Boston College
Lynch School of Education
& Human Development

Department of
Counseling, Applied Developmental, & Educational Psychology

Program of
Applied Developmental & Educational Psychology

QUANTIFYING THE EFFECTS OF THE 'AT-RISK' LABEL: EXPLORING THE DEFICIT-ORIENTED LABELING EXPERIENCES OF LOW-INCOME, FIRST-GENERATION COLLEGE STUDENTS OF COLOR.

Dissertation

by

Shenira A. Perez

David B. Miele, Chair

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

July 2019

ABSTRACT

Quantifying the effects of the 'at-risk' label: Exploring the deficit-oriented labeling experiences of low-income, first-generation college students of color.

Shenira A. Perez, Author

David B. Miele, Chair

Institutional efforts to address attainment gaps in higher education have traditionally centered on deficit-oriented discourses that frame Black and Hispanic students, low-income students, and first-generation college students as 'at-risk' and 'underprepared'. Given the extensive amount of evidence documenting the adverse consequences of labeling and stigmatization, relying on negative descriptors to characterize marginalized students may be detrimental to their motivation and persistence in college—and may inadvertently exacerbate disparities in graduation rates between these students and students from more privileged backgrounds.

A total of three online studies were conducted for this dissertation, which explores the deficit-oriented labeling experiences of low-income, first-generation Black and Hispanic college students (LIFG; $n = 256$) and their non-low-income, continuing-generation White peers (NLIFG; $n = 317$). In Studies 1 and 2, participants were asked to respond to a series of prompts designed to examine the extent to which deficit-oriented labels were applied to them, the contexts in which this occurred, and the motivational and affective consequences they experienced as a result. In Studies 1 and 3, hypothetical scenarios were used to probe participants' interpretations of both deficit-oriented and alternative labels (i.e., *first-generation*

student), as well as the perceived consequences of being characterized by these descriptors. Study 3 also explored potential stereotype threat effects that might result from being characterized by a deficit-oriented label. Specifically, participants were randomly assigned to recall an experience in which they were labeled as an *at-risk* or *first-year student*, and then asked to complete an analytical task. Students' *academic mindsets*, *stereotype vulnerability*, and *racial identity beliefs* were also explored as potential moderators for within and between-group differences in Studies 2 and 3.

Results showed that relative to NLIFG students, LIFG students were significantly more likely to report being labeled by deficit-oriented descriptors. The frequency of these labeling experiences was also significantly associated with negative academic self-perceptions, sense of belonging, and affect, for both LIFG and NLIFG students. Across both sample groups, participants generally indicated that these labels were most often communicated to them by instructors and advisors. Responses to the hypothetical scenarios indicated that LIFG students were more nuanced in their interpretations of different labels, but there were no sample group differences in the extent to which participants expected these labeling experiences to negatively affect hypothetical students. There was no evidence of stereotype threat effects on subsequent performance, but this result may have been due to limitations associated with the manipulation task. Lastly, there was evidence to suggest that endorsing stronger academic growth mindsets may mitigate the negative effects of these stigmatizing experiences. The implications of these findings and recommendations for future work are also discussed.

ACKNOWLEDGEMENTS

Pursuing a Ph.D. at Boston College has been an incredibly rewarding adventure that I will be forever grateful for having had the privilege to experience. For this, I have many people to thank. First and foremost, my dissertation chair and advisor David Miele, whose high expectations pushed me to grow, both as a researcher and writer, and whose flexibility and encouragement allowed me to pursue research interests that were personally meaningful to me. His mentorship has been invaluable and I feel very fortunate for having had the opportunity to work with him. My dissertation committee, Heather Rowan-Kenyon and Marina Vasilyeva, for their guidance throughout this process. Larry Ludlow, an amazing professor and mentor who helped me overcome my fear of statistics head-on by inviting me to co-author a statistical paper with him, and who made that process an incredible learning opportunity. Ines Maturana Sendoya, whose passion for educational equity inspired the topic of this dissertation and influenced my career trajectory, as well as my undergraduate mentors Asia Eaton and Ronald Fisher, both of whom I credit for motivating me to pursue a doctoral degree.

Last—but most certainly not least—I am incredibly thankful for my family's unwavering support, encouragement, and unconditional love. My mom, whose intellectual curiosity and insight inspire me every day. My dad, whose work ethic and determination to succeed I hope to emulate in my own career. My sister, who is my best friend, an amazing mother, the funniest person I know, and an extremely influential force in my life. My brother-in-law (and the second funniest person I know), who not only puts up with my constant presence in his home, but actually welcomes it. My dog Mackie, who, quite frankly, deserves an honorary Ph.D. And finally, the loves of my life—Shashi and Bub—who motivate me to be the best version of me. This dissertation is dedicated to these extraordinary humans I am blessed to call my family.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

CHAPTER 1: INTRODUCTION

Over the past two decades, Black and Hispanic students have been enrolling at 4-year colleges and universities at an increasing rate, such that in 2014, these students comprised approximately 38 percent of the freshmen class (NCES, 2016). However, a recent report from the National Student Clearinghouse Research Center (NSCRC; 2017) indicated that given the current retention rates, over 40 percent of the Black and Hispanic students currently attending college are likely to drop out prior to receiving their degree, which is markedly higher than the 27 percent drop-out rate for White students. Moreover, Black and Hispanic students are also disproportionately more like to be first-generation college students and/or come from low-income backgrounds—both of which are characteristics that are also associated with disproportionately lower retention and graduation rates (NSCRC, 2017). Therefore, despite enrolling in college at rates that are comparable to their White peers, Black and Hispanic students—particularly those who are first-generation and/or from low-income backgrounds—are still considerably less likely to obtain their degree (Cahalan, Perna, Yamashita, Ruiz, & Franklin, 2017). Together, these data indicate that despite advancements in expanding access to higher education to students who have been historically marginalized in these contexts, the gaps in attainment continue to persist.

Several scholars have argued that the enduring nature of these and other achievement gaps are—at least in part—a product of a deficit ideology (Valencia, 1997; 2010). Deficit ideology refers to any framework for explaining disparities in educational achievement and attainment that solely attributes responsibility for these poor outcomes to students the themselves and their families (Marger, 1996; Pearl 1991; Placier, 1996). That is, as opposed to addressing the structural factors that breed inequities in achievement and attainment (Castro, 2014;

Valencia, 1997), deficit ideology characterizes underrepresented students as *at-risk*, *underprepared*, and *disadvantaged*, and primarily relies on educational 'interventions' to remediate underachievement.

Labeling students with a deficit-oriented label, such as 'at-risk' or 'underprepared,' is problematic because, rather than situating underachievement and disparities in attainment within the broader context of an inequitable educational system, these labels imply that the students themselves are at the center of the problem and that intervention programs are designed to address *their* inadequacies (Castro, 2014; Pearl, 1991). However, despite the negative message conveyed by these deficit-oriented labels, they are commonly used in the educational literature and in practical contexts (Castro, 2014; Gray, 2013; Valencia, 1997; 2010). For instance, there is prior research documenting the use of deficit-oriented labels to characterize low-income Black and Hispanic students and first-generation college students participating in educational intervention programs (Castro, 2014). We also know that these labels are associated with negative stereotypes about the academic competence and potential of the students that bear them (Castro, 2014; Gray, 2013; Steinhauer, 2017). For instance, Gray (2013) documented her observations of 'at-risk' students being described as "challenging", ''having a lack of maturity', and having "a checkered history" (p. 1247). Based on the manner in which these students were described to her, Gray concluded that "…these were the students expected to 'wash out of the university'… [they] were not framed as students whose experiences and interests would enrich the University and succeed in college" (p. 1247).

The extensive amount of literature documenting the negative effects of labeling and other stigmatizing experiences in academic environments suggest that the pervasive use of deficit-oriented labels within higher education may have important implications for students' outcomes

in these contexts, particularly for college students from groups that have been historically marginalized within higher education ('marginalized college students', for short; i.e., first-generation college students; low-income students; Black and Hispanic students; Castro, 2014; Pearl, 1991; Valencia, 1997; 2010). For example, we know that students labeled by descriptors that negatively represent their intellectual capacity—such as those designated as *learning disabled*—are subject to prejudiced perceptions of their competence and/or treatment by teachers that can negatively influence their own perceptions of their academic ability (e.g., Graham, 1984; Schifer, 2013; McKown & Weinstein, 2008). Other findings indicate that marginalized college students are more likely to experience diminished feelings of belonging in college, as a result of internalizing the negative academic stereotypes associated with their group membership (Walton & Cohen, 2007; Winograd & Rust, 2014). These students are also more likely than their non-stigmatized peers, to experience the acute effects of stereotype threat on their performance in achievement contexts, and the long-term consequences of chronic stereotype threat on their self-perceptions—which can serve to further exacerbate their sense of belonging in college (Aronson & Steele, 2005). Given the importance of high academic self-perceptions and strong feelings of belonging for students' motivation, persistence, and ultimately, achievement in college, at the very least, students' stigmatizing experiences are likely indirectly associated with lower levels of college retention and completion (e.g., Hausmann, Ye, Schofield, & Woods, 2009; Morrow & Ackermann, 2012; O'Keeffe, 2013).

Despite the seriousness of these implications, questions surrounding the use of deficit-oriented labels within higher education have not received the empirical attention they warrant. The lack of research on this topic is surprising, especially given the considerable amount of criticism these labeling practices have received from both practitioners and researchers alike

(e.g., Pelligrini, 1991; Pearl, 1991; Valencia, 2010). However, this work has been largely theoretical or anecdotal in nature, and not based on actual data. As a result, there is much that we do not understand about students' experiences of being characterized by these labels and the potential consequences of these experiences—which is the focus of the research conducted for this dissertation.

For instance, although we have evidence that these labels are used by staff and administrators to characterize the low-income students or color and/or first-generation college students that participate in their academic support programs, the extent to which these labels are communicated *directly* to students in these contexts—as well as others—remains unclear (Castro, 2014). It is also unclear which types of students might be disproportionately likely to have a deficit-oriented label applied to them. However, given that Black and Hispanic students are overrepresented among low-income college students and first-generation college students, both of which are populations associated with achievement disparities as well, there is reason to believe that these students would also be overrepresented among the 'at-risk' student population. This dissertation also examined the extent to which low-income, first-generation Black and Hispanic college students are characterized by deficit-oriented labels, relative to their non-low-income, continuing generation White peers, and focused on identifying the contexts in which these experiences would be most likely to occur. These are important questions because they can help identify the students and contexts at the center of this issue.

Identifying the students and contexts at the center of this issue would also provide valuable direction to investigations of the potential consequences associated with these experiences. Although these consequences have also yet to be explored, findings from existing work on labeling and stigmatization provide a basis from which to begin speculating about these

consequences. For instance, given that other types of academically stigmatizing experiences in achievement contexts have been shown to adversely influence students' academic self-perceptions and sense of belonging in college, it is possible that this would also apply to students who endure deficit-oriented labeling experiences. There is also reason to believe that the extent of these effects would depend on (a) the extent to which a student endures deficit-oriented labeling experiences, (b) the extent to which they interpret deficit-oriented labels in a negative manner, and (c) whether the extent to which the student experience other types of academically stigmatizing experiences.

For marginalized college students, being characterized by a deficit-oriented descriptor might augment the salience of negative academic stereotypes and trigger experiences of stereotype threat in settings where their academic performance is being evaluated (e.g., during exams). In turn, this could also adversely influence their academic self-perceptions and their sense of belonging in college. Alternatively, students may interpret these negative evaluations as being a product of prejudiced beliefs, which could increase their vigilance for prejudice and foster mistrust in that context. This scenario could be problematic if these experiences were to occur in the context of an academic support program, because a sense of distrust could potentially drive students to discount useful feedback and information that might otherwise be beneficial to them. The research conducted for this dissertation also explored these possibilities by investigating questions related to the potential consequences associated with students' deficit-oriented experiences.

Finally, it is unknown if other factors—in addition to students' background characteristics—might moderate the extent to which students experience negative consequences as a result of these experiences. However, there is some evidence to suggest that students'

academic mindsets (i.e., their beliefs about the malleability of intelligence), may be able to mitigate the negative effects of academic stigmatization. For instance, encouraging African-American students to endorse a growth mindset (i.e., view intelligence and academic ability as being malleable with hard work and effort) has been shown to protect them from the negative effects of stereotype threat on their identification with academics and their actual achievement (Aronson, Fried, & Good, 2002).

