations by Class*

Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline, except if denoted EOW4 (End of Week 4).

Figure 8

Mental Health, Social Identity Threat, and Research Effort by Class



Note. C1: Engaged but Worried (moderately high threat), C2: Disengaged and Worried (high threat), C3: Relaxed and Nonchalant (moderately low threat), C4: Engaged and Confident (low threat). Outcomes measured at baseline, except if denoted EOW4 (End of Week 4).

which can create SIT. Additionally, *Disengaged and Worried* students were higher than the two low-threat classes in neuroticism and lowest in academic self-control, the former of which validates their description as “worried” and the latter of which may portend suboptimal academic outcomes.

Engaged but Worried class.

The *Engaged but Worried* students were fairly well-prepared academically overall, particularly in terms of years of prior experience and perceived preparedness. Interestingly, these students overall had a similar amount of undergraduate research experience as the *Engaged and Confident* students but felt significantly less prepared by that experience—this may indicate that these *Engaged but Worried* students tend to feel less confident than they should, given objective qualifications. This finding may be a sign of SIT, particularly given the majority of this class is female. Moreover, these students tended to pursue their degrees in fields that get the closest to gender parity of the classes, although the average percentage of women by field is still only ~41. The *Engaged but Worried* students, like the *Disengaged and Worried* students, were high in neuroticism and lower than the other two classes in academic self-control (albeit only significantly lower in self-control than the lowest threat class), validating the descriptor “worried” and possibly a sign of relatively worse academic prospects, although academic self-control scores were at about the midpoint of the scale.

Relaxed and Nonchalant class.

The *Relaxed and Nonchalant* students were relatively academically prepared overall, with the highest proportion of students coming in with master’s degrees and tending to be between the lower and higher threat classes on the other academic preparation outcomes. In terms of gender balance in these students’ fields, these students tended to be in male-dominated

fields; this is not necessarily surprising given 64% of them are male. The *Relaxed and Nonchalant* students also scored under the midpoint on the neuroticism scale on average, still higher than the *Engaged and Confident* students but significantly lower than students from the other two classes, supporting the notion that they are not particularly worried on average. For academic self-control, they tended to score higher than students from the highest threat class and lower than the *Engaged and Confident*, the lowest threat, students. Overall, these results further support that *Relaxed and Nonchalant* students tend to fall in the middle in terms of threat compared to the other classes but are more similar to the *Engaged and Confident* class than the higher threat classes.

Engaged and Confident class.

The *Engaged and Confident* students tended to be well-prepared academically; most saliently, they perceived their prior research experience as most helpful in preparing them for graduate school, which validates their “confident” descriptor. These students tend to be in fields that have not achieved gender parity (as is true for every class given that the sample is primarily STEM students) but are closer to parity than the fields of students in the *Relaxed and Nonchalant* or *Disengaged but Worried* classes; the average field’s percentage of women for these students (~39%) is close to the percentage of women in the class (~41%). The *Engaged and Confident* students reported the lowest levels of neuroticism and the highest levels of academic self-control, supporting the idea that they are unworried and the possibility of high academic achievement.

What Class Membership Predicts (Phase 3)

For Phase 3, we explored how class membership predicts distal outcomes. We focus on the control condition students here in order to disentangle any influences of intervention in our examination of what graduate school is like for typical students, that is, those who do not

undergo intervention. The results for the analyses looking specifically at control condition students are evident in the intercepts and the accompanying contrasts to test the intercepts against each other (the first two columns) in Table 8, which also contains the results for Phase 4. For these analyses, we again used the BCH procedure in order to account for the uncertainty associated with assigning individuals to classes. Besides class membership, we included two dummy-coded intervention variables. When a baseline measure of a distal outcome was available, we also included it in the model and allowed it to vary across classes; we grand mean centered the baseline measures so that the intercepts represented class-specific distal means for someone whose baseline score was at the sample mean. Intervention condition was dummy coded with control condition as the reference group, so that intercepts represent expected values for students who do not receive intervention. We examined the class differences in intercepts both as an omnibus test—do the expected distal outcome levels of students in the control condition significantly differ by class?—and using specific contrasts. For the specific contrasts, we compared the *Engaged and Confident* class to each of the others and compared the *Disengaged and Worried* class to the other three. We chose these comparisons because we are interested in ascertaining the relative threats facing students and these two classes are the highest and lowest threat classes, respectively. We describe the rest of the details for these analyses in Phase 4, where we discuss condition effects. For now, we focus on the intercepts and the differences between them so that we can better understand control condition students.

The results for these Phase 3 and 4 analyses are akin to regression results, and as such, the coefficients can be interpreted in a similar way. As mentioned, the intercepts can be interpreted as the expected outcome level of a student in the indicated class and in the control condition at, if applicable, the mean of the outcome at baseline. For example, 6.390 is the

Table 8*End-of-Semester Differences by Latent Class as a Function of Intervention Condition, Controlling for Baseline Levels*

Outcomes	End-of-Semester Mean (Intercept)	Class Differences in Intercepts	Baseline Predictor	Control vs Belonging	Class Differences in Control vs Belonging	Control vs Affirmation	Class Differences in Control vs Affirmation
	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>	<i>B (SE)</i>
Attitudes towards Grad School/Academia							
Interest*							
<i>Engaged & Confident</i>	6.390 (0.134)***	<i>ref</i>	0.229 (0.135)†	-0.069 (0.182)	<i>ref</i>	-0.089 (0.154)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	6.010 (0.117)***	-0.380 (0.194)†	0.560 (0.164)**	0.106 (0.154)	0.174 (0.259)	-0.185 (0.151)	-0.097 (0.234)
<i>Engaged but Worried</i>	6.113 (0.650)***	-0.277 (0.664)	0.438 (1.541)	0.052 (0.171)	0.121 (0.252)	0.024 (0.170)	0.113 (0.232)
<i>Disengaged & Worried</i>	5.514 (0.304)***	-0.877 (0.331)**	0.473 (0.218)*	0.262 (0.269)	0.331 (0.322)	-0.035 (0.248)	0.054 (0.289)
<i>Disengaged vs Other Classes</i>		0.658 (0.380)†			-0.233 (0.290)		-0.049 (0.266)
Right Decision†							
<i>Engaged & Confident</i>	3.685 (0.133)***	<i>ref</i>	0.149 (0.190)	-0.161 (0.128)	<i>ref</i>	0.153 (0.117)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.364 (0.084)***	-0.322 (0.167)†	0.523 (0.090)***	0.134 (0.123)	0.294 (0.192)	0.096 (0.116)	-0.057 (0.179)
<i>Engaged but Worried</i>	3.372 (0.084)***	-0.313 (0.160)†	0.409 (0.095)***	0.033 (0.109)	0.194 (0.171)	-0.064 (0.119)	-0.217 (0.170)
<i>Disengaged & Worried</i>	3.138 (0.177)***	-0.547 (0.220)*	0.314 (0.220)	0.157 (0.164)	0.317 (0.206)	-0.239 (0.181)	-0.392 (0.214)†
<i>Disengaged vs Other Classes</i>		0.336 (0.190)†			-0.155 (0.182)		0.301 (0.198)
Doubt Decision							
<i>Engaged & Confident</i>	2.034 (0.158)***	<i>ref</i>	0.582 (0.114)***	0.478 (0.186)*	<i>ref</i>	-0.047 (0.169)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	2.378 (0.122)***	0.344 (0.216)	0.410 (0.094)***	-0.262 (0.165)	-0.740 (0.270)**	-0.311 (0.155)*	-0.264 (0.248)
<i>Engaged but Worried</i>	2.176 (0.113)***	0.142 (0.197)	0.424 (0.094)***	0.253 (0.152)†	-0.225 (0.244)	0.312 (0.163)†	0.359 (0.238)
<i>Disengaged & Worried</i>	2.435 (0.180)***	0.401 (0.237)†	0.408 (0.153)**	0.169 (0.227)	-0.309 (0.290)	0.132 (0.228)	0.178 (0.281)

<i>Disengaged vs Other</i>		-0.239 (0.200)			-0.012 (0.252)		-0.147 (0.252)
<i>Classes</i>							
Determination 1							
<i>Engaged & Confident</i>	4.901 (0.261)***	<i>ref</i>	0.617 (0.097)***	0.015 (0.378)	<i>ref</i>	-0.088 (0.342)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	4.789 (0.210)***	-0.112 (0.364)	0.580 (0.079)***	0.083 (0.300)	0.068 (0.524)	0.185 (0.287)	0.274 (0.485)
<i>Engaged but Worried</i>	4.824 (0.178)***	-0.077 (0.321)	0.740 (0.063)***	0.070 (0.242)	0.055 (0.455)	-0.027 (0.237)	0.061 (0.422)
<i>Disengaged & Worried</i>	4.224 (0.294)***	-0.677 (0.389)†	0.522 (0.134)***	0.864 (0.408)*	0.849 (0.550)	0.207 (0.426)	0.295 (0.540)
<i>Disengaged vs Other</i>		0.614 (0.326)†			-0.808 (0.454)†		-0.184 (0.467)
<i>Classes</i>							
Determination 2							
<i>Engaged & Confident</i>	5.408 (0.250)***	<i>ref</i>	0.331 (0.203)	-0.707 (0.416)†	<i>ref</i>	0.347 (0.338)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	4.854 (0.202)***	-0.554 (0.346)	0.414 (0.100)***	0.318 (0.296)	1.026 (0.554)†	-0.087 (0.293)	-0.434 (0.481)
<i>Engaged but Worried</i>	5.050 (0.217)***	-0.358 (0.335)	0.516 (0.108)***	-0.193 (0.308)	0.514 (0.527)	0.204 (0.306)	-0.142 (0.463)
<i>Disengaged & Worried</i>	4.939 (0.326)***	-0.469 (0.407)	0.340 (0.148)*	-0.278 (0.421)	0.430 (0.585)	-0.239 (0.403)	-0.585 (0.520)
<i>Disengaged vs Other</i>		0.165 (0.359)			0.084 (0.475)		0.393 (0.453)
<i>Classes</i>							
Determination 3							
<i>Engaged & Confident</i>	4.348 (0.247)***	<i>ref</i>	0.781 (0.088)***	-0.035 (0.381)	<i>ref</i>	-0.155 (0.370)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	4.598 (0.235)***	0.250 (0.372)	0.338 (0.099)**	-0.738 (0.343)*	-0.703 (0.556)	-0.278 (0.317)	-0.122 (0.529)
<i>Engaged but Worried</i>	4.655 (0.221)***	0.307 (0.337)	0.380 (0.073)***	0.254 (0.303)	0.289 (0.494)	-0.181 (0.292)	-0.025 (0.479)
<i>Disengaged & Worried</i>	4.185 (0.261)***	-0.163 (0.354)	0.604 (0.090)***	0.109 (0.394)	0.144 (0.541)	0.231 (0.373)	0.387 (0.518)
<i>Disengaged vs Other</i>		0.349 (0.300)			-0.281 (0.450)		-0.436 (0.426)
<i>Classes</i>							
Determination 4†							
<i>Engaged & Confident</i>	4.297 (0.185)***	<i>ref</i>	0.360 (0.082)***	0.601 (0.293)*	<i>ref</i>	0.567 (0.308)†	<i>ref</i>