There is also evidence to suggest that students' *racial identity beliefs* may be able to mitigate the effects of stigmatizing school environments. Several studies have found that students from racial and ethnic minority groups who identify strongly with their background, seem to be shielded from the negative effects of diminished sense of belonging in college on students' self-worth, as well as the consequences of experiencing discrimination on students' academic achievement (Butler-Barnes et al., 2018; Chavous et al., 2003; 2008). What is unclear, is the extent to which these findings might apply to students' deficit-oriented labeling experiences. That said, the research conducted for this dissertation also addressed questions surrounding the role of students' academic mindsets and their racial identity beliefs in influencing their interpretations of deficit-oriented labels, and their perceptions of the consequence they might experience as a result of being labeled by these types of descriptors.

Objectives

The research conducted for this dissertation had several individual objectives—which in combination—were developed to enhance our overall understanding of students' deficit-oriented labeling experiences.

I. Contrast the extent to which marginalized college students are characterized by deficit-oriented labels, relative to their peers from more privileged backgrounds.

II. Identify the contexts in which these experiences are most likely to occur.

III. Identify and examine the potential consequences associated with students' deficit-oriented labeling experiences.

IV. Examine the extent to which students' academic mindsets and racial identity beliefs moderate their interpretations of deficit-oriented labels and their perceptions of the types of consequences they would be likely to experiences as a result of being labeled by these types of descriptors.

CHAPTER 2: LITERATURE REVIEW

The research that informed this dissertation was drawn from various areas of the broader sociological, psychological, and educational literatures. This literature review provides an overview of existing research both directly and indirectly related to college students' deficit-oriented labeling experiences and focuses on the implications of this work with respect to the research questions addressed in this dissertation.

The work discussed in this chapter is organized into three parts. In *Part I*, I introduce the concept of *educational deficit-thinking* and discuss its history and philosophical evolution, as well as the manner with which this ideology manifests within institutions of higher education. I end this section by explaining the origins of particular *deficit-oriented labels* (e.g., *at-risk*; *underprepared*, *disadvantaged*) and discuss the contexts in which these labels are typically used to characterize students in higher education. In *Part II*, I present a summary of the broader labeling literature as it relates to the consequences of stigmatization both in and out of educational contexts. Throughout this section, I use Labeling Theory as a framework for explaining the mechanisms driving these effects. I end this section by discussing the potential role of students' beliefs— specifically their *academic mindsets* and *racial identity beliefs*—in mitigating the effects of potentially stigmatizing experiences. Finally, in *Part III*, I discuss the implications of the findings presented in this literature review with respect to the potential effects of characterizing college students by deficit-oriented descriptors, with a focus on the marginalized students for whom these experiences may be especially stigmatizing. I conclude this chapter by setting the stage for the research that was conducted for this dissertation.

Part I

The Origins & Evolution of Educational Deficit-Thinking

Educational deficit-thinking (or *deficit-framing*) is the application of a deficit ideology to explain differences in student achievement and educational attainment (Castro, 2014, Marger, 1997; Pearl, 1991; Placier, 1996; Smit, 2011). In *The Evolution of Deficit Thinking*, Valencia (1997) describes this framework as 'an endogenous theory' used to explain the underachievement of low-income Black and Hispanic students, as the onus of failure is placed on the student. He explains that the deficit model posits that these students fail in school as a result of *internal* deficits that manifest as intellectual, linguistic, motivational, and behavioral limitations. At various points in history, these deficits have been attributed to genetic, cultural, and/or socioeconomic factors (Valencia 1997; 2010).

For instance, prior to the Civil Rights Movement, the deficit model was most closely aligned with genetic inferiority theories (e.g., eugenics) and explained disparities in student achievement as a product of inherent, biological differences between races (Pearl, 1991), with Black and Mexican American students viewed as intellectually inferior to White students. The influence of this perspective on the segregation laws that dominated educational policy for over a century is indisputable (Valencia, 2012). Pro-segregationists argued that co-mingling of races in schools would be detrimental to the achievement of White students, and that Black and Mexican American students would be best served through 'practical' curricula they could understand (Walters, 2001). However, with the passing of *Brown v. The Board of Education* in 1954—which effectively banned racial segregation in public schools—the eugenic underpinnings of deficit ideology became increasingly unpopular. In fact, in 1964, as he was preparing to publish his analysis of race relations, Thomas Pettigrew found only three American researchers who were

willing to support genetic explanations for racial differences in IQ scores (Pettigrew, 1964; Pearl, 1991).

As the popularity of genetic theories of intelligence continued to wane, they were replaced by ideas drawn from psychological theories of cultural deprivation (Pearl, 1991; Valencia, 1997). In this evolved iteration of the deficit model, underachievement was attributed to risk factors associated with students' sociocultural background—such as inadequate parenting practices and a lack of assimilation to American culture (Marger, 1997; Placier, 1996). For instance, it was often argued that norms of Black and Hispanic culture promoted parenting practices that did little to support children's learning and development (Pearl, 1991). Economist Thomas Sowell echoed the sentiment of these times in his chapter "*The Mexicans*", in which he wrote that "The goals and values of Mexican Americans have never centered on education"; a cultural 'deficit' he believed accounted for the poor academic achievement of Mexican American children (Ethnic America, 1981; p. 266). This argument—which was commonplace between the 1950's and 80's—spawned the development of educational remediation programs that were designed to provide low-income Black and Hispanic students with the developmental skills they were lacking (Pearl, 1991). In fact, the majority of early childhood education programs, including the HEAD Start program—one of the largest, longest running educational programs in the US—were developed under this premise (Pearl, 1991). The influence of deficit-thinking on the labeling of the children and families these programs were designed to target, can be observed by simply reading the title of a preliminary report on the *Early Training Project* (which inspired

the development of HEAD Start) published by its Principal Investigators, Susan Gray and Rupert Klaus in 1965: "*An Experimental Preschool Program for Culturally Deprived Children*".[1]

Turning to the present day, in its contemporary form, modern day deficit-thinking remains closely aligned with the 'softer' version of deficit ideology that was adopted in the post-Civil Rights Movement era, and continues to influence the ways in which the academic experiences of low-income Black and Hispanic students are framed (Castro, 2014; Marger, 1994; Menchaca, 1997; Placier, 1996; Valencia, 2010; 2012; Yosso, 2005). What is especially interesting is, that over the past century, several theories have been advanced to explain gaps in student achievement. One example is the sociolinguistic perspective posited by anthropologists in the 1960's, which attributed the school failure of minority students to cultural differences in communication styles that they argued encouraged teachers' misperceptions of these students as being 'unmotivated'. Unlike the cultural deprivation theories that preceded it, this argument interpreted miscommunications between teachers and their students as unintentional 'cultural blind spots', rather than placing the blame on a specific party (Erikson, 1987). Other examples include Ogbu's (1978) Oppositional Culture Theory, which portrayed the underachievement of minority students as their intentional rejection of a 'majority' culture that marginalizes people of color and limits their access to 'legitimate' professions and economic means, and 'structural inequality models' that explain school failure in the contexts of macropolitical and economic contexts (Pearl, 1991). However, none of these frameworks achieved and sustained a level of influence over educational policy and practice quite like the deficit model (Valencia, 1997).

[1] It is important to note that although there has been some debate over the effectiveness of the HEAD Start program, overall, the evidence suggests that the program has had favorable, long-term impacts on its participants (Bauer & Schanzenbach, 2016).

So, what is it about the deficit framework that has allowed it to remain both relevant and influential for well over 100 years? There is no doubt that the fluid nature of deficit ideology has facilitated the perception of contemporary deficit-thinking as being disconnected from its historically racist roots (Castro, 2014; Pearl, 1991; Valencia, 1997). Thus, in keeping with its protean nature, rather than focusing on 'weaknesses' of students' ethnic cultures, in its present form, deficit-thinking attributes chronic inequalities in income and education—or a 'culture of poverty'—for breeding disparities in parenting, children's academic performance, motivation, and behavior (Placier, 1996; Valencia, 2010). By attributing achievement gaps to a 'culture of poverty', modern day deficit-thinking avoids making a direct connection between race or ethnic background and academic achievement, and therefore lacks the overtly racist implications associated with prior iterations of the framework. However, in practice, most attempts at structural reform have been inadequate and ineffective. As a result, our current educational system is one that discourages overt racism, but advances policies and practices that negatively and disproportionately affect low-income Black and Hispanic students—which ultimately work to preserve the status quo (Castro, 2014).

A particularly powerful example of this can be found in the countless number of both federally and privately funded educational 'intervention' programs developed in the last 50 years (e.g., Banks, 1993). These programs typically vary in the population of students they target (e.g., high school students; college students) but endorse a common goal of ameliorating achievement gaps by helping students acquire the skills they need to be academically successful (*see* Gray & Klaus, 1970, for an example of one of the first educational 'intervention' programs). The idea of developing interventions for students in order to address problems that are borne from flawed structural systems, is one of the more nuanced aspects of modern-day deficit thinking. Moreover,

the manner in which these programs are framed, have been instrumental in sustaining deficit-thinking within our educational system (Castro, 2014). A major concern with the educational interventions is that oftentimes, students targeted for these initiatives are students from marginalized groups that are characterized as 'at-risk', 'underprepared', and 'disadvantaged'. Given that these programs center on students rather than focusing on the structural factors that breed and sustain disparities in achievement, the framing of marginalized students as 'at-risk' in these contexts lend themselves particularly well to being interpreted as reflecting deficit-thinking.

Although some anti-deficit scholars have called for an outright rejection of the 'intervention approach' to addressing achievement gaps—entertaining that argument is beyond the scope of this dissertation, and not a course of action I personally endorse. What is of primary interest to this dissertation, are the deficit-oriented labels that are used to characterize marginalized college students in the discourse surrounding attainment gaps, the pervasiveness of these characterizations, and the specific contexts in which they seem to be most prevalent—all of which are discussed in the remaining sections of *Part I*.

The 'At-risk', 'Underprepared', & 'Disadvantaged' Labels

In the context of discussions related to the attainment gaps in higher education and/or the educational interventions designed to close them, labels such as 'at-risk', 'underprepared', and 'disadvantaged' are often used to characterize the students typically targeted by these programs (Castro, 2014; Pearl, 1991; Smit, 2011; Valencia, 1997; 2010). Despite generally being used to describe students with similar characteristics, each of these labels have somewhat different origins and implications. For example, the 'at-risk' label—which was appropriated from the medical literature—has its roots in the field of epidemiology, where it has been historically used

to reflect the extent to which certain groups of people may be vulnerable to specific medical conditions. In educational contexts, the 'at-risk' label has been historically used to characterize students who are perceived as being increasingly likely to fail. In some instances, the label is used to designate students based on concrete indicators, such as GPAs or test scores below a certain threshold. However, this label is also commonly used in a vague manner to characterize students that may or may not share some common characteristics, but not necessarily based on specific criteria (Castro, 2014). The variability in the way the 'at-risk' label is utilized in achievement contexts creates ambiguity as to what the descriptor actually implies about students characterized as such. In the absence of additional context, this lack of clarity may encourage interpretations that attribute students' at-risk status to internal factors (e.g., intelligence; motivation) or background characteristics (e.g., race; SES; Castro, 2014; Pelligrini, 1991).

In contrast to the 'at-risk' label, the characterization of students as 'underprepared' originated in educational contexts, and is most often used to describe college students (typically lower classmen or incoming freshmen) who lack—or are perceived as lacking—the adequate amount of pre-college academic preparation required to succeed in college level courses (Lundell & Collins, 1999). Similarly to the 'at-risk' label, when used vaguely and in the absence of context, characterizing students as 'underprepared' can be interpreted as implying some fault on the part of the student for lacking the adequate amount of preparation. And lastly, the label 'disadvantaged'—which originated from cultural deprivation theories (e.g., '*culturally disadvantaged children [students]'*), is now used more commonly in educational contexts to characterize students from 'economically disadvantaged backgrounds', of which a disproportionate number are Black and Hispanic (Smit, 2011).

Throughout this dissertation, I refer to these three descriptors—'at-risk', 'underprepared', and 'disadvantaged'—both individually by name and collectively as deficit-oriented labels. I refer to them as such, because *all* labels carry with them assumptions about the individuals who bear them, and in this particular case, the 'at-risk', 'underprepared', and 'disadvantaged' labels communicate a similar underlying deficit-oriented message about the students characterized as such. More specifically, these labels imply that students are lacking in some way in comparison to some ideal—a message that is likely associated with negative motivational implications for the students characterized by these descriptors (Castro, 2014; Gray, 2013).