<i>Relaxed & Nonchalant</i>	4.507 (0.218)***	0.210 (0.307)	0.326 (0.110)**	-0.419 (0.277)	-1.020 (0.440)*	-0.237 (0.287)	-0.804 (0.455)†
<i>Engaged but Worried</i>	4.734 (0.177)***	0.437 (0.260)†	0.261 (0.091)**	-0.212 (0.257)	-0.813 (0.396)*	-0.267 (0.266)	-0.834 (0.414)*
<i>Disengaged & Worried</i>	3.896 (0.273)***	-0.401 (0.326)	0.288 (0.143)*	0.112 (0.400)	-0.489 (0.489)	0.096 (0.375)	-0.471 (0.480)
<i>Disengaged vs Other</i>		0.617 (0.302)*			-0.122 (0.440)		-0.075 (0.420)
<i>Classes</i>							
Preference for Academic							
Career							
<i>Engaged & Confident</i>	0.986 (0.265)***	<i>ref</i>	0.957 (0.049)***	-0.411 (0.533)	<i>ref</i>	-0.589 (0.479)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	0.503 (0.387)	-0.483 (0.511)	0.758 (0.071)***	0.488 (0.503)	0.899 (0.795)	0.201 (0.504)	0.790 (0.755)
<i>Engaged but Worried</i>	0.575 (0.319)†	-0.411 (0.420)	0.954 (0.054)***	-0.080 (0.367)	0.331 (0.656)	-0.075 (0.408)	0.514 (0.638)
<i>Disengaged & Worried</i>	0.997 (0.428)*	0.011 (0.496)	0.809 (0.080)***	-0.109 (0.587)	0.302 (0.783)	0.299 (0.580)	0.888 (0.743)
<i>Disengaged vs Other</i>		-0.310 (0.480)			0.108 (0.660)		-0.453 (0.653)
<i>Classes</i>							
Grad School Overall							
Eval***							
<i>Engaged & Confident</i>	78.170 (2.659)***	<i>ref</i>		-2.479 (4.603)	<i>ref</i>	1.392 (4.349)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	69.826 (2.736)***	-8.344 (4.144)*		1.898 (3.818)	4.377 (6.491)	-2.180 (4.000)	-3.571 (6.411)
<i>Engaged but Worried</i>	63.786 (3.094)***	-14.384 (4.155)**		2.711 (4.225)	5.191 (6.363)	8.074 (4.073)*	6.682 (6.065)
<i>Disengaged & Worried</i>	52.360 (3.441)***	-25.810(4.312)***		3.368 (6.115)	5.847 (7.586)	-4.503 (5.635)	-5.895 (7.067)
<i>Disengaged vs Other</i>		18.234 (3.913)***			-2.658 (6.731)		6.932 (6.250)
<i>Classes</i>							
Grad School Fun***							
<i>Engaged & Confident</i>	71.053 (3.035)***	<i>ref</i>		-1.472 (5.647)	<i>ref</i>	1.683 (5.451)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	65.768 (3.245)***	-5.285 (4.834)		0.746 (4.869)	4.538 (7.893)	-5.446 (4.994)	-7.129 (8.029)
<i>Engaged but Worried</i>	60.143 (3.243)***	-10.909 (4.519)*		4.180 (4.502)	4.055 (7.197)	2.934 (4.786)	1.251 (7.368)
<i>Disengaged & Worried</i>	45.160 (4.412)***	-25.893(5.307)***		-2.266 (7.321)	-4.049 (10.038)	2.170 (7.197)	0.487 (8.942)

<i>Disengaged vs Other</i>		20.495 (4.895)***		3.417 (8.023)		-2.446 (7.947)	
<i>Classes</i>							
Institutional Trust***							
<i>Engaged & Confident</i>	6.316 (0.124)***	<i>ref</i>		-0.338 (0.284)	<i>ref</i>	0.033 (0.174)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	6.011 (0.108)***	-0.305 (0.179)†		0.059 (0.170)	0.397 (0.359)	-0.270 (0.195)	-0.303 (0.283)
<i>Engaged but Worried</i>	5.772 (0.109)***	-0.545 (0.168)**		0.000 (0.182)	0.338 (0.344)	0.218 (0.187)	0.184 (0.259)
<i>Disengaged & Worried</i>	5.374 (0.208)***	-0.942 (0.241)***		-0.126 (0.375)	0.211 (0.466)	-0.173 (0.350)	-0.206 (0.387)
<i>Disengaged vs Other</i>		0.659 (0.224)**			0.033 (0.403)		0.167 (0.375)
<i>Classes</i>							
Burnout***							
<i>Engaged & Confident</i>	2.279 (0.176)***	<i>ref</i>		0.311 (0.332)	<i>ref</i>	-0.201 (0.265)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	2.755 (0.228)***	0.476 (0.313)		-0.210 (0.310)	-0.521 (0.493)	0.183 (0.316)	0.384 (0.449)
<i>Engaged but Worried</i>	3.578 (0.212)***	1.299 (0.280)***		-0.374 (0.268)	-0.685 (0.433)	0.095 (0.321)	0.296 (0.423)
<i>Disengaged & Worried</i>	3.555 (0.259)***	1.276 (0.309)***		0.016 (0.412)	-0.295 (0.523)	0.006 (0.438)	0.207 (0.507)
<i>Disengaged vs Other</i>		-0.684 (0.293)*			-0.107 (0.457)		0.020 (0.485)
<i>Classes</i>							
Interpersonal Relations and Perceived Fit							
Perceived Social Support**							
<i>Engaged & Confident</i>	4.337 (0.134)***	<i>ref</i>	0.445 (0.133)**	-0.046 (0.190)	<i>ref</i>	-0.166 (0.203)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.936 (0.122)***	-0.401 (0.195)*	0.498 (0.095)***	0.283 (0.179)	0.330 (0.283)	0.269 (0.162)†	0.435 (0.281)
<i>Engaged but Worried</i>	4.275 (0.112)***	-0.062 (0.177)	0.525 (0.114)***	-0.128 (0.163)	-0.081 (0.255)	-0.150 (0.165)	0.015 (0.266)
<i>Disengaged & Worried</i>	3.573 (0.180)***	-0.764 (0.223)**	0.547 (0.099)***	0.253 (0.211)	0.299 (0.281)	0.302 (0.261)	0.468 (0.327)
<i>Disengaged vs Other</i>		0.610 (0.197)**			-0.217 (0.240)		-0.318 (0.286)
<i>Classes</i>							

Similarity to Colleagues**							
<i>Engaged & Confident</i>	64.560 (3.609)***	<i>ref</i>	10.803 (1.887)***	-6.707 (5.107)	<i>ref</i>	-2.600 (4.737)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	53.724 (3.596)***	-10.836 (5.549)†	10.784 (1.760)***	2.433 (4.958)	9.140 (7.752)	-2.627 (4.756)	-0.027 (7.301)
<i>Engaged but Worried</i>	54.549 (3.247)***	-10.011 (4.932)*	9.052 (1.266)***	5.357 (4.580)	12.065 (6.979)†	2.355 (4.606)	4.956 (6.708)
<i>Disengaged & Worried</i>	44.534 (4.941)***	-20.027 (6.072)**	1.528 (2.512)	-0.700 (6.618)	6.007 (8.277)	-4.871 (6.273)	-2.270 (7.783)
<i>Disengaged vs Other</i>		13.078 (5.454)*			1.061 (7.337)		3.913 (7.004)
Classes							
Isolation***							
<i>Engaged & Confident</i>	2.341 (0.121)***	<i>ref</i>		0.162 (0.193)	<i>ref</i>	-0.091 (0.192)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	2.746 (0.118)***	0.405 (0.184)*		-0.179 (0.168)	-0.376 (0.269)	-0.067 (0.154)	0.024 (0.266)
<i>Engaged but Worried</i>	2.723 (0.106)***	0.382 (0.163)*		0.138 (0.143)	-0.007 (0.238)	0.194 (0.162)	0.285 (0.255)
<i>Disengaged & Worried</i>	3.337 (0.141)***	0.996 (0.184)***		-0.229 (0.232)	-0.309 (0.335)	-0.403 (0.208)†	-0.312 (0.280)
<i>Disengaged vs Other</i>		-0.734 (0.159)***			0.269 (0.257)		0.415 (0.235)†
Classes							
Belonging*							
<i>Engaged & Confident</i>	5.521 (0.301)***	<i>ref</i>	0.567 (0.274)*	-0.322 (0.198)	<i>ref</i>	0.002 (0.181)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	5.319 (0.123)***	-0.201 (0.342)	0.430 (0.324)	0.043 (0.169)	0.365 (0.284)	-0.067 (0.152)	-0.068 (0.256)
<i>Engaged but Worried</i>	5.034 (0.111)***	-0.487 (0.323)	0.685 (0.113)***	0.089 (0.152)	0.410 (0.254)	0.166 (0.167)	0.164 (0.251)
<i>Disengaged & Worried</i>	4.011 (0.460)***	-1.510 (0.549)**	-0.248 (0.425)	0.156 (0.290)	0.477 (0.348)	-0.230 (0.324)	-0.231 (0.370)
<i>Disengaged vs Other</i>		1.281 (0.476)**			-0.219 (0.311)		0.263 (0.343)
Classes							
Belonging Uncertainty 1**							
<i>Engaged & Confident</i>	4.309 (0.290)***	<i>ref</i>	0.412 (0.106)***	0.378 (0.450)	<i>ref</i>	-0.317 (0.344)	<i>ref</i>