Finally, with respect to the contexts in which deficit-oriented labels are used—which is the topic I turn to next—given that this dissertation focuses on college students, my discussion is primarily limited to the use of deficit-oriented labels within higher education. However, it is important to note that there is literature documenting their use in primary and secondary education as well (see Croninger & Lee, 2001, for an example from secondary education, and Elbaum, Vaughn, Tejero-Hughes, & Watson-Moody, 2000; Macaruso, Hook, & McCabe, 2006 for examples from primary education).

The use of 'deficit-oriented' labels in higher education. There is considerable evidence that deficit-oriented labels are often used to characterize certain types college students. For instance, two quick searches on Google Scholar using the key words *"at-risk [underprepared] college students" and "academic performance"* yielded over 1,000 hits for the search using the *at-risk* label and approximately 700 for *underprepared*—and this was after I limited the results of both searches to materials published *after* the year 2000.[2] Of course, one could argue that these labels are used primarily as jargon within the literature, and that prevalent use by

[2] I included the keywords "*academic performance*" to eliminate medical research focused on populations of college students who are increasingly at-risk for experiencing certain adverse health outcomes.

researchers is not indicative of prevalent use in practical settings. However, references to certain types of college students as 'at-risk’, ‘underprepared’, and ‘disadvantaged’ can be easily found in the popular media and within practical settings.

For example, a 2017 article in the *New York Times* by Jennifer Steinhauer discussed the efforts of several large research-focused universities in using predictive data technology to identify ‘at-risk’ students. In describing the outcomes associated with dramatic increases in the proportion of first-generation students enrolling in 4-year colleges, the author explains “…with that growth came attendant failures among first-generation students, who were often ill prepared for the rigors of college life, and educational institutions that were not designed to serve them”. Because the author’s characterization of first-generation college students’ academic struggles ‘as failures due to their ill-preparedness’ lacks context, her statement can be easily interpreted as placing the burden of blame on the students for ‘lacking preparedness’ for college. Moreover, the author’s assertion that ‘the institutions these students attend are not designed to serve them’, also lacks context and does little to alleviate the culpability the former statement seemingly imposes on the students.

Another example can be found on the official website of the National Academic Advising Association (NACADA)—the largest advising-focused, professional organization in the country—where they have a section titled “*At-Risk Students*”. That section, which contains a brief article, titled “*Advising At-Risk Students*”, describes the characteristics of ‘at-risk’ college students by borrowing a definition from Maxwell (1997). Maxwell defines ‘at-risk’ college students as those whose “skills, knowledge, motivation, and/or academic ability are significantly below those of the 'typical' student in the college or curriculum in which they are enrolled” (p. 2). This characterization of ‘at-risk’ fails to state any concrete indicators for these students’ at-risk

status—for example—GPAs and/or standardized test scores that fall below a certain threshold. Instead, this description frames these students as lacking in contrast to their peers, by comparing their low levels of 'skills, knowledge, motivation, and/or academic ability' to those of the 'typical' student at their college. The implications of this description send a strong message about the deficit-thinking that underlies Maxwell's beliefs about these students and signals a tacit endorsement by NACADA as well. Of particular concern, is the fact that this information is being presented on the website of a large, national organization with over 12,000 practitioners who represent hundreds of colleges and universities across all 50 states. This scenario is more problematic than if this information was being presented on the website belonging to a single institution, because it is indicative of the extent to which these labels are endorsed within higher education.

Despite the evidence that deficit-oriented labels are used to characterize certain types of college students, to my knowledge, there is only one existing study (Castro, 2014) that is explicitly focused on documenting the use of these labels by practitioners in particular higher education settings. In this qualitative study, Castro (2014) used critical discourse analysis to examine deficit-thinking by STEM intervention programming staff from several large, research universities. To do so, she drew on data from structured interviews that were collected through a large-scale, NSF funded study investigating the underrepresentation of women and Black and Hispanic students in STEM fields. During the interviews, participants were asked various questions about the programs they oversaw and the students they typically worked with.[3] After reviewing 55 interviews, the author identified six in which the interviewees used the labels 'at-risk' and 'underprepared' to characterize the students they worked with. For example, given the

[3] The actual prompts and questions included in these structured interviews were not published in Castro's (2014) paper and therefore, not available to me.

lack of context in one administrator's characterizations of students as 'underprepared', can be easily interpreted as placing the locus of blame on the student for lacking preparedness (e.g., "we see underrepresented students certainly at all academic levels, a disproportionate number of *underprepared* that need various sorts of clearly directed assistance"; "the concern that I have at this university is bringing [in] significantly underprepared students"; p. 412). In contrast, another administrator describes his students as "a population of first-generation college students, students from low-income backgrounds, students that may come from academically underprepared high schools..." (p. 412), placing the blame for any academic shortcomings squarely on the '*underprepared* high schools', not the students.

Castro (2014) also found that several administrators used the label 'at-risk' in an ambiguous manner, lacking any explanation or context for the students' at-risk status (e.g. "I'm actually a former student of [This Program], because [This Program] focuses on high- potential and high-achieving students that, once they get here, are 'at-risk' as compared to the rest of the population.."). Another administrator's use of the *at-risk* label characterized students as a 'risk' or 'burden' to the institution, "We take some *high-risk students*, but we don't take very many…when you bring in an *at-risk student* and they fail, then people are so done with you" (p. 415). Another staff member combined the *at-risk* label and their reference to students' background characteristics "…working with an at-risk population…I started seeing that there was a trend…there were a lot of African American men on probation…a lot of student athletes on probation" (p. 415). In this case, the staff member does provide some concrete reason for the students' 'at-risk' status (i.e., they are on academic probation or liability); however, they also use the label in the same context as other student characteristics, such as race. Conflating students' background characteristics with their lived experiences of being denied equal access to

educational opportunities is extremely problematic, because it establishes an association between the characteristic (e.g., race) and students' at-risk status, rather than linking students' at-risk status to their experiences in an inequitable educational system. Importantly, despite characterizing students as 'at-risk' or 'underprepared', Castro (2014) noted that all of the administrators and program staff displayed a genuine interest in helping the students they served.

Castro's (2014) findings offer several insights with respect to the use of deficit-oriented labels within higher education. One, her findings suggest that the use of these labels is not exclusive to the literature or limited to online discourse—but that practitioners actively use these labels in the context of their work with students. Surprisingly, in at least some cases, the administrators and staff interviewed for Castro's study seemed to use these labels in a manner that was so casual, it could easily be interpreted as indifference (or perhaps, a lack of awareness) towards their implications for students. With respect to the prevalence in which these labels were used, just over 10 percent of the participants who were interviewed for that study used these labels to characterize their students in a manner that reflected deficit-thinking. If these findings are generalizable to the overall population of practitioners within higher education, they suggest that at least one in ten who work with similar populations of students would potentially rely on similar descriptors to characterize them. However, given the small sample of the study, this is likely a conservative estimate. Moreover, the ambiguous manner in which Castro observed the *at-risk* label being used, suggests that in some cases, these labels may be used as general descriptors for certain types of students, rather than as an official classification based on objective and concrete indicators.

The findings from this work provide a valuable foundation from which to build upon—particularly with respect to *students'* experiences with these labels. More specifically, given that

Castro's analysis centered on the administrators' and program staff's use of labels that communicated deficit-thinking *about* their students to an interviewer, it remains unknown whether these administrators might also use those labels during interactions *with* their students, thus communicating those deficit-oriented characterizations *directly* to them. Although this has yet to be explored empirically, there is evidence to suggest that this scenario is not unlikely. For instance, the unreserved manner in which the administrators in Castro's study communicated these deficit-oriented characterizations to the individuals interviewing them suggests that labels like 'at-risk' and 'underprepared' may be deeply embedded in discourse surrounding educational interventions. Further, the prevalence of deficit-oriented labels in the literature and on institutional and organizational websites—all of which are accessible to students—implies that although they may be perceived as negative terms, they are not necessarily perceived as socially unacceptable (or 'politically incorrect').

Given the substantial amount of existing literature documenting the negative effects of labeling, the question of whether these labels are communicated directly to students is a legitimate one with important implications (e.g., Field, Hoffman, Peter, & Sawilowsky, 1992; Shifrer, 2013). Thus, in the next section of this literature review, I use related evidence from the psychological and educational literatures as a basis from which to discuss the potential consequences that might be associated with communicating deficit-oriented characterizations directly to students.

Part II

The Consequences of Labeling

There is an impressive body of research—dating back to the 1950's—examining the effects of labeling (Becker, 1951; 1963; Lemert, 1967; 1974). Much of this work followed the development of Labeling Theory, which originated in the field of sociology, primarily with the purpose of explaining societal influence in perpetuating criminal behavior (Becker, 1951; 1963; Lemert, 1967; 1974; Link et al., 1999). Early labeling theorists Edwin Lemert and Howard Becker proposed that labeling individuals as 'deviants' and 'criminals' produced *secondary deviance*—essentially, subsequent criminal behavior resulting from the labeling process itself (Becker 1951). They argued that the negative stereotypes attached to these descriptors adversely impacted the self-perceptions of labeled individuals, as well as the manner in which they were perceived and subsequently treated by others (Lemert, 1967). Lemert and Becker portrayed chronic criminal behavior as the result of both internal and external processes, such that through the course of interactions with others in their communities, individuals labeled as 'deviants' and 'criminals' would eventually internalize those identities and use it to condone subsequent behavior congruent with the label (Becker, 1951; Lemert, 1974). Findings from correlational studies examining associations between receiving an official designation of *felon* and rates of recidivism largely support this argument (e.g., Chiricos et al., 2007). Studies exploring the *mechanisms* driving these potential associations indicate that *both* the stigmatization associated with these labels and existing structural limitations that often stem from this stigmatization (e.g., difficulties finding employment and housing) prevent ex-felons from reintegrating into society through legitimate means and work in tandem to influence recidivism (e.g., Reisig, Bales, Hay, & Wang, 2007).

The influence of Lemert and Becker's early work on subsequent labeling research is irrefutable; since the 1970's, researchers have extended their work on the effects of labeling to include individuals from other marginalized groups, such as those living with mental illness and/or physical disabilities, the elderly, and members of racial and ethnic minority groups (e.g., Angermeyer & Matschinger 1997; Link et al., 1999; Martin et al., 2000; Tepper, 2014). Drawing from Lemert and Becker's representation of labeling effects, much of the labeling research has focused on exploring those consequences as a result of one or more of the following mechanisms: (a) social stigmatization, (b) self-stigmatization, and/or (c) perceived stigmatization (Becker 1951; Lemert, 1967; Link & Phelan, 2001).

Social stigmatization (or *social stigma*, for short) refers to societal beliefs and attitudes with respect to individuals pertaining to a particular group—and includes the accompanying stereotypes, prejudice, and discrimination experienced by individuals as a result of their membership to a stigmatized group (Link & Phelan, 2001). There has been a considerable amount of work exploring social stigma in the context of psychiatric and psychological pathology, specifically in relation to the stigmatization of individuals designated as 'mentally ill', or with respect to labels associated with specific diagnoses (e.g., *psychopath*; *schizophrenic*). Findings from this work consistently show that individuals labeled as *mentally ill* or with respect to a specific diagnosis, are often subject to prejudice and discrimination as a result of being perceived as possessing 'undesirable' characteristics. For instance, in a study by Angermeyer & Matschinger (2005), the authors examined the social stigma associated with individuals designated as 'schizophrenic'. To do this, they presented adult participants with hypothetical case study vignettes that described as individual who displayed behavioral patterns typically associated with Schizophrenia. Some participants were presented with vignettes that referenced

the individual as being 'schizophrenic', whereas other participants were presented with vignettes about the same individual that did not use a label to characterize their behavior. The researchers found that participants were more likely to perceive the individuals in the vignettes as 'dangerous and unpredictable' when the individual's behavior was characterized as a 'schizophrenic', relative to participants who were presented with the vignette that did not include information about the individual's Schizophrenia diagnosis. Similar work with college students have yielded consistent findings, such that college students who are labeled as 'mentally ill' are more likely to be perceived as 'dangerous' by their peers, relative to students not designated as such (Phelan & Basow; 2007). Several studies have also reported associations between participants' perceptions of people who are labeled as 'mentally ill' and the desire for greater social distance (Angermeyer & Matschinger 1997; Link et al., 1999; Martin et al., 2000).