<i>Relaxed & Nonchalant</i>	4.629 (0.235)***	0.320 (0.401)	0.506 (0.134)***	0.232 (0.377)	-0.146 (0.636)	-0.009 (0.332)	0.308 (0.520)
<i>Engaged but Worried</i>	5.464 (0.237)***	1.154 (0.379)**	0.342 (0.093)***	-0.043 (0.267)	-0.421 (0.530)	-0.490 (0.307)	-0.173 (0.468)
<i>Disengaged & Worried</i>	5.097 (0.304)***	0.787 (0.416)†	0.304 (0.170)†	0.083 (0.400)	-0.295 (0.593)	0.294 (0.411)	0.612 (0.530)
<i>Disengaged vs Other</i>		-0.296 (0.346)			0.106 (0.460)		-0.566 (0.464)
<i>Classes</i>							
Belonging Uncertainty 2							
<i>Engaged & Confident</i>	5.733 (0.195)***	<i>ref</i>	0.284 (0.094)**	-0.021 (0.292)	<i>ref</i>	0.130 (0.251)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	5.634 (0.144)***	-0.099 (0.263)	0.587 (0.142)***	0.207 (0.229)	0.227 (0.403)	0.119 (0.226)	-0.011 (0.366)
<i>Engaged but Worried</i>	6.073 (0.152)***	0.340 (0.251)	0.120 (0.116)	0.014 (0.194)	0.034 (0.356)	-0.065 (0.224)	-0.196 (0.342)
<i>Disengaged & Worried</i>	5.460 (0.232)***	-0.273 (0.300)	0.596 (0.130)***	0.112 (0.320)	0.133 (0.428)	-0.181 (0.345)	-0.311 (0.422)
<i>Disengaged vs Other</i>		0.354 (0.255)			-0.045 (0.355)		0.242 (0.379)
<i>Classes</i>							
Belonging Uncertainty 3***							
<i>Engaged & Confident</i>	4.124 (0.278)***	<i>ref</i>	0.479 (0.139)**	-0.717 (0.424)†	<i>ref</i>	0.329 (0.358)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.137 (0.213)***	-0.987 (0.377)**	0.480 (0.133)***	0.219 (0.338)	0.935 (0.589)	0.540 (0.321)†	0.211 (0.521)
<i>Engaged but Worried</i>	3.016 (0.252)***	-1.108 (0.381)**	0.420 (0.101)***	-0.256 (0.286)	0.461 (0.519)	-0.104 (0.310)	-0.432 (0.481)
<i>Disengaged & Worried</i>	2.774 (0.210)***	-1.350 (0.345)***	0.365 (0.112)**	0.429 (0.362)	1.145 (0.549)*	-0.175 (0.347)	-0.504 (0.493)
<i>Disengaged vs Other</i>		0.652 (0.263)*			-0.680 (0.421)		0.430 (0.405)
<i>Classes</i>							
Self-Evaluations							
Self-efficacy							
<i>Engaged & Confident</i>	3.761 (0.394)***	<i>ref</i>	-0.082 (0.570)	-0.007 (0.125)	<i>ref</i>	-0.008 (0.114)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.171 (0.088)***	-0.589 (0.415)	0.133 (0.189)	-0.062 (0.125)	-0.055 (0.188)	-0.079 (0.120)	-0.071 (0.177)
<i>Engaged but Worried</i>	3.123 (0.084)***	-0.638 (0.405)	0.525 (0.072)***	0.070 (0.113)	0.077 (0.171)	0.075 (0.115)	0.083 (0.164)
<i>Disengaged & Worried</i>	2.901 (0.195)***	-0.860 (0.437)*	0.436 (0.203)*	0.151 (0.176)	0.158 (0.214)	-0.083 (0.214)	-0.075 (0.241)

<i>Disengaged vs Other</i>		0.450 (0.239)†			-0.150 (0.195)		0.079 (0.230)
<i>Classes</i>							
Self-esteem**							
<i>Engaged & Confident</i>	3.617 (0.173)***	<i>ref</i>	0.468 (0.123)***	0.024 (0.233)	<i>ref</i>	0.215 (0.231)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.412 (0.174)***	-0.205 (0.268)	0.196 (0.143)	0.119 (0.238)	0.094 (0.362)	0.070 (0.200)	-0.145 (0.330)
<i>Engaged but Worried</i>	3.148 (0.141)***	-0.469 (0.227)*	0.475 (0.072)***	0.042 (0.201)	0.017 (0.313)	-0.042 (0.204)	-0.256 (0.314)
<i>Disengaged & Worried</i>	2.461 (0.342)***	-1.155 (0.381)**	0.151 (0.191)	0.593 (0.372)	0.568 (0.435)	0.001 (0.310)	-0.214 (0.384)
<i>Disengaged vs Other</i>		0.931 (0.360)*			-0.531 (0.401)		0.080 (0.342)
<i>Classes</i>							
Imposter Syndrome							
<i>Engaged & Confident</i>	3.168 (0.157)***	<i>ref</i>	0.561 (0.126)***	0.152 (0.241)	<i>ref</i>	-0.022 (0.179)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.138 (0.162)***	-0.029 (0.246)	0.582 (0.179)**	0.013 (0.166)	-0.139 (0.320)	-0.043 (0.148)	-0.020 (0.253)
<i>Engaged but Worried</i>	3.509 (0.135)***	0.341 (0.209)	0.649 (0.104)***	-0.024 (0.127)	-0.176 (0.276)	-0.205 (0.140)	-0.182 (0.231)
<i>Disengaged & Worried</i>	3.344 (0.139)***	0.177 (0.208)	0.430 (0.114)***	-0.189 (0.179)	-0.341 (0.296)	0.143 (0.195)	0.165 (0.261)
<i>Disengaged vs Other</i>		-0.073 (0.167)			0.236 (0.209)		-0.233 (0.219)
<i>Classes</i>							
Sense of Accomplishment***							
<i>Engaged & Confident</i>	3.632 (0.206)***	<i>ref</i>		-0.022 (0.340)	<i>ref</i>	-0.454 (0.279)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	3.168 (0.166)***	-0.464 (0.288)		-0.083 (0.256)	-0.062 (0.462)	-0.110 (0.250)	0.344 (0.408)
<i>Engaged but Worried</i>	3.012 (0.187)***	-0.620 (0.284)*		-0.215 (0.249)	-0.194 (0.429)	0.186 (0.267)	0.639 (0.393)
<i>Disengaged & Worried</i>	2.186 (0.205)***	-1.445 (0.288)***		0.411 (0.307)	0.433 (0.453)	-0.617 (0.303)*	-0.163 (0.408)
<i>Disengaged vs Other</i>		1.084 (0.236)***			-0.518 (0.354)		0.491 (0.346)
<i>Classes</i>							
Mental Health							
Distress							

<i>Engaged & Confident</i>	6.627 (1.025)***	<i>ref</i>	0.695 (0.220)**	0.400 (1.118)	<i>ref</i>	-0.577 (0.984)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	7.218 (0.747)***	0.591 (1.364)	0.491 (0.243)*	-0.046 (1.034)	-0.446 (1.657)	-1.590 (0.823)†	-1.013 (1.390)
<i>Engaged but Worried</i>	8.242 (0.617)***	1.615 (1.212)	0.658 (0.105)***	-1.883 (0.847)*	-2.283 (1.425)	-0.795 (0.897)	-0.217 (1.354)
<i>Disengaged & Worried</i>	6.275 (1.169)***	-0.351 (1.543)	0.670 (0.155)***	1.860 (1.352)	1.460 (1.734)	0.219 (1.361)	0.796 (1.661)
<i>Disengaged vs Other</i>		1.086 (1.275)			-2.370 (1.500)		-1.206 (1.487)
<i>Classes</i>							
Social Identity Threat							
Stereotype Threat							
(Gender)							
<i>Engaged & Confident</i>	2.593 (0.229)***	<i>ref</i>	0.527 (0.145)***	-0.325 (0.266)	<i>ref</i>	-0.179 (0.253)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	2.775 (0.215)***	0.182 (0.340)	0.562 (0.164)**	-0.357 (0.220)	-0.032 (0.375)	-0.001 (0.230)	0.178 (0.370)
<i>Engaged but Worried</i>	3.008 (0.176)***	0.415 (0.293)	0.653 (0.061)***	0.037 (0.219)	0.362 (0.351)	-0.032 (0.245)	0.148 (0.359)
<i>Disengaged & Worried</i>	3.011 (0.193)***	0.418 (0.296)	0.642 (0.106)***	-0.127 (0.325)	0.198 (0.415)	0.102 (0.322)	0.282 (0.405)
<i>Disengaged vs Other</i>		-0.219 (0.231)			-0.088 (0.360)		-0.173 (0.360)
<i>Classes</i>							
Stereotype Threat							
(Race)†							
<i>Engaged & Confident</i>	2.162 (0.228)***	<i>ref</i>	0.265 (0.138)†	-0.295 (0.244)	<i>ref</i>	-0.447 (0.222)*	<i>ref</i>
<i>Relaxed & Nonchalant</i>	2.760 (0.161)***	0.598 (0.301)*	0.596 (0.089)***	-0.322 (0.232)	-0.027 (0.366)	-0.380 (0.230)†	0.066 (0.348)
<i>Engaged but Worried</i>	2.425 (0.158)***	0.263 (0.281)	0.803 (0.051)***	0.124 (0.207)	0.419 (0.324)	0.297 (0.225)	0.743 (0.321)*
<i>Disengaged & Worried</i>	2.853 (0.176)***	0.691 (0.285)*	0.694 (0.097)***	-0.421 (0.301)	-0.126 (0.383)	-0.166 (0.315)	0.281 (0.381)
<i>Disengaged vs Other</i>		-0.404 (0.209)†			0.257 (0.336)		-0.011 (0.349)
<i>Classes</i>							
Research Effort							
Research Hours							

<i>Engaged & Confident</i>	11.954 (2.580)***	<i>ref</i>	-0.126 (3.907)	<i>ref</i>	-5.241 (3.356)	<i>ref</i>
<i>Relaxed & Nonchalant</i>	8.592 (2.105)***	-3.362 (3.621)	-1.279 (3.137)	-1.152 (5.439)	4.677 (3.397)	9.918 (5.202)†
<i>Engaged but Worried</i>	9.408 (1.821)***	-2.546 (3.210)	2.747 (3.157)	2.873 (5.114)	3.342 (2.909)	8.583 (4.507)†
<i>Disengaged & Worried</i>	16.349 (3.052)***	4.395 (3.954)	-1.846 (5.837)	-1.720 (6.952)	-10.943 (3.989)**	-5.702 (5.140)
<i>Disengaged vs Other</i>		-6.364 (3.368)†		2.293 (6.295)		11.869 (4.501)**
<i>Classes</i>						

Note: This table provides the coefficients for each class, specifically the coefficients for the intercepts, the baseline predictor where applicable, and the two dummy-coded intervention variables (where control condition is the reference group). Class differences represent the difference between each class and the *Carefree and Confident* class; the last contrast tests the differences between the *Disengaged and Worried* class and the other three classes. Asterisks beside outcome variable name denote that the end-of-semester means (intercepts) are significantly different from each other overall. Parameter estimates are unstandardized. All outcomes reported are from the same time point, the end of the semester (end of week 4), except for institutional trust, measured at the end of week 3 (in the latter half of the semester).

*** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$

expected level of interest of a student in the *Engaged and Confident* class and in the control condition whose baseline level of interest is at the sample mean (see Table 8). The -0.877 in the next column to the right indicates the difference in end-of-semester interest between the *Disengaged and Worried* and *Engaged and Confident* classes, which is statistically significant. This indicates that among students who did not receive intervention, interest at the end of the semester was significantly lower in the *Disengaged and Worried* class than in the *Engaged and Confident* class, controlling for baseline levels of interest. The estimates for baseline slopes can be interpreted as the average change in the outcome for students in the indicated class and in the control condition when the baseline level increased by 1. For instance, 0.229 is how much end-of-semester interest can be expected to increase on average for students in the *Engaged and Confident* class and in the control condition when baseline interest increases by 1.

Class membership predicted multiple distal outcomes in control condition students, including attitudes towards graduate school and academia, interpersonal relations and perceived fit, and self-evaluations. It did not predict mental health, social identity threat (SIT), or research effort in control condition students.³ Regardless of threat profile, then, students who did not receive intervention and who started the semester at the average levels of these outcomes had similar levels of mental health symptoms and SIT across classes at the end of the first semester of graduate school; they also reported similar amounts of time spent on research across classes. We present results by class in order from highest to lowest threat.

³ At baseline, collapsed across conditions, class membership did predict mental health and SIT (see Table 7).

Disengaged and Worried class.

The *Disengaged and Worried* control condition students had overall negative attitudes towards graduate school and academia at (or near) semester's end.⁴ On average, they were the least interested in their studies and field, evaluated graduate school the most negatively, trusted their institution the least, and reported the highest levels of burnout. In terms of interpersonal relations and perceived fit, they felt the least supported, the least similar to their colleagues, the most isolated, the lowest sense of belonging, and relatively high belonging uncertainty. These students also evaluated themselves the most negatively, with the lowest self-esteem and the lowest sense of accomplishment. While the omnibus test for self-efficacy was not significant, individual comparisons showed that *Disengaged and Worried* students reported significantly lower self-efficacy than the *Engaged and Confident* students. Overall, these students are clearly struggling the most both before and at the end of the first semester of graduate school.

Engaged but Worried class.

The *Engaged but Worried* students had mixed attitudes towards graduate school and academia. On the one hand, they maintained relatively high interest in their studies and field at the end of the semester. On the other hand, their graduate school evaluations and institutional trust were lower than the average for students from the two lower threat classes, and their reported burnout was higher. These students also reported mixed outcomes in terms of interpersonal relations and perceived fit. On the one hand, they were similar to the lowest threat class in their (high) perceptions of social support. On the other hand, they felt more isolated and less similar to their colleagues than the lowest threat students and felt relatively uncertain about their sense of belonging. In terms of self-evaluation, *Engaged but Worried* students reported

⁴ All outcomes in Table 8, the ones described in this section, are from the end of the semester, which is end of week 4 (EOW4) except for institutional trust, measured at end of week 3 (EOW3; in the latter half of the semester).

relatively low levels of self-esteem and sense of accomplishment at semester's end. Overall, these students were faring better at the end of the semester than the *Disengaged and Worried* students but worse than the students from the other two classes, which is the same pattern found before the semester.

Relaxed and Nonchalant class.

The *Relaxed and Nonchalant* students, like the *Engaged but Worried* students, had mixed attitudes towards graduate school and academia; however, their attitudes were overall more positive than those of the *Engaged but Worried* class. These students expressed similar interest in their field and research to the *Engaged and Confident* and *Engaged but Worried* students; however, there was a nonsignificant trend for the *Relaxed and Nonchalant* students to have lower interest than the lowest threat class, and this pattern aligns with initial class descriptions. The *Relaxed and Nonchalant* students also evaluated graduate school at similarly high levels as the students from the lowest threat class (although not quite as high for the overall evaluation) and felt similarly high and low levels of institutional trust and burnout, respectively. These findings align with the initial description of the class (e.g., as not maximally engaged but also not facing much threat) and emphasize its distinctions from the *Engaged but Worried* class. Regarding interpersonal relations and perceived fit, the *Relaxed and Nonchalant* students tended to perceive less social support than the *Engaged and Confident* students but felt about as similar to their colleagues as the lowest threat class students (with a trend towards the *Relaxed and Nonchalant* students perceiving less similarity)—this, too, emphasizes their distinction from the *Engaged but Worried* students. The *Relaxed and Nonchalant* students also felt about as isolated as the *Engaged but Worried* students, which is more so than the lowest threat class, and had somewhat more uncertainty about belonging than the lowest threat class. Additionally, these students had

relatively positive self-evaluations, with their self-esteem and sense of accomplishment levels being about as high as the lowest threat class students. Overall, these students fared relatively well at the end of the semester; they were most similar to the lowest threat class, although still not faring quite as well as those students, particularly in terms of interpersonal relations.

Engaged and Confident class.

The *Engaged and Confident* control condition students, as expected, had the best end-of-semester outcomes. These students had the most positive attitudes towards graduate school and academia at the end of the semester, with the highest interest, most positive evaluations of graduate school, highest institutional trust, and lowest burnout. They also felt the best about interpersonal relations and their perceived fit, with the most perceived social support, highest perceived similarity to colleagues, least isolation, highest belonging, and least belonging uncertainty. Unsurprisingly given their “confident” descriptor, they also had the most positive self-evaluations, with the highest self-esteem and greatest sense of accomplishment. Overall, the *Engaged and Confident* students who did not receive intervention lived up to their class name and faced the least psychological threat, further supporting their status as low risk students.

Intervention Effects (Phase 4)

For Phase 4, we tested whether any effects of the intervention condition differed by latent class (Table 8). Specifically, we used the BCH procedure as described in the Phase 3 results, that is, including class membership in a model with two dummy-coded intervention variables as moderators of the latent classes predicting distal outcomes and baseline covariates where available. For each outcome, we examined the effect of each intervention condition compared to the control by class. We also made comparisons of the efficacy of each intervention among classes, with the *Engaged and Confident* class being compared to each of the others, and the

Disengaged and Worried class being compared to the other three (the same comparisons we used for the intercepts). Given that we conducted many tests of comparison, significance tests of individual outcomes should be interpreted cautiously and de-emphasized relative to patterns of results that are likely more reliable than any given individual significant effect.

We described how to interpret the first three columns of Table 8 for the previous section on control condition students; we now explain how to interpret the rest of Table 8 so that we can understand effects of intervention. The coefficients for control vs belonging and control vs affirmation represent the difference in a distal outcome for students in the intervention compared to the control condition for a student in the designated class who began the semester at the mean level of that outcome (where applicable). For instance, 0.478 is the amount by which end-of-semester doubt is higher in the belonging condition than the control condition for a student in the *Engaged and Confident* class who began the semester at the mean level of doubt. Hence, for a typical student in the relatively unthreatened *Engaged and Confident* class, end-of-semester doubt is (surprisingly) higher in the belonging than control condition. The tests of comparison (“class differences”) compare the condition slopes across classes. For instance, the belonging intervention relative to the control condition is associated with an average decrease in doubt of 0.740 in the *Relaxed and Nonchalant* class compared to the *Engaged and Confident* class. Hence, the belonging intervention works better (i.e., at decreasing doubt) in the *Relaxed and Nonchalant* class than the *Engaged and Confident* class.

There were relatively few effects of either intervention overall. We present results by class in order from highest to lowest threat.

Disengaged and Worried class.

For the *Disengaged and Worried* students, the three outcomes significantly affected were research effort, sense of accomplishment, and the first determination item (“Even if I could hardly maintain my social life, I would still continue graduate school”). These highest risk students reported spending on average 11 fewer hours on research in the preceding week if they were in the affirmation condition than if they were in the control condition. Compared to the *Disengaged and Worried* students, students from the other three classes spent on average almost 12 hours more on research in that week due to being in the affirmation instead of the control condition. Moreover, the *Disengaged and Worried* students felt significantly less accomplished at semester’s end if they were in the affirmation relative to the control condition. We did not expect either intervention to decrease research effort or sense of accomplishment for any of the classes. Instead, we expected the highest threat students to benefit most from intervention, which did not happen. However, we did see that the belonging intervention affected one outcome for this class in a more positive way: these highest threat students reported significantly greater determination to persist in graduate school if they were in the belonging rather than the control condition even controlling for baseline determination. It is important to note, though, that this was one of several items regarding determination and there was no effect for this class on the other items. Moreover, the effects of intervention on sense of accomplishment and determination here were not accompanied by significant tests of comparisons of intervention efficacy between classes (i.e., the intervention did not affect this class significantly differently than it affected other classes on these two outcomes, which did not see significant intervention effects); this calls for a more cautious interpretation of these results.

Engaged but Worried class.

For the *Engaged but Worried* students, intervention affected end of semester distress and the overall evaluation of graduate school. The belonging intervention significantly benefitted the mental health of the *Engaged but Worried* students even controlling for baseline distress—among students in this class, distress at the end of the semester was 1.883 points lower in the belonging than in the control condition. Moreover, the affirmation intervention significantly improved these students' overall evaluations of graduate school compared to the control condition. However, neither of these effects were found to be significantly different than the nonsignificant intervention effects on the lowest threat class on these two outcomes. We expected the *Engaged but Worried* students to benefit more from intervention than the two lower threat classes, but there is little evidence for that.

Relaxed and Nonchalant class.

For the *Relaxed and Nonchalant* students, the affirmation intervention significantly decreased doubt in the decision to pursue a Ph.D., controlling for baseline levels, which is a beneficial effect of intervention for this class. Additionally, these students reported significantly more determination for item 3 (“I would quit studying if it became apparent that there were no jobs after graduation”) if they were in the belonging relative to the control condition, controlling for baseline determination. Again, neither of these effects were found to be significantly different than the nonsignificant intervention effects on the lowest threat class for these two outcomes. We did not expect many intervention effects for these students given their relatively low levels of psychological threat.

Engaged and Confident class.

For the *Engaged and Confident* students, we did not expect many intervention effects given their low levels of psychological threat. However, we did find that the affirmation intervention affected race-based stereotype threat and the belonging intervention affected doubt in the decision to pursue a Ph.D. and willingness to study one's field despite wasted time at the end of the semester (determination item 4). Students in this class in the affirmation condition reported 0.447 less race-based stereotype threat at semester's end, controlling for baseline levels, than their control condition counterparts. The class comparisons revealed that the *Engaged and Confident* students benefitted significantly more from affirmation intervention on this outcome than the *Engaged but Worried* students. On the other hand, controlling for baseline doubt, *Engaged and Confident* students in the belonging condition reported 0.478 more doubt about their decision to pursue a Ph.D. at the end of the semester than their control condition counterparts—this is an unintended, adverse effect of intervention. The class comparisons revealed that the belonging intervention hurt the *Engaged and Confident* students significantly more than only the *Relaxed and Nonchalant* students. Hence, the belonging intervention worked better in the *Relaxed and Nonchalant* class than the *Engaged and Confident* class, primarily due to a backfiring effect in the *Engaged and Confident* class. However, the belonging intervention also had a positive effect on the lowest threat students in that it increased their willingness to study their field even if they end up wasting time on certain topics (determination item 4). Students in this class in the belonging condition reported 0.601 more determination than their control condition counterparts controlling for baseline determination. The class comparisons showed that the students in this lowest threat class benefitted significantly more from the belonging intervention on this outcome than the students in the *Relaxed and Nonchalant* and

Engaged but Worried classes. We did not expect the lowest threat class to benefit more from intervention than other, higher threat classes.

CHAPTER 4. DISCUSSION

In this study, we used latent class analysis (LCA) to identify incoming doctoral students who might be at heightened risk of attrition. Results revealed four subgroups of Ph.D. students who varied in their experiences of psychological threat. We explored how these subgroups, or classes, differed by demographic characteristics as well as psychological experiences. Results validated our interpretations of the classes and gave us a fuller picture of the students belonging to each. Moreover, we found that the classes differed across end-of-semester outcomes in meaningful ways, with higher threat classes faring worse than their lower threat counterparts. Finally, we tested the efficacy of two interventions by class and found few effects overall with an ambiguous pattern—results that have multiple potential explanations that require further research to disambiguate. Overall, we suggest the utility of the LCA approach to describing and identifying at-risk students, with some caveats regarding the potential extensions and implications of the approach.

We first summarize the highlights of the results for each class in order by threat level starting with the lowest threat class, as a primary goal of the current study was to better understand doctoral student experiences and we found ample evidence for these experiences differing by threat profile, or class.

The *Engaged and Confident* class was characterized as the lowest threat class, with high interest in and engagement with graduate school and an overall perception of preparedness and confidence with minimal threat. At baseline, the students in this lowest threat class overall seemed the most academically prepared. Low-SES students were notably underrepresented in this class, which was not unexpected given low-SES students face psychological threat that other students may not, and the *Engaged and Confident* students were the least threatened.

International students were also slightly underrepresented—we did not have clear expectations about how international students would do in terms of threat, but this underrepresentation in the lowest threat class points to the potential for international students to find graduate school more threatening than the best performing non-international students.

Furthermore, the *Engaged and Confident* students tended to have the most positive outcomes at baseline and the end of the semester (as a reminder, for the distal outcomes, we focused on control condition students to ascertain what experiences are like when intervention is not involved). A typical student in this class had positive attitudes towards graduate school and academia, so might feel very trusting of their institution and greatly enjoy their experience as a graduate student. As these students tended to perceive their interpersonal lives positively as well, a typical student might feel especially supported and like they fit in well with their department. Typical students in this class evaluated themselves positively, so might feel quite accomplished after a semester of work. Intervention or no intervention, these students seem to be at low risk of attrition (or other adverse outcomes) given their all-around low-threat psychological state. Given their low threat, past research would suggest that the interventions, which have tended to work best among students experiencing more threat, would not be very effective for students in this class. In fact, among the few significant effects of intervention found for this class, we found mixed results. Compared to control, the affirmation intervention decreased their race-based stereotype threat, and the belonging intervention increased their doubt about pursuing a Ph.D. while also increasing one element of their determination to finish their Ph.D. Considering the lack of consistency and overall lack of significant effects, we suggest further research is needed to determine if such low threat students can benefit from intervention.

The *Relaxed and Nonchalant* class was characterized as neither greatly worried about nor greatly engaged in graduate school, with responses tending towards the middle rather than the scale extremes. This class was slightly overrepresented in terms of international students. We did not have strong expectations regarding how international students would be most likely to be categorized, but there are various explanations for this finding. It could be that Americans simply tend to respond at scale extremes more often than people from certain other countries (e.g., Lee et al., 2002; Yeh et al., 1998). There may also be a confounding variable that would better elucidate the distribution of international students across classes; we expand on this idea when we discuss the *Disengaged and Worried* class.

The *Relaxed and Nonchalant* students were in the middle for most outcomes but did have generally more positive than negative outcomes. For instance, they were well-prepared for graduate school with the highest proportion of students entering with master's degrees. They also had positive self-evaluations and attitudes towards graduate school even by the end of the first semester, although their interpersonal outcomes (e.g., isolation) were not as positive—it would be interesting to examine if this latter result is directly related to the overrepresentation of international students in this class. For instance, it could be that international students, although well-prepared for the academic aspects of doctoral programs, face more difficulties integrating socially than non-international students given cultural differences. Overall, the *Relaxed and Nonchalant* class was more similar to the *Engaged and Confident* class than any other class on most outcomes. As such, we would expect them to be at relatively low risk of attrition due to psychological threat. Of the two intervention effects for this class, both were benefits. The affirmation intervention, relative to control, decreased their doubt about deciding to pursue a Ph.D. and reduced their distress; the belonging intervention increased their determination to

persist in their studies even if job prospects look poor. We did not expect intervention to particularly help this class given their relative lack of psychological threat, and it will be interesting to examine if this becomes a pattern in future research.

The *Engaged but Worried* class was characterized by high interest in and engagement with graduate school but also high threat, with concerns about fitting in and being judged. Although this class demonstrated high preparation for graduate school and high engagement in it initially, their relatively high negative affect and uncertainty leads us to predict that they might be at slightly heightened risk of attrition. At baseline, the *Engaged but Worried* students were fairly well-prepared academically in terms of prior research experience but, true to their initial description, reported relatively high levels of neuroticism. In general, they evaluated themselves rather negatively, which likely does not reflect their actual competencies given their prior research experience and acceptance into their doctoral programs. Students in this class, despite the relatively high levels of SIT and impostor syndrome, were determined to get their degrees and maintained high interest in their research and field even by the end of the semester. Unfortunately, alongside that, these students reported high levels of burnout and uncertainty about belonging at semester's end.

In terms of demographics, the *Engaged but Worried* class had the most women, most queer students, the fewest international students, and although non-significant, the most under-represented minority (URM) students. Given this class is characterized by more SIT than others, particularly on the gender axis, we found this unsurprising. We expected female, queer, first-generation, and URM students to face more SIT than other students and for them to be classified more often into a higher threat class. While not the highest threat class, the *Engaged but Worried* class is relatively high risk, and in future research, we would want to take a closer look at long-

term outcomes for this class. If outcomes continue to trend negatively or worsen over time, this class could be an important target for intervention. Moreover, given their relatively high threat, we might expect members of this class to be particularly benefitted by social belonging and values-affirmation interventions. However, there were only intervention effects on two outcomes for this class—intervention benefitted this class' mental health (i.e., reduced their distress) and improved their overall evaluations of graduate school. In the future, particularly given intervention benefitted the mental health of the *Relaxed and Nonchalant* students as well, we may want to explore other outcomes related to mental health—perhaps interventions in this population are more likely to affect mental health compared to other kinds of psychological outcomes. Given the prevalence of mental health issues among graduate students (Evans et al., 2018), this is an important avenue for future research.

The *Disengaged and Worried* class was characterized as the highest risk class, with the highest levels of threat on most of the indicators. At baseline, the *Disengaged and Worried* students were the least academically prepared and, as suggested by their moniker, high in neuroticism and low in academic self-control. They tended to be in fields with relatively few women despite the class being split evenly by gender, which may be one reason they experience gender-based SIT more than other students. As we expected, these students had the most negative outcomes in nearly every domain examined. A typical student in this class was relatively uninterested in their work, not enjoying graduate school, and feeling distrustful and highly burned out at the end of the semester. These students felt relatively different from their peers, unsupported, and isolated. It was unsurprising given their baseline negative affect that they continued to evaluate themselves negatively after the semester, reporting relatively low self-

esteem and low feelings of accomplishment. This class was clearly the highest threat and as such, we expect they are at the highest risk of attrition.

We found it notable that the *Disengaged and Worried* class was not overrepresented in terms of high-risk social categories given its characterization as the highest threat class. This suggests the importance of looking beyond social group membership when designing ways to mitigate threat and reduce attrition for STEM Ph.D. students; indeed, the approach we took with LCA instead of using social categories as proxies for threat seems to provide significant nuance. International students were slightly overrepresented in this highest threat class—perhaps a next step to disambiguate the threat facing international students would be to break down the category further (e.g., international students who speak English as a first language compared to those who do not). For instance, it may be the case that international students who do not speak English as fluently as others encounter SIT. It might also be the case that only English proficiency matters rather than international student status.

We might suspect, given the class' characterization as highly threatened, that *Disengaged and Worried* students would particularly benefit from intervention, but that was not evident in the outcomes examined. Instead, the direction of the two affirmation intervention effects for this class was unexpected, with students in the affirmation condition spending 11 fewer hours on research and feeling significantly less accomplished than their control condition counterparts in the week examined at the end of the first semester. Although there was also a beneficial effect of the belonging intervention, it was only found for one of the four examined determination items (i.e., determination to persist in graduate school even if unable to maintain a social life). It is possible that the observed negative intervention effects were random given the many significance tests conducted, but it is also possible that the values-affirmation intervention truly caused the

decrease in the amount of time *Disengaged and Worried* students spent on research and their lowered sense of accomplishment. For instance, considering the effect on research effort, perhaps an affirmed student already inclined to disengage who is reminded of her value as an individual may respond to continued struggles in graduate school by disengaging further, choosing to spend more time with loved ones or invested in affirming hobbies instead of research. Alternatively, we can imagine a more pessimistic scenario. The affirmation intervention is meant to bolster a student's self-integrity by highlighting that there are other domains in life besides school to derive self-worth from, but could this message fall flat for a student who has low self-worth generally and/or may not derive much self-worth from other domains? It is also clear that number of hours spent on research in a single week is not the only way to assess research effort—in the future, we can examine more comprehensive outcomes in this domain to determine whether there is a clear pattern of results.