Both perceived stigmatization and self-stigmatization (perceived stigma and self-stigma, for short) are mechanisms that focus on *internal* psychological processes as driving the effects of labeling. *Perceived stigmatization* refers to the extent to which individuals in a stigmatized group believe that membership in that group is associated with negative stereotypes, and that they will experience devaluation and/or discrimination by others as a result of those stereotypes (Link & Phelan, 2001). Previous studies examining this mechanism have reported associations with several negative psychosocial and behavioral outcomes, such that individuals who perceive greater levels of stigmatization have also been found to report lower levels of overall quality of life, display higher levels of helpless behavior, and report experiencing difficulty with social interactions and relationships more frequently (Alonso et al., 2009; Bedini, 2000).

In contrast, *self-stigmatization* is the process by which an individual *internalizes* the negative social or perceived stigma associated with a given label or identity, such that the effects of perceived stigma typically operate as a function of the internalization processes associated with self-stigma (Pattyn, Verhaeghe, Sercu, & Bracke, 2014). For example, in a study by Tepper (2014), the author explored participants' perceptions of being labeled as a *senior citizen* and the influence of their perceptions on subsequent use of senior citizens discounts. Through a qualtitative component of the study that included in-depth interviews with participants aged 50 and over, Tepper probed participants' perceptions of being labeled as a *senior citizen*—both in terms of what the label itself implied (e.g., 'being old in a society that values youth') and in reference to perceptions surrounding senior citizens who take advantage of discounts offered through the American Association of Retired Persons (AARP). The author's analysis of this data suggested that participants' percepeptions of the stigma associated with the elderly prompted concerns that they would inevitably internalize negative characterizations, which underlied their hesitation to use AARP discounts, particularly for participants younger than 55. Existing research on the effects of the internalization processes associated with self-stigma, often report connections between self-stigma and maladaptive psychological outcomes (e.g., Corrigan, Larson, & Ruesch, 2009; Corrigan & Rao, 2012; Hatzenbuehler, 2009; Latalova, Kamaradova, & Prasko, 2014; Phelan et al., 2015). For instance, using a sample of medical students, Phelan and colleagues (2015) found that internalizing negative beliefs about being overweight were associated with increased levels of stress, depression, and anxiety, and lower levels of self-esteem in medical students who were clincally overweight, compared to those who were considered normal or underweight by clinical standards.

The moderating role of perceptions, beliefs, and attitudes. Findings from the body of research on labeling provide powerful evidence of the negative consequences of stigmatization psychological processes involved. However, some findings also reflect the nuanced nature of stigmatization processes. Specifically, more recent work on labeling has explored individual differences in the extent to which people experience the adverse effects of stigmatization—be it socially inflicted, self-inflicted, or perceived (e.g., Cadinu, Maass, Lombardo, & Frigerio, 2006; Martens, Johns, Greenberg, & Schimel, 2006). Findings from some of this work suggest that the effects of stigmatization are often moderated by individuals' perceptions, beliefs, and attitudes.

For instance, a study by Camp, Finlay, and Lyons (2002), explored the role of beliefs in mitigating the effects of stigma on self-esteem and self-perceptions, with a group of women who had received a mental illness diagnosis in the past year. The authors sought to challenge the assumption that most individuals labeled as 'mentally ill' incorporate that label as a central element of their own identity and accept the pervasive negative characterizations of mental illness that are commonplace within society. Through a series of in-depth interviews and subsequent qualitative analyses, the authors examined the manner in which participants' understood the social perceptions surrounding mental illness, the medical representation of their diagnosis as defined in the Diagnostic and Statistical Manual of Mental Disorders (DSM), and their interpretations of those portrayals. The results of their analyses showed that in general, most of the women were explicitly aware of the negative stereotypes associated with people who are living with mental illness (e.g., "They just think you're some kind of lunatic or idiot, that's what I feel people look as [sic] you as unless they understand"; "When my children found out they just kept saying I was mad…"; p. 827) and often attributed what they perceived as inaccurate stereotypes to misrepresentations of mental illness in the media (e.g., "You feel like

saying…Well what do you base your facts on? And they wouldn't know, they'd most probably say in the media."; p. 828). Further, a majority of the women outright rejected the specific label assigned with their diagnosis, because they believed that the label oversimplified their lived experiences or because they disagreed with the manner in which their disorder was defined in the DSM (e.g., "The definition I have in my book included lots of other things that in no way anything like the person I am…I am completely the opposite...I don't see myself that way at all."; p. 828). Finally, all of the women offered positive self-perceptions of themselves, some of which were in relation to progress they had made as a result of their hard work (e.g., "I've been living on my own for a whole year now and I'm really proud of that...I'm proud that I've done a lot of hard work on myself."; "I am quite proud of the achievements I have made despite every effort of people to the contrary…"; p. 828). The authors also noted that a common challenge among the women was lacking a sense of belonging in their social circles—feelings that some of the women acknowledged often led them to avoid social situations they previously enjoyed partaking in, but also exacerbated their feelings of social isolation. Overall, Camp, Finlay, and Lyons' (2002) findings are consistent with prior research documenting the effects of stigma on sense of belonging, but also suggest that the extent to which individuals' perceptions of stigma and their own beliefs are misaligned, may reduce the likelihood of experiencing the effects associated with internalized stigma (i.e., self-stigma).

Findings from similar research exploring the influence of other types of beliefs and attitudes have also found that they can serve to either accentuate or minimize the negative effects of stigma (Major & Sawyer, 2009; Moore, Stuewig, & Tangney, 2016; Moore et al., 2016). For instance, one study found that social identity beliefs that centralize the stigmatized identity—such that the individual's identification with a stigmatized identity or group is central to their

self-concept—may increase the likelihood of experiencing both self-stigma and perceived stigma (Moore, Stuewig, & Tangney, 2016). Another study reported an indirect link between possessing an optimistic worldview and diminished effects of social stigma, such that expecting optimistic outcomes for oneself was positively associated with an external attributional style, which is characterized by an individual's tendency to explain outcomes as being a result of environmental or contextual factors that are beyond their control (or perceived to be). In turn, an external attributional style was negatively associated with the effects of discrimination and prejudice (Major & Sawyer, 2009). Another study yielded a direct, negative association between optimism and experiences of self-stigmatization (Moore et al., 2016). Altogether, the findings from this work support rejecting the notion that the effects of stigmatization are universally experienced by all stigmatized persons to the same extent.

Labeling & Stigmatization in Achievement Contexts

Specifically, within the educational literature, there has been a fair amount of research focused on examining the effects of labels on students in achievement contexts (e.g., Field, Hoffman, Peter, & Sawilowsky, 1992; Elbaum, 2002; Haring et al., 1992; Osterholm, Nash, & Kritsonis, 2011; Shifrer, 2013). Findings from the educational work are consistent with other labeling research, in that they also indicate that (a) the consequences of labeling students are often driven by stigmatization processes, (b) stigmatization processes can have profound influence over individuals' self-perceptions in relation to a stigmatized identity and their psychosocial outcomes, and (c) beliefs play an important role in moderating these effects. In this section, I highlight some of the research on the effects of labeling students in achievement contexts. Given that the effects of students' labeling experiences operate as a function of mechanisms associated with stigmatization processes, I discuss the research on labeling in the

broader context of academically stigmatizing experiences. I focus particularly on the consequences of these experiences on students' *academic*, their *sense of belonging* and 'fit' in academic (or achievement) contexts, and experiences of *stereotype threat*. Lastly, I present evidence surrounding the role of students' motivational beliefs in moderating the effects of academic stigmatization.

The consequences of students' stigmatizing experiences on their academic self-perceptions and sense of belonging. Research on the social stigma associated with labeling students as *learning disabled (LD)* or as requiring 'special education', have found that these students often experience prejudice and discrimination as a result (e.g., Field, Hoffman, Peter, & Sawilowsky, 1992; Shifrer, 2013). In a study by Shifrer (2013), the author found that when middle school teachers were asked to access the comptency of students with similar achievement records, they were more likely to perceive learning disabilities in students who were labeled *LD* compared to those who were not. Moreover, the *LD* label influenced teachers' longterm expectations for students as well, such that they were significantly less likely to indicate that they expected a student would obtain a 4-year college degree, if the student was labeled *LD*, relative to those were not labeled *LD*.

In turn, findings from related research on teachers' expectations indicate that teachers' biased beliefs about stigmatized students have implications for students' academic self-perceptions and ultimately, their academic achievement (McKown & Weinstein, 2002; Weinstein, 2002). For instance, findings from this work show that students' expectations often mirror their teachers' expectations, such that students whose teachers report low expectations for them, typically have lower expectations for themselves (McKown & Weinstein, 2002; Weinstein, Marshall, Sharp, & Botkin, 1987).). Studies examining the mechanisms for these

associations suggest that teachers' often communicate their high or low expectations for students through differential treatment—such as offering low expectation students unsolcited help—as well as other non-verbal behavioral cues (Graham, 1984; McKown & Weinstein, 2002; Stipek et al., 2001). Given the strong link between students' academic self-perceptions and their achievement, these findings have important implications for students who are subject to biased perceptions of their academic competence as a result of social perceptions associated with a given label (e.g., learning disabled; Fall & Roberts, 2012; McKown & Weinstein, 2002; Pollard, 1993; Zimmerman, 1990).

Findings from related research on academic stigmatization indicate that these experiences are often associated with students' diminished sense of belonging in academic contexts (Walton & Cohen, 2007; Winograd & Rust, 2014). For instance, in a study by Winograd & Rust (2014), the authors examined the effects of self-stigma on students' academic help seeking habits and sense of belonging at their university, with a sample of primarily Black and Hispanic first-generation college freshmen. The authors used a measure of self-stigma to assess the extent to which students had internalized negative beliefs about seeking academic help (e.g., "Seeking help would make me feel less intelligent"), a measure of sense of belonging that assessed the extent to which students perceived the college environment as warm, supportive, and comfortable (e.g., "I feel as though no one cares about me personally on this campus"), and measures of students' awareness and use of academic support services (e.g., "I [know about] have gone to the Tutoring Center"). Findings from a correlational analysis showed that students' scores on the measure of self-stigma were significantly and negatively associated with students' sense of belonging at their university and their *awareness* of support services, such that on average, the more students' reported internalizing the negative characterizations associated with

seeking academic help, the less likely they were to report feeling like they belonged at their university and knowing how and where to access academic support services. Winograd & Rust's (2014) results are consistent with other work and indicate that experiences of stigmatization can diminish feelings of belonging or 'fit' and ultimately encourage socially isolating behavioral patterns (Camp, Finlay, & Lyons, 2002). Similar associations between students' experiences of stigmatization and their feelings of belonging in college have been reported by several studies (e.g., Aronson, 2004; Inzlicht & Good, 2006; Walton & Cohen, 2007; 2011).

Stigmatization, stereotype vulnerability, and stereotype threat. Within the educational literature, there has also been a considerable amount of work exploring *stereotype vulnerability* and *stereotype threat* (Aronson & Steele, 1995; 2004; 2005; Chavous et al., 2004; Inzlicht & Good, 2006; also see Spencer, Logel, & Davies, 2016, for review). Similar to the concept of *perceived stigma—stereotype vulnerability* refers to the extent to which an individual believes that (a) there are negative stereotypes associated with a social group they belong to and (b) they are subjected to prejudiced beliefs and discrimination as a result of their membership in that group. Aronson and Steele (2005)—who were the pioneering researchers on stereotype vulnerability—developed this construct in an effort to explain the achievement gaps between Black and White students who attended comparable schools and came from similar middle-class backgrounds. The researchers posited that relative to their White peers, the pervasive negative stereotypes about Black students' academic competence made them increasingly susceptible to believing these characterizations were widely endorsed by others, and therefore, were more likely to experience the adverse effects associated the 'threat' of confirming those stereotypes (i.e., *stereotype threat*). Aronson and Steele (2005) characterized this feeling of 'stereotype threat' as the manner with which "students react to stereotypes" (p. 440). They argued that in

contexts where the negative stereotypes about Black students' academic competence were more likely to be activated (or made salient)—such as high stakes testing or other achievement contexts—these students were more likely to experience psychological distress resulting from a fear of confirming those stereotypes, which ultimately hindered their subsequent performance on evaluative tasks (e.g., exams). In contrast, White students from middle-class backgrounds who did not have to contend with negative stereotypes about their academic competence, would be less likely to experience the additional 'threat' of confirming social perceptions during evaluative situations (Aronson & Steele, 1995; 2005).