While we do find appreciable differences among classes initially and at the end of the first semester of graduate school, we did not find much evidence for effects of either intervention or moderation of those effects by class. One intervention was not clearly superior to another, and overall, the effects and trends did not reveal a consistent pattern. There were some ironic effects when interventions backfired for specific classes on certain outcomes, but we do not feel comfortable concluding that any of the observed significant effects were due to more than chance given the high number of significance tests conducted. Given the mixed results and past research indicating social psychological interventions can backfire (Bayly & Bumpus, 2019; Brady et al., 2016; Cohen et al., 2006; Dee, 2015; Liu & Huang, 2019; Miyake et al., 2010), it seems wise to at least proceed with caution and continue exploring why interventions may work for some people but not others and in some contexts but not others. One idea to consider in research going

forward is the contextual affordances doctoral programs are likely to proffer threatened students (see Walton & Yeager, 2020). If doctoral programs are so demanding of one's time and energy, for instance, that students find it difficult to consider and tend to other aspects of their life that might validate their worth as individuals, then the values affirmation intervention may not work. Similarly, if doctoral programs have a culture wherein social support is difficult to obtain for certain kinds of students, the belonging intervention may not work. These possibilities could be particularly relevant for universities such as the ones examined in the current study, which are relatively well-regarded for their research and may have a more competitive atmosphere.

Walton and Yeager (2020) are not the only researchers to offer ideas about the inconsistency of social psychological interventions like the belonging and values-affirmation ones used in this study. Complementing their idea of contextual affordances, Easterbrook and Hadden (2021) proposed a way to consider how and when to apply interventions using the *Identities in Context* model, which emphasizes how interactions between social identities and local contexts can cause SIT and feelings of identity incompatibility differentially across educational settings. In the present research, at this point it is unclear exactly how the context influences the efficacy of the interventions. As alluded to previously, it may be that Ph.D. students do not react to these interventions the same way similarly positioned students at lower levels of education would. Moreover, beyond the prestige of the university, the context of the local campus, field of study, department, or cohort may be influential. We did check to see that our selected latent class model was robust in that the same classes emerged even when we used subsamples (i.e., we ran individual LCAs for cohort 1, cohort 2, PSU, and Stanford/Columbia—for Stanford and Columbia we combined subsamples given constraints on sample size). However, even if campus and cohort—and possibly other characteristics of the context like field

of study or department—do not affect the class construction meaningfully, they may still alter intervention efficacy. Thus, future work can examine contextual characteristics in a more nuanced way. For instance, we may want to use campus as a covariate in analyses using condition, so that we can examine if intervention effects manifest differently based on campus.

The limited scope of outcomes examined here, primarily at the end of the first semester, may also be a reason we did not find many effects. It is imperative to take a closer look at the efficacy of the interventions, which we can do by examining more distal and different kinds of outcomes (e.g., GPA, attrition), as well as more outcomes within the domains already identified (e.g., mental health outcomes besides psychological distress). Furthermore, we can more closely scrutinize intervention efficacy by utilizing the smartphone data to explore nuances about any potential patterns of fluctuation in psychological experiences over different windows of time. For instance, we can look at fluctuations in sense of belonging over time as a function of condition, which may give a better sense of belonging uncertainty than the 3-item measure of belonging uncertainty we used (analyzed as 3 separate items because of the poor reliability).

Other limitations to the current study include constraints of the analytic technique used. For instance, LCA is a technique that is difficult to generalize, as the construction of classes is very sample-specific. While we conducted robustness analyses by cohort campus, and they lend support for the model selected, we cannot conclude that these same classes would emerge in other groups of students. Moreover, we had to trichotomize some of the indicator variables in sample-specific ways (i.e., using the distribution of values in the current sample rather than theoretically determined cutoffs due to issues with low cell size). Although making adjustments to how the variables were trichotomized was necessary due to cell size issues, it could exacerbate the issue of the selected model being overfit to the current sample because what is considered

low in this sample, for instance, may be considered medium in another. Of course, the categorization of continuous variables can be considered a limitation in and of itself given loss of information; however, we did this because latent profile analysis (the complement to LCA that only differs in its use of continuous rather than categorical indicators) failed to work in the current study (i.e., the models were not stable), which we expected given the strict assumptions that need to be met for latent profile analysis.

Although we believe we demonstrated the utility of LCA in this study through the clear identification of classes and the ability to use those classes to predict outcomes relevant to attrition risk, we do have reservations about the extensions and implications of the approach. Regarding extensions, we would be interested in using class membership in different kinds of analyses, such as time-varying effect models, to give us a more nuanced picture of intervention effects in particular; however, there is not currently a way to do this without taking an approach contraindicated by the literature (i.e., a classify-analyze approach). The classify-analyze approach is not recommended due to its failure to account appropriately for measurement error (Nylund-Gibson et al., 2019). The BCH procedure we used is the ideal approach to take into account the fact that assignment of individuals to classes is uncertain (Bakk & Vermunt, 2016; Dziak et al., 2016), but this procedure is still relatively nascent and requires further development of theory and software before it can be used for more complicated analyses. Furthermore, the implications of the class structure identified in the current sample are unclear given aforementioned limitations on generalizability using LCA. Ideally, we want future researchers to be able to use the approach we took to risk identification in a way that allows them to tailor interventions to students who most need them. One idea would be to measure the same indicators we used here in the population of interest (i.e., incoming doctoral students), apply the same

approach to trichotomizing those indicators (i.e., based on theoretically based cutoffs where possible and sample-specific cutoffs where not), and then, if similar classes emerge, target interventions to students who are unlikely to be negatively impacted by intervention. Before this, however, it is important to gather more information about intervention effects, as the evidence found here is unclear. To do this, as highlighted previously, we can consider context and moderators and expand the scope of outcomes and time points examined.

Future work could also explore the question of whether class membership changes over time as a function of intervention condition. Do initially higher risk students transition to lower risk classes if they received intervention? Researchers can address this kind of question by using latent transition analysis (LTA), an extension of LCA. One challenge with using LTA for this kind of study, however, is the large changes in sample size from baseline to distal time points. More generally, too, issues of participant attrition from the study or likeliness to continue into the longitudinal part of the study at all are important to consider when drawing conclusions about class membership over time.

CHAPTER 5. CONCLUSION

In this study, we used LCA to identify doctoral students who may be at heightened risk of attrition. We found that incoming Ph.D. students can be categorized into four different psychological threat profiles, or classes, that portend different levels of risk for negative psychological outcomes at the end of the first semester. We believe such psychological experiences may have implications for later attrition, although future work is necessary to examine this question more directly. Furthermore, we explored the potential moderation of social belonging and values-affirmation interventions by threat profile but did not find conclusive results about the efficacy of the interventions overall by the end of the first semester or about moderation of any effects by threat profile. Future work should expand the investigation into potential intervention effects by considering contextual moderators, different timelines for effects to manifest (e.g., outcomes after one or more years), and different indicators of intervention efficacy (e.g., more performance-based outcomes). Overall, we consider LCA a useful descriptive and predictive tool, but require further data before we can determine the efficacy of using LCA to tailor interventions.

REFERENCES

- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health psychology, 19*(6), 586-592. <https://doi.org/10.1037/0278-6133.19.6.586>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723. doi:10.1109/ TAC.1974.1100705
- Anderman, L. H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *The Journal of Experimental Education, 72*(1), 5-22. <https://doi.org/10.1080/00220970309600877>
- Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White Men Can't Do Math: Necessary and Sufficient Factors in Stereotype Threat. *Journal of Experimental Social Psychology, 35*(1), 29–46. <https://doi.org/10.1006/jesp.1998.1371>
- Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim & C. Stangor (Eds.), *Prejudice* (pp. 83-103). Academic Press. <https://doi.org/10.1016/B978-012679130-3/50039-9>
- Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling: A Multidisciplinary Journal, 23*(1), 20–31. <https://doi.org/10.1080/10705511.2014.955104>
- Bayly, B. L., & Bumpus, M. F. (2019). An exploration of engagement and effectiveness of an online values affirmation. *Educational Research and Evaluation, 25*(5-6), 248-269. <https://doi.org/10.1080/13803611.2020.1717542>
- Binning, K., & Browman, A. (2019). *Theoretical, Ethical, and Policy Considerations for*

Conducting Social-Psychological Interventions to Close Educational Achievement Gaps.

<https://psyarxiv.com/w68b9/download?format=pdf>

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. <https://doi.org/10.1111/j.1467-8624.2007.00995.x>

Blascovich, J., & Mendes, W. B. (2001). Challenge and threat appraisals. In *Feeling and thinking: The role of affect in social contagion* (pp. 59–82). Cambridge.
https://books.google.com/books?hl=en&lr=&id=PSiU9wsJ13QC&oi=fnd&pg=PA59&dq=Blascovich+J,+Mendes+WB.+Challenge+and+threat+appraisals:+The+role+of+affective+cues&ots=ekJr2PuzRJ&sig=LD_6uS3qiaBMp7_L_p_tu_t8nsU

Borman, G. D., Rozek, C. S., Pyne, J., & Hanselman, P. (2019). Reappraising academic and social adversity improves middle school students' academic achievement, behavior, and well-being. *Proceedings of the National Academy of Sciences of the United States of America, 116*(33), 16286–16291. <https://doi.org/10.1073/pnas.1820317116>

Bower-Brown, S., Zadeh, S., & Jadva, V. (2021). Binary-trans, non-binary and gender-questioning adolescents' experiences in UK schools. *Journal of LGBT Youth, 1-19*.
<https://doi.org/10.1080/19361653.2021.1873215>

Bowman, N. A., Hill, P. L., Denson, N., & Bronkema, R. (2015). Keep on truckin' or stay the course? Exploring grit dimensions as differential predictors of educational achievement, satisfaction, and intentions. *Social Psychological and Personality Science, 6*(6), 639–645.
<https://doi.org/10.1177/1948550615574300>

Brady, S. T., Reeves, S. L., Garcia, J., Purdie-Vaughns, V., Cook, J. E., Taborisky-Barba, S.,