Over two decades and hundreds of studies later, in addition to observing stereotype threat effects in students from other groups stigmatized by pervasive academic stereotypes, such as Hispanic students (Guyll et al., 2010; Gonzales, Blanton, & Williams, 2002; Woodock et al., 2012), students from low SES backgrounds (Croizet & Clair, 1998), and women in the context of STEM domains (Beasley & Fischer, 2012; Keller, 2007; Shapiro & Williams, 2012), findings from this work have also yielded several important insights on this phenomenon. For instance, findings from laboratory based studies that involve inducing stereotype threat through experimental manipulations, indicate that although membership in a stigmatized group is *not* a requisite for experiencing stereotype threat, but these experiences are typically still driven by an established stereotype (e.g., gender differences in sensitivity; Leyens, Désert, Croizet, & Darcis, 2000). That said, students from stigmatized groups are more likely to experience stereotype threat than students from non-stigmatized groups (Aronson et al., 1999; Chavous et al., 2004; Spencer, Logel, & Davies, 2016). The evidence also indicates that unlike the effects of self-stigma, the negative effects of stereotype threat do not seem to seem to operate as a function of internalization. More specifically, students need not endorse the negative stereotypes associated

with their group in order to temporarily experience the 'threat' associated with confirming that stereotype; instead, their awareness (or perception) that the stereotype exists is enough to induce stereotype threat (Aronson et al., 1999). However, *chronic* experiences of stereotype threat have been shown to diminish students' sense of belonging in academic environments and promote deidentification with academics and/or achievement—both of which are believed to function through internalization processes (Spencer, Logel, & Davies, 2016). These findings suggest that although students from non-stigmatized groups may experience the acute effects of stereotype threat on their performance in certain contexts, stigmatized students are more likely to experience chronic stereotype threat and therefore, also more likely to internalize those characterizations, and suffer long-term negative effects on their sense of belonging and self-perceptions as a result.

The importance of students' academic self-perceptions and sense of belonging in college. The well-documented effects of stigmatization on students' academic self-perceptions and sense of belonging are extremely problematic and a serious cause for concern. This is due to the fact that over several decades of research, using samples of students from diverse racial (Chavous et al., 2004; Freeman, Anderman & Jensen, 2007; Soria & Stebelton, 2012), ethnic (Soria & Stebelton, 2012), and socioeconomic backgrounds (Soria & Stebelton, 2012), findings have consistently indicated that students' academic self-perceptions and sense of belonging in academic contexts are critical to their success in those environments. More importantly—as I discuss further in this section—there is evidence that for students from groups associated with stigmatized academic identities (e.g., students from low-income backgrounds, first-generation college students, and/or Black and Hispanic students), these factors can disproportionately influence their academic outcomes (Chavous et al., 2004; Inzlicht & Good, 2006).

Drawing from self-determination theory—which argues that the need for social connectedness and belonging is a basic psychological need—educational researchers argue that students' motivation and achievement in academic environments largely depend on their feelings of belonging in that environment (Deci & Ryan, 1985; 2000; Walton & Carr, 2012; Walton & Cohen, 2007). With respect to college students, *sense of belonging* typically refers to the extent to which these students feel like they 'fit in' and are a valued member of their university's community (Good & Dweck, 2012). Research exploring college students' sense of belonging typically find that these feelings of belonging play a critical role in students' motivation, persistence, and ultimately achievement in college (Good & Dweck, 2012; Soria and Stebelton, 2012; Walton & Carr, 2012; Walton & Cohen, 2007). For instance, sense of belonging has been shown to predict female undergraduate students' desire to pursue math careers in the future—a domain in which women are often stigmatized—and can predict their actual achievement in college math courses as well (Good & Dweck, 2012). Another study by Soria and Stebelton (2012) explored associations between students' sense of belonging, their self-reported academic engagement (e.g. 'I interact with faculty during lectures [class]'), and retention rates, with a sample of first-generation and continuing-generation college freshmen. The authors found that students' sense of belonging predicted both their academic engagement and their first to-second-year retention, such that greater levels of sense of belonging in college were associated with both higher odds that students would return their second year and greater levels of academic engagement. However, relative to continuing-generation students, first-generation students reported significantly lower levels of academic engagement and were 45% less likely to return to college after their freshman year.[4] Findings from other studies have found positive associations

[4] Odds ratio after holding all other demographic variables constant.

between college students' sense of belonging and their intrinsic motivation and self-efficacy beliefs—both of which are strong predictors of academic achievement (e.g., Freeman, Anderman & Jensen, 2007; Komarraju & Nadler, 2013; Pajares, 2003).

Moreover, Walton & Cohen (2004) argue that in addition to lacking a feeling of belonging or 'fit' at their university, college students from groups that have been historically stigmatized with regard to academics—such as low-income students, first-generation students, and Black and Hispanic students—are also more likely to feel a sense of *belonging uncertainty* in academic contexts—which they define as a "global concern about the quality of one's social ties" (p. 83). To explore this phenomenon, the authors conducted a series of studies with undergraduate students. In one study, they found that assigned to an experimental condition that involved an uncertainty manipulation designed to make students questions the number of friends they had in their academic program, Black students reported weaker feelings of 'fitting in' within their academic department and lower perceptions of their potential to succeed in their program, relative to Black participants in two other conditions that were not designed to make students feel isolated in their majors. In contrast, White students were unaffected by the manipulation. However, in a subsequent study in which the authors tested an intervention designed to ease *belonging uncertainty*, they found that Black students experienced increases in achievement, whereas White students did not.

Existing educational research has traditionally included students' *academic self-perceptions* as a component of *academic self-concept*, along with students' beliefs and attitudes about themselves in relation to their academic skill and performance (Guay, Marsh, & Boivin, 2003; Marsh, Byrne, & Shavelson, 1988; Marsh et al., 2005). In addition to the work on self-concept, there is also research that focuses only on students' academic self-perceptions in

relation to other outcomes related to motivation and achievement (e.g., Banks & Woolfson, 2008; Meece & Courtney, 1992). For the purpose of this dissertation, I limit my discussion to the latter literature—focusing primarily on college students. Findings from this work indicate that college students' academic self-perceptions are strong predictors of student achievement (e.g., Banks & Woolfson, 2008). In fact, students' academic self-perceptions have been found to out-predict some objective measures of ability (Colbeck, Cabrera, & Terenzini, 2001; Hackett et al., 1992; Pajares & Miller, 1994). This is likely due to the fact that students' academic self-perceptions are associated with other factors related to achievement, that are not typically assessed in traditional measures of ability, such as effort (e.g., Meltzer et al., 2004). For instance, using self-reported data from both students in the 4^{th}-9^{th} grades who were labeled as *LD* and their teachers, Meltzer and colleagues (2004) found that students designated as *LD* with higher academic self-perceptions reported exerting significantly more effort on their school work and using effective learning strategies to a greater extent than *LD* students with lower academic self-perceptions. Moreover, teachers' data indicated that they perceived *LD* students with higher academic self-perceptions as displaying competence levels comparable to those of their non-*LD* peers, which reflects the associations between teachers' expectations, students' academic self-perceptions, and achievement found in related studies. Associations between students' academic self-perceptions and negative affective and psychosocial outcomes have also been reported by previous work, such that students' lower perceptions of academic competence were associated with higher levels of concern and tension regarding their academic performance (e.g., Putwain, Woods, & Symes, 2010). Lastly, findings from two meta-analyses showed that increases in college students' academic self-perceptions were associated with higher scores on cognitive measures and higher retention rates (Rhee & Hurtado, 2009; Robbins et al., 2004).

Altogether, these findings underscore the multitude of ways in which students' academic self-perceptions and sense of belonging influence their achievement outcomes in college. They also point to the nuanced manner in which these factors disproportionately affect students from stigmatized groups, relative to those from non-stigmatized groups. For instance, relative to students from non-stigmatized groups, those who do belong to stigmatized groups generally enter higher education with lower levels of belonging in college, and are also more vulnerable to experiencing diminished feelings of belonging throughout their time in college (e.g., Walton & Cohen, 2007; Winograd & Rust, 2014). The cyclical nature of students' academic self-perceptions and their achievement can pose unique challenges for stigmatized students who experience academic struggles, such that it may reinforce students' negative perceptions about their academic competence, which in turn, would serve to further cement their feelings of not belonging in college.

Mitigating the effects of stigma and stereotype threat through students' motivational beliefs. Considering the serious implications associated with students' experiences of stigmatization, some recent work has focused on targeting students' beliefs in an attempt to mitigate these deleterious effects (Aronson, Fried, & Good, 2002; Good, Aronson, & Inzlicht, 2003). Students' motivational beliefs are important because they act as a lens—or framework—from which students draw on to interpret their experiences—which in turn, influence their responses to those experiences (Dweck & Leggett, 1988; Halligan, 2006; 2007, Hoffman, 2015; Lai, 2011). The basis for using students' beliefs as a strategy for mitigating the consequences of stigmatization stems from a considerable amount of evidence documenting the influence of students' beliefs over virtually every factor related to achievement—as well as achievement itself (e.g., Dweck & Leggett, 1988; Dweck & Molden, 2017; Eccles & Wigfield, 2002). In this

section, I limit my discussion to two specific types of beliefs—students' *academic mindsets* and their *racial [ethnic] identity beliefs*—both of which have shown particular promise in this regard (Aronson, Fried, & Good, 2002; Butler-Barnes et al., 2018; Chavous et al., 2008; Good, Aronson, & Inzlicht, 2003).

Academic mindsets. Students' *academic mindsets* (i.e., their implicit theories of intelligence), a concept first developed by psychologist Carol Dweck, refers to students' beliefs about the nature and malleability of intelligence and academic ability (Diener & Dweck, 1978; 1980; Dweck, 1999). More specifically, some students have a *growth mindset*—viewing intelligence and academic capacity as malleable skills that can be developed through hard work and effort, whereas other students have a *fixed mindset* and perceive intelligence and academic ability as being relatively inherent and stable. When faced with academic challenges or adversity, students' academic mindsets have been shown to influence their subsequent *affect*, *cognition*, *motivation*, and *behavio*r (Diener & Dweck, 1978; 1980; Dweck & Bandura, 1986; Dweck & Molden, 2017; Ehrlinger, Mitchum, & Dweck, 2016; Good, Aronson, & Inzlicht, 200; Haimovitz, Wormington, & Corpus, 2011). For instance, students who endorse a *growth mindset* are more likely to react to their struggles with a renewed sense of motivation, increased effort, use of adaptive coping strategies, perceive the challenge as a learning opportunity, and display more positive affect than students with a *fixed mindset*—who in contrast, are more likely to display helpless behavioral patterns and negative affect (Dweck & Leggett, 1988; Dweck, 1999). Dweck (1999) argues that students' academic mindsets, which reflect in their perceptions of their ability as being malleable or fixed, influence their interpretations of performance cues in relation to their ability. More specifically, students with a growth mindset tend to interpret challenge as an indication that they need to increase their effort and work harder, whereas students with a

fixed mindset are likely to interpret the same struggle as an indication that they have reached the limits of their capacity.

Findings from other research on academic mindsets have revealed other valuable insights about these beliefs. For instance, Claro, Paunesku, & Dweck (2016) found that the positive influence of a growth mindset extended to students across all SES backgrounds and seemed to temper the adverse effects of poverty on students' academic achievement—which indicates that students from low-income backgrounds may benefit disproportionately from endorsing a growth mindset. There is also evidence that these beliefs can be changed to reflect a stronger growth orientation in the short-term using experimental tasks and in the long-term through targeted interventions (Aronson, Fried, & Good, 2002; Blackwell, Dweck, & Trzesniewski, 2007; Miu & Yeager, 2015; Yeager et al., 2016).