- Tomasetti, S., Davis, E. M., & Cohen, G. L. (2016). The psychology of the affirmed learner: Spontaneous self-affirmation in the face of stress. *Journal of Educational Psychology, 108*(3), 353-373. <https://doi.org/10.1037/edu0000091>
- Bratter, J. L., Rowley, K. J., & Chukhray, I. (2016). Does a self-affirmation intervention reduce stereotype threat in black and Hispanic high schools?. *Race and Social Problems, 8*(4), 340-356. <https://doi.org/10.1007/s12552-016-9187-4>
- Bravata, D. M., Watts, S. A., Keefer, A. L., Madhusudhan, D. K., Taylor, K. T., Clark, D. M., Nelson, R. S., Cokley, K. O., & Hagg, H. K. (2019). Prevalence, predictors, and treatment of impostor syndrome: a systematic review. *Journal of general internal medicine, 1*-24. <https://doi.org/10.1007/s11606-019-05364-1>
- Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing Inequality in Academic Success for Incoming College Students: A Randomized Trial of Growth Mindset and Belonging Interventions. *Journal of Research on Educational Effectiveness, 11*(3), 317–338. <https://doi.org/10.1080/19345747.2018.1429037>
- Choe, N. H., Borrego, M. J., Martins, L. L., Patrick, A. D., & Seepersad, C. C. (2017, June). A quantitative pilot study of engineering graduate student identity. In *2017 ASEE Annual Conference & Exposition, Columbus, OH*.
- Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools, 42*(2), 197–205. <https://doi.org/https://doi.org/10.1002/pits.20048>
- Clance, P. R. (1985). *The imposter phenomenon*. Peachtree, Atlanta.
- Cohen, G. L., & Garcia, J. (2005). "I am us": negative stereotypes as collective threats. *Journal of Personality and Social Psychology, 89*(4), 566-582. <https://doi.org/10.1037/0022->

3514.89.4.566

- Cohen, G. L., & Garcia, J. (2008). Identity, belonging, and achievement: A model, interventions, implications. *Current Directions in Psychological Science*, *17*(6), 365-369.
<https://doi.org/10.1111/j.1467-8721.2008.00607.x>
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307–1310.
<https://doi.org/10.1126/science.1128317>
- Cohen, G. L., Purdie-Vaughns, V., & Garcia, J. (2012). An Identity Threat Perspective on Intervention. In *Stereotype Threat: Theory, Process, and Application*.
<https://doi.org/10.1093/acprof:oso/9780199732449.003.0018>
- Cohen, G. L., & Sherman, D. K. (2014). The Psychology of Change: Self-Affirmation and Social Psychological Intervention. *Annual Review of Psychology*. <https://doi.org/10.1146/annurev-psych-010213-115137>
- Collins, L. M., & Lanza, S. T. (2010). Repeated measures latent class analysis and latent transition analysis. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken: Wiley, 181-224.
- Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology*, *102*(3), 479-496.
<https://doi.org/10.1037/a0026312>
- Davey, G. C., Hampton, J., Farrell, J., & Davidson, S. (1992). Some characteristics of worrying: Evidence for worrying and anxiety as separate constructs. *Personality and Individual Differences*, *13*(2), 133-147. [https://doi.org/10.1016/0191-8869\(92\)90036-O](https://doi.org/10.1016/0191-8869(92)90036-O)

- Dee, T. S. (2015). Social identity and achievement gaps: Evidence from an affirmation intervention. *Journal of Research on Educational Effectiveness*, 8(2), 149-168.
<https://doi.org/10.1080/19345747.2014.906009>
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of Personality Assessment*, 91(2), 166-174.
<https://doi.org/10.1080/00223890802634290>
- Dziak, J., Bray, B., Zhang, J., Zhang, M., & Lanza, S. (2016). Comparing the performance of improved classify-analyze approaches for distal outcomes in latent profile analysis. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 12(4), 107–116. <https://doi.org/https://doi.org/10.1027/1614-2241/a000114>
- Easterbrook, M. J., & Hadden, I. R. (2021). Tackling educational inequalities with social psychology: Identities, contexts, and interventions. *Social Issues and Policy Review*, 15(1), 180-236. <https://doi.org/10.1111/sipr.12070>
- Ehrenberg, R. G., Jakubson, G. H., Groen, J. A., So, E., & Price, J. (2007). Inside the Black Box of Doctoral Education: What Program Characteristics Influence Doctoral Students' Attrition and Graduation Probabilities? *Journals.Sagepub.Com*, 29(2), 134–150.
<https://doi.org/10.3102/0162373707301707>
- Erez, A., & Judge, T. A. (2001). Relationship of core self-evaluations to goal setting, motivation, and performance. *Journal of Applied Psychology*, 86(6), 1270-1279.
<https://doi.org/10.1037/0021-9010.86.6.1270>

- Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T., & Vanderford, N. L. (2018). Evidence for a mental health crisis in graduate education. *Nature Biotechnology*, *36*(3), 282-284.
<https://doi.org/10.1038/nbt.4089>
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142.
<https://doi.org/10.3102/00346543059002117>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. In *Review of Educational Research* (Vol. 74, Issue 1, pp. 59–109). American Educational Research Association.
<https://doi.org/10.3102/00346543074001059>
- Gerber, A. S., & Green, D. P. (2017). Field experiments on voter mobilization: An overview of a burgeoning literature. *Handbook of Economic Field Experiments*, *1*, 395-438.
<https://doi.org/10.1016/bs.hefe.2016.09.002>
- Gnambs, T., & Hanfstingl, B. (2016). The decline of academic motivation during adolescence: An accelerated longitudinal cohort analysis on the effect of psychological need satisfaction. *Educational Psychology*, *36*(9), 1691-1705.
<https://doi.org/10.1080/01443410.2015.1113236>
- Good, C., Rattan, A., & Dweck, C. S. (2012). Why Do Women Opt Out? Sense of Belonging and Women's Representation in Mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. <https://doi.org/10.1037/a0026659>
- Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, *61*(2), 231.
<https://academic.oup.com/biomet/article/61/2/215/268006>
- Goyer, J. P., Garcia, J., Purdie-Vaughns, V., Binning, K. R., Cook, J. E., Reeves, S. L., Apfel,

- N., Taborsky-Barba, S., Sherman, D. K., & Cohen, G. L. (2017). Self-affirmation facilitates minority middle schoolers' progress along college trajectories. *Proceedings of the National Academy of Sciences*, *114*(29), 7594-7599. <https://doi.org/10.1073/pnas.1617923114>
- Groen, J. A., Jakubson, G. H., Ehrenberg, R. G., Condie, S., & Liu, A. Y. (2008). Program design and student outcomes in graduate education. *Economics of Education Review*, *27*(2), 111–124. <https://doi.org/https://doi.org/10.1016/j.econedurev.2006.09.010>
- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in Context: School Moderation of the Impact of Social Identity Threat on Racial/Ethnic Achievement Gaps. *Sociology of Education*, *87*(2), 106–124. <https://doi.org/10.1177/0038040714525970>
- Hanselman, P., Rozek, C. S., Grigg, J., & Borman, G. D. (2017). New evidence on self-affirmation effects and theorized sources of heterogeneity from large-scale replications. *Journal of Educational Psychology*, *109*(3), 405-424. <https://doi.org/10.1037/edu0000141>
- Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology*, *106*(2), 375-389. <https://doi.org/10.1037/a0034679>
- Hay, C., & Forrest, W. (2006). The development of self-control: Examining self-control theory's stability thesis. *Criminology*, *44*(4), 739-774. <https://doi.org/10.1111/j.1745-9125.2006.00062.x>
- Hayes, L., Zinner, L., Wise, J., & Carton, J. (2019). Effects of a self-affirmation intervention on grades in middle school and first-year college students. *Journal of Articles in Support of the Null Hypothesis*, *16*(1), 57-80.

- Hughes, M. E., Waite, L. J., Hawkey, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on Aging, 26*(6), 655-672. <https://doi.org/10.1177/0164027504268574>
- Inzlicht, M., Tullett, A. M., Legault, L., & Kang, S. K. (2011). Lingering effects: Stereotype threat hurts more than you think. *Social Issues and Policy Review, 5*(1), 227–256. <https://doi.org/10.1111/j.1751-2409.2011.01031.x>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). *Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues*. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (p. 114–158). The Guilford Press.
- Jones, B. D., Ruff, C., & Paretti, M. C. (2013). The impact of engineering identification and stereotypes on undergraduate women's achievement and persistence in engineering. *Social Psychology of Education, 16*(3), 471-493. <https://doi.org/10.1007/s11218-013-9222-x>
- Jones, E. O., & Huey, S. J. Jr. (2020). Affirmation and majority students: Does affirmation impair academic performance in white males?. *Basic and Applied Social Psychology, 42*(3), 176-191. <https://doi.org/10.1080/01973533.2020.1732389>
- Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?. *Journal of Personality and Social Psychology, 83*(3), 693-710. <https://doi.org/10.1037/0022-3514.83.3.693>
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L. T., Walters, E. E., & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological*

- Medicine*, 32(6), 959-976. <https://doi.org/10.1017/S0033291702006074>
- Kinias, Z., & Sim, J. (2016). Facilitating women's success in business: Interrupting the process of stereotype threat through affirmation of personal values. *Journal of Applied Psychology*, 101(11), 1585–1597. <https://doi.org/10.1037/apl0000139>
- Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science*, 355(6322), 251-252. <https://doi.org/10.1126/science.aag2063>
- La Guardia, J., & Ryan, R. (2002). What adolescents need. *Academic motivation of adolescents*, 2, 193-218.
- Lazarsfeld, P., & Henry, N. (1968). *Latent structure analysis*. Houghton Mifflin.
- Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a Likert scale. *Research in Nursing & Health*, 25(4), 295-306.
<https://doi.org/10.1002/nur.10041>
- Lent, R., Brown, S., & Larkin, K. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology*, 31(3), 356–362.
<https://doi.org/10.1037/0022-0167.31.3.356>
- Liu, C. H., & Huang, P. S. (2019). Beneficial effects of self-affirmation on motivation and performance reduced in students hungry for others' approval. *Contemporary Educational Psychology*, 56, 1-13. <https://doi.org/10.1016/j.cedpsych.2018.10.008>
- Liu, X., & Flick, R. (2019). The relationship among psychological need satisfaction, class engagement, and academic performance: Evidence from China. *Journal of Education for Business*, 94(6), 408-417. <https://doi.org/10.1080/08832323.2018.1541855>
- Lovitts, B. E., & Nelson, C. (2000). The hidden crisis in graduate education: Attrition from Ph.D. programs. *Academe*, 86(6), 44–50.