Specifically of interest to the current research, is the evidence that adopting a stronger growth mindset can help alleviate the effects of stereotype threat in students from academically stigmatized groups (Good, Aronson, & Inzlicht, 2003; Aronson, Fried, & Good, 2002). For example, in a study by Aronson, Fried, & Good (2002), the authors tested an experimental intervention designed to reduce experiences of stereotype threat in Black undergraduate students by encouraging the adoption of a growth mindset. The study tested an intervention condition designed to promote students' endorsement of a growth mindset, against another experimental condition designed encourage participants to think about intelligence in a 'fixed' but very domain-specific manner that emphasized everyone has strengths and weaknesses, and a control condition that did not address students' views about intelligence at all. In addition to measuring students' perceived stereotype threat and their academic mindsets, they also measured their identification with achievement (e.g., "Considering all the things that matter to you and make

you who you are, how important is academic achievement?"), and their enjoyment of academics (e.g., "How much do you enjoy the educational process—studying, going to class, taking tests, etc.—at Stanford?"), both several days and several weeks after the intervention concluded. They also used students' grades at the end of the semester and their prior SAT scores to assess their academic performance and to control for differences in prior achievement. The results of this study showed that for the Black students in the sample, participating in the intervention condition successfully increased their views of intelligence as being malleable, both in the short and long-term. Moreover, relative to Black students in the other two conditions, those in the intervention condition reported higher identification with academic achievement, higher enjoyment of academics, and greater achievement gains at the end of the year. Interestingly, Black students' perceptions of stereotype threat did not differ between the three conditions—a result that the authors' note may imply that the effect of the intervention functioned through its influence on students' reactions to a stereotype threatening environment, rather than influencing their perception of those environments. In conjunction with these findings, the malleable nature of students' growth mindsets and positive, far-reaching influence, suggest that they may be effective in abating the negative effects associated with other types of stigmatizing experiences as well.

Racial identity beliefs. There is a strong consensus within the psychological and educational literatures regarding the important role of *racial identity beliefs* on the academic outcomes of Black and Hispanic students (Fordham & Ogbu, 1986; Penn, Gaines, & Phillips, 1993; Sellers et al., 1997). Historically though, there has been less agreement with respect to the nature of that role (Sellers et al., 1997). Some scholars argued that strong racial identification places these students at a disadvantage due to the negative academic characterizations about

Black and Hispanic students (Fordham & Ogbu, 1986; Penn, Gaines, & Phillips, 1993). In contrast, others argued that strong identification with their racial or ethnic background should shield these students from the consequences of stigmatization (Chavous et al., 2008; Sellers et al., 1997). However, contemporary frameworks of racial identity represent these identities as multi-faceted and dynamic systems, which would seem to contradict earlier arguments that characterized their influence on students' educational outcomes as being broad and linear (Sellers et al., 1997).

An example of one such framework, is the Multidimensional Model of Racial Identity (MMRI) developed by Sellers and colleagues (1997), which argues for three stable and measurable dimensions of racial identity: (1) *centrality*—which is the extent to which a person normatively defines themselves with regards to their race, (2) *ideology*— which represents a person's beliefs, opinions, and attitudes with respect to the way they believe members of their race should act, and (3) *regard*—which represents the extent to which a person feels positively or negatively towards members of their race. Existing research using the MMRI as a theoretical framework, suggest that some dimensions of racial identity may be more advantageous for buffering students from the negative effects of stereotypes and discrimination than others (Chavous et al., 2008; Gummadam, Pittman, & Ioffe, 2016). For instance, in a study by Chavous and colleagues (2003), the authors employed a cluster analysis to explore various profiles of Black identity beliefs with a sample of African-American high school seniors, and used the Multidimensional Inventory of Black Identity (MIBI)—designed by Sellers and colleagues (1997)—as a measure of students' racial identity beliefs. The results of their analysis indicated that in general, students with higher *centrality beliefs* (i.e., perceived their racial identity as more central to their self-concept), higher *private regard beliefs* (i.e., positive beliefs about themselves

as a member of their race), and lower *public regard beliefs* (i.e., positive beliefs about African-Americans), attended school more regularly and were more likely to go to college, relative to those with moderate *centrality* and *private regard beliefs*, but higher *public regard beliefs*. However, compared to students in both of these clusters, students who were low across all three beliefs (i.e., *centrality*, *private regard*, and *public regard*) were associated with the lowest high school attendance and college enrollment rates. Additionally, other studies have also found similar buffering effects of higher centrality beliefs, such that they also seem to mitigate the negative effects of stigmatizing school environments on students' motivation—particularly for African-American girls (Butler-Barnes and colleagues, 2018; Chavous et al., 2008).

However, other studies have yielded evidence to suggest that centrality beliefs may be more beneficial for some academic outcomes—such as engagement and retention—but immaterial and possibly somewhat disadvantageous with respect to others (e.g., Awad, 2007; Cokley, McClain, Jones, & Johnson, 2012; Harper & Tuckman, 2006). For instance, a study by Cokley and colleagues (2012) found a moderately negative association between Black high school students' racial centrality beliefs and their GPA. However, this study relied on a sample of 96 students, whereas Chavous and colleagues (2003) had a sample size of over 600 students. In another study by Harper & Tuckman (2006), the authors argued that they were unable to replicate some key findings from Chavous and colleagues (2003), despite relying on the same methodology as the former and a comparable sample of Black 9th and 12th grade high school students. The authors argued that in contrast to the earlier study, their analysis did not yield a cluster group of students higher in centrality and private regard beliefs but lower in public regard beliefs, who were also associated with the highest high school attendance rates and subsequent college enrollment rates in Chavous and colleagues (2003). They also found that the cluster

group with lower overall identification beliefs (i.e., lower in centrality, private regard, and public regard beliefs) had an average GPA that was significantly higher than that of the group with higher beliefs across all three dimensions. The authors portrayed this particular finding as being in direct conflict with Chavous and colleagues' finding that the students with lower overall identification beliefs were associated with the poorest academic outcomes. However, Chavous and colleagues did not find significant differences in high school GPA between any of the cluster groups—instead, they reported that the group of students with lower beliefs across all three dimensions were associated with the lowest high school attendance and subsequent college enrollment rates. Moreover, the difference in GPAs between the lower and higher identification groups in Harper & Tuckman's study may have been statistically significant, but for each sets of GPAs they reported (one for freshmen and one for seniors), the difference was less than the standard deviation reported for either of the means.

In addition to the work on racial identity, another subset of the broader *social identity* literature has focused on students' *ethnic identity beliefs*—defined as "an enduring, fundamental aspect of the self that includes a sense of membership in an ethnic group and the attitudes and feelings associated with that membership" (Phinney, 1996; p. 222). The prior research on students' ethnic identity beliefs have explored the influence of these beliefs on factors related to students' academic outcomes—using measures similar to those used in the racial identity research (i.e., they assess students' perceptions, sense of belonging, and attitudes towards their ethnic group; e.g., Gummadam, Pittman, & Ioffe, 2016; Phinney & Alpuria, 1990; Umana-Taylor et al., 2014). The findings from this work extend those from the racial identity research, because they incorporate students from other stigmatized groups that are typically excluded from the former (e.g., non-Black Hispanics; Gummadam, Pittman, & Ioffe, 2016; Valencia, 2012). For

instance, a study by Gummadam, Pittman, & Ioffe (2016) explored associations between students' ethnic identity beliefs and their sense of belonging in school, global self-worth, psychosocial outcomes, and academic achievement, with a sample of African American, Hispanic, and Asian-American college students. The authors assessed students' ethnic identity beliefs using the Multigroup Ethnic Identity Measure-Revised (MEIM-R; Phinney & Ong, 2007), which is a 6-item scale designed to measure students' *commitment* to their ethnic group membership (i.e., the extent to which students feel a sense of belonging to their ethnic group; e.g., "I feel a strong attachment towards my own ethnic group") and *exploration* of their ethnic group (i.e., what that membership represents; e.g., " I have spent time trying to find out more about my ethnic group, such as its history, traditions, and customs"). The results of this study indicated that students' ethnic identity beliefs were positively associated with both their global self-worth and their academic achievement. Most notably though, the authors also found that in the absence of feeling a sense of belonging in their college, greater identification with their ethnic group shielded students' self-worth from the negative effects of feeling socially isolated at their institutions. Given that Black and Hispanic students college students are more likely than their White peers to be underrepresented within their institutions, they are also more likely to lack a sense of belonging in that context (NCES, 2016; Walton & Cohen, 2004; 2011). Therefore, evidence suggesting that students' racial and ethnic identity beliefs may protect their self-worth in the absence of feeling like they belong at their college or university, has particularly important implications for this population of students.[5] Additionally, Gummadam, Pittman, and Ioffe's (2016) findings also provide a potential mechanism for Chavous and colleagues' (2003) finding, in which they observed the lowest college enrollment rate in the

[5] Underrepresented is operationalized as (a) in relation to the proportion of students of color at an institution and (b) in comparison to the proportion of people of color in the general population.

group of students who had lower overall identification beliefs (i.e., lower centrality, private regard, and public regard beliefs).

Collectively, the findings from this work suggest that the extent to which students from racial and ethnic minority groups centralize their racial or ethnic identity (referred to as *racial identity* from this point forward) within their self-concept may be particularly effective in buffering them from the negative consequences of academically stigmatizing experiences (Chavous et al., 2003; 2008; Gummadam, Pittman, & Ioffe, 2016). Given that *centrality beliefs* also encompass beliefs about students' sense of connectedness and attachment with other members of their racial group, which were also measured in Gummadam, Pittman, & Ioffe (2016), it is possible that students' stronger centrality beliefs positively influence their academic outcomes—in part—by protecting their sense of belonging in academic environments. There is also a basis for arguing that identifying as Black and/or Hispanic includes valuing education, such that stronger racial identification should include striving for high academic achievement (Phinney, 1990; Valencia, 2012). Lastly, there is also evidence that centrality beliefs are negatively associated with maladaptive racial identity beliefs—such as assimilation beliefs, which refer to beliefs regarding the extent to which people of color should stress an American (or White) identity versus a Black or Hispanic identity (e.g., "Blacks should act more like Whites to be successful in this society"; Sellers, Chavous, & Cooke, 1998).

Part III

Implications for Students' Deficit-Oriented Labeling Experiences

When interpreted collectively, the findings presented in this literature review offer a foundation from which to begin addressing several important aspects of college students' deficit-oriented labeling experiences. For instance, there is strong evidence to suggest that Black and Hispanic students may be disproportionately represented within the 'at-risk' college student population (Placier, 1996; Marger, 1991; Valencia, 1997; 2010). There is also evidence to suggest that these types of labels may be communicated to students through their participation in educational intervention programs or academic support services—and that the nature of these contexts may facilitate students' perceptions of these environments as 'threatening'(Castro, 2014; Aronson & Good, 2016). These findings also provide some basis from which to argue that these experiences are both stigmatizing to students and likely associated with adverse consequences to their academic self-perceptions and sense of belonging in college. They also suggest that students who belong to one or more groups that have been historically stigmatized in academic contexts, may experience effects that are disproportionate in nature, relative to students from non-stigmatized groups (Aronson & Steele, 2005). Finally, there is also encouraging evidence with respect to potentially mitigating any negative effects of these experiences through students' motivational beliefs (Aroson, Fried, & Good, 2012; Butler-Barnes et al., 2018; Chavous et al., 2003). In this section, I discuss the evidence presented in the literature review as it pertains to questions surrounding students' deficit-oriented labeling experiences and thoughtfully consider the critical questions that remain unanswered.

Do the 'at-risk' and 'underprepared' labels serve as proxies for low-income and/or first-generation Black and Hispanic students? Historically, the use of these labels in education

originated in the context of racial and ethnic achievement gaps and were used *explicitly* to characterize low-income Black and Hispanic students—a practice that would not be considered socially acceptable under the current sociopolitical climate (Marger, 1991; Pearl, 1961; Valencia, 1997). More recently, in lieu of explicitly targeting students based on race and ethnic group membership, many institutions of higher education commonly rely on indicators of low achievement and/or students' low-income and first-generation status as a basis for determining who is 'at-risk' (Castro, 2014; Valencia, 2010).

What remains unclear is the extent to which these labels are used to characterize Black and Hispanic students from low-income backgrounds, relative to White students from similar backgrounds. Given that the indicators commonly used to identify 'at-risk' students could also serve as proxies for identifying Black and Hispanic students, also suggests that these students will continue to represent a disproportionate percentage of this student population. Using factors that disproportionately affect Black and Hispanic students as indicators for 'at-risk' status, also encourage associations between the 'at-risk' label and students' background characteristics (Castro, 2014; Gray, 2013; Valencia, 1997; 2010). Moreover, in the absence of concrete criteria for categorizing students as *at-risk* or *underprepared*, the evidence also seems to suggest that Black and Hispanic students may be more likely than their White peers to be incorrectly characterized by these labels, simply based on their overrepresentation in that population of students. Moreover, identifying these students would be advantageous in identifying potential mechanisms for mitigating the effects of these experiences. One goal of this dissertation was to establish some preliminary understanding of the types of students who may be disproportionately likely to be characterized by deficit-oriented labels.

Are educational interventions 'threatening' environments for students? Both the anecdotal and empirical evidence suggests that deficit-oriented labels are often used by faculty, staff, and administrators, in the context of educational 'interventions' and/or support services, that serve underachieving college students, or those perceived to be 'at-risk' to underachieve (Castro, 2014; Gray, 2013; Valencia, 2010). Interestingly, there seems to be no evidence to suggest any systemic issues with relation to ill will or malicious intent underlying the use of the *at-risk* or *underprepared* labels in these contexts (Castro, 2014). Rather, the evidence suggests these contexts are designed to be constructive and supportive for students, and that the individuals who work with these students have a genuine interest in helping them succeed (Castro, 2014). That said, the manner in which students are framed in these contexts, lend themselves to being interpreted as deficit-oriented (Valencia, 1997; 2010).

What is unknown, is the extent to which this is problematic. For instance, it is possible that communicating deficit-oriented labels *directly* to students in those contexts may encourage students' perceptions of those environments as 'threatening'. Inzlitch & Good (2006) describe 'threatening environments' as "settings where people come to suspect that they could be devalued, stigmatized, or discriminated against because of a particular social identity" (p. 145). It may seem counterintuitive to suggest that students might perceive well-intentioned offers of academic support as 'threatening'; however, there is evidence that students sometimes perceive unsolicited academic help from teachers as a low-ability cue (Graham, 1984). Therefore, if students were to interpret low-ability evaluations as the result of prejudiced beliefs, this might encourage perceptions of that context as being 'threatening'. For instance, for Black and Hispanic students—who have a history of being academically stigmatized—being labeled by a descriptor that also communicates a negative message regarding their academic competence

trigger feelings of devaluation and discrimination, even in the absence of the intervention. Conversely, in the absence of the label, offering these students unsolicited academic support may also be sufficient to trigger these feelings. However, my intuition is, that in combination, both the label and the offer of unsolicited support may reinforce students' interpretations of that context as an endorsement of negative stereotypes and breed feelings of prejudice and discrimination to an even greater extent. Although White students from low-income backgrounds and/or those who are first-generation students may also perceive these contexts as 'threatening', I suspect that the strength of these perceptions would vary between students who belong to one academically stigmatized group versus those who belong to multiple academically stigmatized groups (Castro, 2014; Croizet & Clair, 1998; Gray, 2013).

Are there other contexts in which these labels may be communicated to students? It is also unclear whether there are other contexts within their university or college, in which these labels might be communicated to students. Given that students also work closely with instructors in class and with other staff, such as academic advisors, it is important to identify all of the contexts in which students might be vulnerable to these stigmatizing experiences. This dissertation also sought clarity with respect to this question.

Are there consequences associated with labeling students as 'at-risk'? The findings from the research examining the consequences of labeling students in achievement contexts, indicate these experiences adversely impact students through various processes of stigmatization (e.g., Inzlicht & Good, 2006; Walton & Cohen, 2004). The broader research on academically stigmatizing experiences also indicates that college students who belong to academically stigmatized groups, may be disproportionately susceptible to experiencing diminished academic self-perceptions and feelings of belonging in college. These students are also more likely to

experience stereotype threat, which can impede their performance in achievement contexts and have adverse, long-term effects on their self-perceptions and sense of belonging in college (e.g., Aronson & Steele, 2005).

What is unknown, is the extent to which (a) students' deficit-oriented labeling experiences would elicit the same consequences associated with other labels used in achievement contexts (e.g., *learning disabled*), and (b) whether these consequences would be experienced universally by all students to the same extent, or if they might vary based on students' background characteristics. With respect to the first point, there is evidence to suggest that the consequences of students' deficit-oriented labeling experiences would be similar to those associated with other academically stigmatizing experiences, because the processes that underlie those consequences center around inferences about the students' academic competence. For instance, similarly to students who are labeled as *LD*, students who are labeled as *at-risk* or *underprepared* in the context of participating in a support program may be subjected to biased perceptions of their competency to a greater extent than students with similar achievement records who are *not* characterized by these descriptors (and thus, not targeted by these programs). If these biased perceptions influence the manner in which faculty, staff, and/or administrators treat students (which includes communicating the labels directly to them)—which the evidence suggests is possible—then these students might also experience diminished self-perceptions of their academic competence as a result of internalizing those characterizations. Alternatively, even if a faculty, staff, or instructor's low academic perception about a student are an accurate representation of the student's current competency level, they may allow these perceptions to influence their behavior towards the student in a manner that the student perceives as 'threatening'. As a result, rather than internalizing these negative perceptions—some students

may interpret them as a result of prejudiced beliefs, which could foster a sense of distrust and drive students to discount beneficial feedback and information provided to them in that context—again, ultimately undermining those efforts to help them (Inzlicht & Good, 2006).

With respect to the second point, it remains unclear the extent to which background characteristics might influence the types and severity of consequences students experience as a result of being labeled by a deficit-oriented descriptor. However, my intuition is that the extent to which students from different backgrounds experience the negative consequences of being labeled by deficit-oriented descriptors, would likely vary—at least in part—in terms of the extent to which students are vulnerable to negative stereotypes about one or more groups to which they belong. For instance, I would expect White students from low-income backgrounds or those who are first-generation college students to experience the consequences of these labeling experiences to a greater extent, than White students who are neither from low-income families or first-generation college students, given that students from both low-income backgrounds and first-generation students are often stigmatized in achievement contexts, and therefore more likely to be increasingly vigilant to perceptions of those stereotypes (Aronson & Steele, 2005; Croizet & Clair, 1998; Gray, 2013). However, I would expect Black or Hispanic students who are first-generation college students or from low-income backgrounds to experience the effects of these labeling experiences to a greater extent than the others, given that they associate with multiple identities that are pervasively stigmatized in academic contexts.

Further, I suspect that students' background characteristics might be particularly salient and influential with respect to potential stereotype threat effects. More specifically, it is possible that being labeled in a deficit-oriented manner may be sufficient to trigger stereotype threat effects, simply based on the negative information these labels convey about students' academic

potential. Given that students need not belong to a stigmatized group in order to experience the effects of stereotype threat, the *acute* effects of stereotype threat might extend to all students who endure these labeling experiences, simply based on the negative stereotypes associated with students who are labeled with descriptors such as 'at-risk' or 'underprepared' (Gray, 2013). However, low-income students and/or first-generation students—particularly those who are also Black and Hispanic—are disproportionately more likely to experience *chronic* stereotype vulnerability, which could lead to frequent experiences of impeded academic performance that can result in long-term effects on students' academic self-perceptions and sense of belonging in college that are cyclical in nature. For example, chronic experiences of poor performance can diminish students' self-perceptions about their academic ability and sense of belonging in college—both of which could also influence performance through their influence on other factors like academic engagement and academic help-seeking behaviors (e.g., Winograd & Rust, 2014).

Overall, the evidence suggests that at the very least, students' deficit-oriented labeling experience may impact students' self-perceptions and sense of belonging in college, and even their academic performance—all of which can negatively influence their motivation and academic persistence in college. Moreover, students who stand to benefit the most from high academic self-perceptions and a strong sense of belonging in college, may be particularly likely to endure stigmatizing experiences associated with consequences to both outcomes. Given that these outcomes are often specifically targeted in support programs, the effects of students' deficit-oriented experiences may potentially undermine institutional equity efforts. Thus, a primary aim of this dissertation was to develop a better understanding of the types and severity of consequences associated with these labeling experiences.

Can students' motivational beliefs mitigate the potential effects of these stigmatizing experiences? Despite the troublesome implications of students' deficit-oriented labeling experiences, there is encouraging evidence to suggest that certain beliefs systems may shield students from the negative consequences of these experiences. For instance, an intervention designed to enhance African American students' view of intelligence as being malleable was found to mitigate the effects of stereotype threat on their academic achievement. The evidence also indicates that their perceptions of their race and/or ethnicity as central to their self-concept can buffer them from the negative effects of potentially stigmatizing experiences. Moreover, students' stereotype vulnerability may also moderate the consequences of these labeling experiences, both with respect to stereotype threat effects as well as other consequences (Aronson & Steele, 2005; Inzlitch & Good, 2006). These findings are particularly promising because they seem to suggest that stronger racial identity beliefs disproportionately benefit the types of students who are most likely to experience stigmatization, by mitigating the negative consequences associated with those experiences.

However, there is no direct evidence regarding the extent to which these beliefs may buffer students from the potential consequences of being labeled as 'at-risk' or other deficit-oriented labels. Specifically, given that most of the work on college students' racial and ethnic identity beliefs is correlational, it is unknown if these beliefs are a cause or effect of students' resiliency to stigmatization. Moreover, considering that the negative academic stereotypes associated with Black and Hispanic students relate precisely to students' race and ethnic membership, it remains to be seen if a strong identification with one or the other, will buffer or accentuate the effects of stigmatizing experiences that draw directly from those negative academic stereotypes. For instance, it is possible that a stronger identification with race or ethnic

membership may heighten Black and Hispanic students' vigilance to prejudice and make them more vulnerable to experiences of stereotype threat. However, there is also a basis for arguing that stronger centrality beliefs may protect Black and Hispanic students by making them more resilient to experiencing diminished feelings of belonging in college. Moreover, the current research assessed students' racial identity beliefs in relation to their identities as *college students*, which I argue should further strengthen any associations between racial identification and valuing of education. Enhancing our understanding of the manner with which Black and/or Hispanic students' racial identification may influence their interpretations and responses to deficit-oriented labeling experiences was an important aim of this dissertation.

Current Research

The findings presented throughout this literature review provide critical insights with respect to the use of deficit-oriented labels in higher education, and the potential consequences that might stem from communicating these characterizations directly to students. However, they also highlight several important gaps of knowledge that limit our present understanding of the consequences potentially associated with these stigmatizing experiences—and in broader terms, how they might influence students' overall success in college. Addressing these gaps in the literature is important and warrant further investigation because it may help unearth the underlying mechanisms driving the disparities in students' motivation and academic persistence that—at least in part—continue to perpetuate the poor retention and graduation rates of low-income and/or first-generation Black and Hispanic college students.

That said, the goal of the current research was to address several of these voids through empirical investigation. A total of three online studies were conducted for this dissertation—each designed to enhance our understanding of college students' deficit-oriented labeling experience

by testing a particular set of research questions. Study 1 was conducted as a preliminary investigation that explored research questions related to descriptive aspects of students' deficit-oriented labeling experiences—including the frequency of these experiences, the contexts in which they occurred, and the effects students experienced as a result. Using vignettes about hypothetical students and their advisors, this study also examined research questions related to students' interpretations of a context in which a hypothetical student was being offering unsolicited academic support and also characterized by a *deficit-oriented* label or a *neutral* label. Study 2 employed refined methods (developed as a result of the findings from the first study) to continue exploring research questions related to the frequency and contexts in which students experience being labeled by specific deficit-oriented descriptors, and the effects they experienced as a result. This study also explored associations between students' *academic mindsets*, their *stereotype vulnerability*, their *racial identity beliefs*, and the extent to which students reported experiencing motivational and affective consequences as a result of their deficit-oriented labeling experiences. Study 3 used hypothetical scenarios very similar to those in Study 1, to continue examining research questions related to students' interpretations of contexts in which a student was being offering unsolicited academic support while being characterized by a *deficit-oriented* or *neutral* label. However, this study expanded on Studies 1 and 2 by examining the moderating influence of students' academic mindsets and their racial and ethnic identity beliefs on their interpretations of these hypothetical contexts, and by exploring potential stereotype threat effects associated with students' deficit-oriented labeling experiences.