- Magnus, K., Diener, E., Fujita, F., & Pavot, W. (1993). Extraversion and neuroticism as predictors of objective life events: a longitudinal analysis. *Journal of Personality and Social Psychology, 65*(5), 1046. <https://doi.org/10.1037/0022-3514.65.5.1046>
- Major, B., & Schmader, T. (2018). *Stigma, social identity threat, and health*. In B. Major, J. F. Dovidio, & B. G. Link (Eds.), *Oxford library of psychology. The Oxford handbook of stigma, discrimination, and health* (pp. 85–103). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190243470.013.3>
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior, 2*(2), 99-113. <https://doi.org/10.1002/job.4030020205>
- Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *Oxford library of psychology. The Oxford Handbook of Quantitative Methods. Statistical analysis* (pp. 551–611). Oxford University Press. <https://psycnet.apa.org/record/2013-01010-025>
- Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008), 1234-1237. <https://doi.org/10.1126/science.1195996>
- Moore, R. T., & Schnakenberg, K. (2012). blockTools: Blocking, Assignment, and Diagnosing Interference in Randomized Experiments. R package version 0.5-7.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide (7th ed)*. Los Angeles, CA: Authors.
- National Research Council. (2011). *A Data-Based Assessment of Research-Doctorate Programs in the United States (with CD)*. National Academies Press. <https://doi.org/10.17226/12994>
- National Science Foundation. (2018). Survey of Earned Doctorates. Retrieved October 1, 2020,

from <http://www.nsf.gov/statistics/srvydoctorates/>

Nicpon, M. F., Huser, L., Blanks, E. H., Sollenberger, S., Befort, C., & Kurpius, S. E. R. (2006).

The relationship of loneliness and social support with college freshmen's academic performance and persistence. *Journal of College Student Retention: Research, Theory & Practice*, 8(3), 345-358. <https://doi.org/10.2190/A465-356M-7652-783R>

Nieuwenhuis, M., Manstead, A. S., & Easterbrook, M. J. (2019). Accounting for unequal access

to higher education: The role of social identity factors. *Group Processes & Intergroup Relations*, 22(3), 371-389. <https://doi.org/10.1177/1368430219829824>

Nieuwhof, M. G., ThJ ten Cate, O., Oosterveld, P., & Soethout, M. B. (2004). Measuring

strength of motivation for medical school. *Medical Education Online*, 9(1), 4355.

<https://doi.org/10.3402/meo.v9i.4355>

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in

latent class analysis and growth mixture modeling: A Monte Carlo simulation study.

Structural Equation Modeling: A Multidisciplinary Journal, 14(4), 535-569.

<https://doi.org/10.1080/10705510701575396>

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from latent classes: A

demonstration of different approaches to include distal outcomes in mixture

models. *Structural Equation Modeling: A Multidisciplinary Journal*, 26(6), 967-985.

<https://doi.org/10.1080/10705511.2019.1590146>

Oberski, D. (2016). Mixture models: Latent profile and latent class analysis. In J. Robertson &

M. Kaptein (Eds.), *Modern statistical methods for HCI. Human-computer interaction*

series. (pp. 275-287). Springer, Cham. https://doi.org/10.1007/978-3-319-26633-6_12

Patrick, A., Borrego, M., & Prybutok, A. (2018) Predicting Persistence in Engineering through

- an Engineering Identity Scale. *International Journal of Engineering Education*, 34(2(A)), 351-363. <https://doi.org/10.15781/T2ZC7SB9J>
- Protzko, J., & Aronson, J. (2016). Context moderates affirmation effects on the ethnic achievement gap. *Social Psychological and Personality Science*, 7(6), 500-507. <https://doi.org/10.1177/1948550616646426>
- Ramage, K. (1990). Motivational factors and persistence in foreign language study. *Language Learning*, 40(2), 189-219. <https://doi.org/10.1111/j.1467-1770.1990.tb01333.x>
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 27(2), 151-161. <https://doi.org/10.1177/0146167201272002>
- Rockinson-Szapkiw, A. J., Spaulding, L. S., & Spaulding, M. T. (2016). Identifying significant integration and institutional factors that predict online doctoral persistence. *The Internet and Higher Education*, 31, 101-112. <https://doi.org/10.1016/j.iheduc.2016.07.003>
- Rosenthal, L., Levy, S. R., London, B., Lobel, M., & Bazile, C. (2013). In pursuit of the MD: The impact of role models, identity compatibility, and belonging among undergraduate women. *Sex Roles*, 68(7-8), 464-473. <https://doi.org/10.1007/s11199-012-0257-9>
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115(2), 336-356. <https://doi.org/10.1037/0033-295X.115.2.336>
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461-464. <https://doi.org/10.1214/aos/1176344136>
- Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate

- analysis. *Psychometrika*, 52, 333–343. doi:10.1007/BF02294360
- Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety research*, 4(1), 27-41. <https://doi.org/10.1080/08917779108248762>
- Sellers, R. M., Rowley, S. A., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional inventory of black identity: A preliminary investigation of reliability and construct validity. *Journal of Personality and Social Psychology*, 73(4), 805-815. <https://doi.org/10.1037/0022-3514.73.4.805>
- Settles, I. H. (2004). When multiple identities interfere: The role of identity centrality. *Personality and Social Psychology Bulletin*, 30(4), 487-500. <https://doi.org/10.1177/0146167203261885>
- Settles, I. H., Jellison, W. A., & Pratt-Hyatt, J. S. (2009). Identification with multiple social groups: The moderating role of identity change over time among women-scientists. *Journal of Research in Personality*, 43(5), 856-867. <https://doi.org/10.1016/j.jrp.2009.04.005>
- Sheldon, K. M., & Hilpert, J. C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion*, 36(4), 439-451. <https://doi.org/10.1007/s11031-012-9279-4>
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., Tomassetti, S., David Nussbaum, A., & Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/a0031495>
- Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering

against identity threat. *Personality and Social Psychology Bulletin*, 39(5), 663-676.

<https://doi.org/10.1177/0146167213480816>

Shrauger, J. S., & Sorman, P. B. (1977). Self-evaluations, initial success and failure, and improvement as determinants of persistence. *Journal of Consulting and Clinical Psychology*, 45(5), 784-795.

<https://doi.org/10.1037/0022-006X.45.5.784>

Shryock, K. J., & Froyd, J. E. (2017, June). Board# 135:“I AM an engineer!” Three Scales used in Measuring Identification of Engineering as First-Year Students. In *2017 ASEE Annual Conference & Exposition*.

<https://doi.org/10.18260/1-2--27740>

Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67, 415–437.

<https://doi.org/10.1146/annurev-psych-073115-103235>

Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, 45(4), 249-257.

<https://doi.org/10.1080/00050067.2010.482109>

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self.

In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302).

Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60229-4](https://doi.org/10.1016/S0065-2601(08)60229-4)

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629.

<https://doi.org/https://doi.org/10.1037/0003-066X.52.6.613>

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797-811.

<https://doi.org/10.1037/0022-3514.69.5.797>

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The

- psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379–440. [https://doi.org/10.1016/S0065-2601\(02\)80009-0](https://doi.org/10.1016/S0065-2601(02)80009-0)
- Strayhorn, T. L. (2014). What role does grit play in the academic success of black male collegians at predominantly white institutions? *Journal of African American Studies*, 18(1), 1–10. <https://doi.org/10.1007/s12111-012-9243-0>
- Tao, K. W., & Gloria, A. M. (2019). Should I stay or should I go? The role of impostorism in STEM persistence. *Psychology of Women Quarterly*, 43(2), 151-164. <https://doi.org/10.1177/0361684318802333>
- Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, 23(1), 73-82. <https://doi.org/10.1177/0963721413512856>
- Walton, G. M., & Brady, S. T. (2020). The social-belonging intervention. In *Handbook of Wise Interventions: How Social-Psychological Insights Can Help Solve Problems*. Guilford Press. <https://studentexperienceproject.org/wp-content/uploads/2019/11/WaltonBrady2019.pdf>
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/0022-3514.92.1.82>
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*. <https://doi.org/10.1126/science.1198364>
- Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering. *Journal of Educational Psychology*. <https://doi.org/10.1037/a0037461>

- Walton, G. M., Murphy, M. C., Logel, C., & Yeager, D. S., & The College Transition Collaborative. (2017). *The social-belonging intervention: A guide for use and customization*.
- Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science*, 20(9), 1132–1139. <https://doi.org/10.1111/j.1467-9280.2009.02417.x>
- Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, 125(5), 617–655. <https://doi.org/10.1037/rev0000115>
- Walton, G. M., & Yeager, D. S. (2020). Seed and Soil: Psychological Affordances in Contexts Help to Explain Where Wise Interventions Succeed or Fail. *Current Directions in Psychological Science*, 29(3), 219–226. <https://doi.org/10.1177/0963721420904453>
- Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, 49(1), 287-293. <https://doi.org/10.1037/0022-3514.49.1.287>
- Woodcock, A., & Bairaktarova, D. (2015). Gender-biased self-evaluations of first-year engineering students. *Journal of Women and Minorities in Science and Engineering*, 21(3), 255-269. <http://doi.org/10.1615/JWomenMinorScienEng.2015013125>
- Woodcock, A., Hernandez, P., & Estrada, M. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology*, 103(4), 635–646. <https://doi.org/10.1037/a0029120>
- Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D’Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: a self-transcendent purpose for learning

fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*(4), 559–580. [https://doi.org/ 10.1037/a0037637](https://doi.org/10.1037/a0037637)

Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., Hessert, W. T., Williams, M. E., & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, *143*(2), 804-824. [https://doi.org/ 10.1037/a0033906](https://doi.org/10.1037/a0033906)

Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., Kamentz, D., Ritter, G., Duckworth, A. L., Urstein, R., Gomez, E. M., Markus, H. R., Cohen, G. L., & Dweck, C. S. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences*, *113*(24), E3341-E3348. <https://doi.org/10.1073/pnas.1524360113>

Yeh, L. L., Kim, K. O., Chompreeda, P., Rimkeeree, H., Yau, N. J. N., & Lundahl, D. S. (1998). Comparison in use of the 9-point hedonic scale between Americans, Chinese, Koreans, and Thai. *Food Quality and Preference*, *9*(6), 413-419. [https://doi.org/10.1016/S0950-3293\(98\)00028-7](https://doi.org/10.1016/S0950-3293(98)00028-7)