A note about sample groups. For each of the three studies conducted for this dissertation, the final sample was limited to undergraduate students who were categorized into one of two sample subgroups based on their background characteristics. The first subgroup

consisted of Black and/or Hispanic students from low-income backgrounds, who were also first-generation college students (LIFG students, for short), whereas the second subgroup consisted of White students who were neither from low-income backgrounds or first-generation college students (NLIFG students, for short). My decision to limit the focus of this research on these specific subgroups of students was based on several factors. First, as is the case with most research, I had limited funds to work with. Because I conducted three studies overall, this limited the number of participants I could recruit for each study. Moreover, because I incorporated several experimental manipulations into two of the three studies, including more than two subgroups would have severely impacted sample sizes for each subgroup. Two, given my limitation of funds and the lack of prior research on this topic, I wanted to address my research questions using a group of students—which the evidence indicated—would be *most* likely to have experienced being labeled as 'at-risk' or 'underprepared' (i.e., LIFG students), and contrast their experiences with those of students who would be *least* likely to have had these experiences (i.e., NLIFG students). More specifically, I felt that recruiting one group of students associated with multiple academically stigmatized identities and comparing their experiences to another group of students—who from the outset at least—were not associated with any academically stigmatized identities, would be the most efficient route to identifying the potential consequences associated with these labeling experiences. Further, I limited the LIFG sample group to students who identified as Black and/or Hispanic for two reasons. One, Black and/or Hispanic students represent a disproportionate percentage of Black and Hispanic students currently enrolled in college (NCES, 2017), and two, the negative academic stereotypes associated with these students

do not apply to all students of color, and in fact, some students of color are associated with academic stereotypes that are at the opposite end of the spectrum (e.g., Asian students).[6]

In the Chapters 3, 4, & 5, I discuss each of the studies conducted for this dissertation in detail, including the methodology used in each study, the results of this work, and a discussion of the findings.

[6] Race and ethnicity were not mutually exclusive, such that participants who identified as Black but also as White and/or with other groups, would also be considered LIFG students. Similarly, students who identified as Hispanic were categorized as LIFG students, regardless of which other racial/ethnic group they reported identifying with.

CHAPTER 3: STUDY 1

Study 1 served as a preliminary exploration of LIFG and NLIFG students' deficit-oriented labeling experiences and was conducted as an online survey administered through Qualtrics. This study investigated students' deficit-oriented labeling experiences in two ways. One portion of the study focused on research questions related to *descriptive* aspects of students' deficit-oriented labeling experiences, such as the frequency and context in which these experiences occurred, as well as their affective and motivational consequences. The research questions related to these aspects of students' labeling experiences were addressed by collecting both qualitative and quantitative data. First, participants were asked to describe specific details about their deficit-oriented labeling experiences through a series of open-ended prompts. Considering the lack of existing research on this topic, I wanted to avoid being overly presumptuous and provide students themselves with the opportunity to detail their labeling experiences in their own words. However, given that I had little control over the quantity and quality of data yielded through open-ended items, I also incorporated a set of closed-ended items designed to assess the same elements of students' labeling experiences that were probed by the open-ended prompts, but with respect to four specific deficit-oriented labels (i.e., at-risk, underprepared, disadvantaged, and underrepresented).

My intuition was that LIFG students would report being labeled by deficit-oriented labels more frequently than NLIFG students. I also expected that relative to NLIFG students, LIFG students would also report experiencing negative affect as a result of these experiences to a greater extent—particularly because being characterized by a deficit-oriented label might increase the saliency of the negative academic stereotypes associated with one or more of their academically stigmatized identities.

Another portion of this study used hypothetical scenarios and closed-ended items to examine research questions related to participants' interpretations of a context in which a student was provided feedback from their advisor on their poor academic performance that semester. For this section of the study, I wanted to examine participants' interpretations of a deficit-oriented label versus a more neutral label, so participants were told that the advisor characterized the student as either 'at-risk' or 'first-generation' while providing them with feedback. Hypothetical scenarios were used in lieu of drawing on students' actual labeling experiences in order to account for students who had not experienced being labeled by a deficit-oriented descriptor in the past. After reading the scenario, participants responded to items that examined the extent to which they perceived the advisor's feedback (including the characterization of the student) as communicating *positive* or *negative beliefs* about the hypothetical student and the extent to which they believed the advisor was *intentionally* communicating their positive or negative beliefs about the student through their feedback. In addition to examining these interpretations, I also explored participants' perceptions regarding the extent to which they expected the hypothetical student in the scenario to experience a series of affective and motivational consequences as a result of their interaction with the advisor.

My decision to contrast the *at-risk* label with the 'first-generation' label was a result of wanting to test a potential alternative for the 'at-risk' label that could potentially apply to a large proportion of the 'at-risk' college student population. Given that there is sufficient evidence to suggest that status as a first-generation college student is an academically stigmatized identity within higher education, one might question my decision to use that label as a 'neutral' option. However, this decision was based on the fact that the extent to which the 'first-generation' label is interpreted as concrete and objective depends less on situational factors and context, whereas

the 'at-risk' label can be vague and ambiguous when used without additional context. Moreover, the inherently negative connotations associated with the 'at-risk' label leave little room for interpreting this descriptor in a positive manner, whereas the underlying tone of the 'first-generation' label is more open to interpretation. In fact, for many students, bearing the first-generation label represents a source of pride and accomplishment for being the first in their families to go to college—another element that makes this descriptor uniquely different from the 'at-risk' label (e.g., Donovan & Johnson, 2005).

Concerning this section of the study, I had several hypotheses regarding participants' interpretations of the *at-risk* v. the *first-generation* label, and the types of consequences they believed the hypothetical student might experience as a result of receiving feedback that characterized them as one or the other. First, given the inherently negative nature of the 'at-risk' label, I expected that both LIFG and NLIFG students would interpret the that label more negatively than the 'first-generation' label, and that they would indicate that they expected the student labeled as 'at-risk' to experience negative affective and motivational consequences to a greater extent than the student labeled as a 'first-generation' student. However, because I hypothesized that (compared to NLIFG students) LIFG students would report being more frequently labeled by deficit-oriented descriptors, my intuition was that reading the hypothetical scenario might easily activate memories of their own experiences and influence their responses as a result. Thus, I also expected that, relative to NLIFG students, LIFG students would interpret the 'first-generation student' label as communicating less negative beliefs and as resulting in negative affective and motivational consequences to a lesser extent. And, I expected to find similar between-group differences for the 'at-risk' label, but in the opposite direction.

In the remainder of this chapter, I describe the methods used to collect the data for this study, including specific details regarding the recruitment of participants, the survey (including all prompts, items, and measures), and the procedure participants followed to participate in the study. Following this, I describe the analyses conducted on the data and the results yielded from those analyses. Finally, I end with a summary of the findings, the conclusions drawn from these findings, and their implications for Studies 2 and 3.

Method

Inclusion/Exclusion Criteria

Prospective participants were deemed eligible to participate in this study only if they met the criteria required to be categorized into the LIFG or NLIFG samples. In order to determine their eligibility, prospective participants were required to complete a brief 11-item pre-screening questionnaire. Their responses to seven of those items were used to determine whether they were eligible to be categorized into one of the two sub-samples. The remaining four questions were demographic questions that were included in the pre-screening process in order to prevent prospective participants from guessing the specific criteria being used to determine their eligibility for the study. The pre-screening questions of primary interest are listed below, followed by the response criteria that was used to categorize participants into the sub-samples (see Appendix B for full list of pre-screening questions and response options).

a. *Are you currently a full-time student at a 4-year university or college?*
b. *Are you currently eligible for the Federal Pell Grant?*
c. *Are you currently eligible for Federal Work Study?*
d. *What is the highest level of education completed by your mother or female guardian?*
e. *What is the highest level of education completed by your father or male guardian?*

f. Are you Hispanic or Latino (i.e., a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race)?

g. Please indicate your racial background (select all that apply):

Criteria for LIFG sample. Prospective participants were required to meet the following criteria to be categorized as a LIFG student for this study: *1)* Questions (a), (b), and (c): Choose "yes", *2)* Questions (d) and (e): Choose: High School *or* GED or 2-year college *or* vocational degree, *and 3)* Questions (f) and (g): Choose "yes" for (f) *or* Black or African American for (g).[7]

Criteria for NLIFG sample. Prospective participants were required to meet the following criteria to be categorized as an NLIFG student for this study: *1)* Question (a): Choose "yes", *2)* Questions (b), and (c): Choose "no", *3)* Questions (d) and (e): Choose: Bachelor's degree (e.g., BS, BA, AB) *or* Master's degree (e.g., MS, MA, MBA) *or* Professional degree (e.g., JD, LLB, MD, DDS, DVM) *or* Doctorate (e.g., PhD, DSc, EdD), *4)* Question (f): Choose "no", and *5)* Question (g): Choose ONLY "White".

Participants

The final sample of the study was $N = 119$ ($n = 59$ LIFG; $n = 60$ NLIFG) and consisted of students between the ages of 18-29 ($M = 22.01$, $SD = 2.28$), of which 57% were male. The majority of students were upperclassmen (74%), majoring in STEM fields (38%), social sciences (27%), the humanities and education (9% each), as well as the creative arts (6%). Sixty-nine percent of students indicated that they attended a public university or college and approximately 19% indicated that their institution was religiously affiliated. Sixty-eight percent of the LIFG sub-sample identified as Hispanic or Latino and 46% identified as solely or partly Black or

[7] Participants who indicated that they identify with more than one race—including *Black or African American*—were categorized as Black for the purposes of this study.

African American.[8] One hundred percent of the participants in the NLIFG sub-sample identified as non-Hispanic White (see *Table 3.1* for additional demographics by sample groups). Lastly, an independent *t*-test examining participants' responses to the subjective SES question on the demographics questionnaire (i.e., "How financially well off were you growing up?") indicated that on average, LIFG participants perceived themselves as being significantly less "well off" (M= 2.14; SD= .73) than NLIFG students (M= 3.60, SD= .72), $t(117) = 11.03$, $p < .001$, $d = 2.02$ (see *Table 3.1* on pg. 187 for additional demographic information by sample groups).

Table 3.1. Summary of *Study 1* demographic information for the full sample and by sample groups.

Sample	Age (*M*)	Sex *(Male)*	Hispanic	Black	White	Upperclassmen	Public Institution	Religious Affiliation
Full Sample (*n*= 119)	22.02	57.1	33.6	22.7	68.9	73.9	68.9	18.5
LIFG (*n*= 59)	22.20	49.2	67.8	45.8	37.3	69.5	81.4	20.3
NLIFG (*n*= 60)	21.83	65	0	0	100	78.4	56.7	16.7

Note: All numbers in columns 3-9 are expressed in percentages.

Exclusion of cases from initial sample. A total of 1,455 Mechanical Turk workers consented to participate in this study, indicated they were full-time college students at a 4-year college or university, and completed the pre-screening questionnaire, from which 168 met the inclusion criteria for one of the two sample groups and completed some portion of the study. Of those, twenty-four participants were excluded from the final sample because they indicated they were over the age of 29. This age cutoff was employed to try and ensure that the participants in

[8] Neither race nor ethnicity were mutually exclusive.

the final sample were relatively representative of the "typical" LIFG or NLIFG college student. An additional 25 participants were also removed from the final sample because their responses to the open-ended questions suggested they had not taken the study seriously (e.g., responses were uninterpretable; $n = 7$), or because their response to the subjective SES item (e.g., "How financially well off were you growing up?") was not consistent with their status as a low-income college student *or* non-low-income college student ($n = 9$ LIFG students; $n = 9$ NLIFG students).[9]

Recruitment

Participants were recruited for this online study via Mechanical Turk and redirected to Qualtrics Research Suite to complete the Consent Form, pre-screening questionnaire—and if deemed eligible—the study survey. Prior to being redirected to the consent form on Qualtrics, prospective participants were able to view general information about the study on Mechanical Turk—which included information about the estimated time for completion of the study, as well as the consent and pre-screening process.

Compensation

Participants were compensated with $1.50 (~ $.0.15/per minute) for completing the study. Compensation was only granted to participants who completed the study in its entirety. In order to receive their compensation, participants were provided with a completion code on the last page of the study survey—which they used to redeem their compensation on Mechanical Turk's website.

[9] Income status was considered inconsistent for LIFG students if they indicated they were "extremely well-off" or "well-off" while they were growing up. For NLIFG students, income status was considered inconsistent if they indicated that their families were "not very well off" or "poor".

Materials & Measures

Section 1: Open-ended prompts about labeling experiences. This study utilized one general open-ended prompt that asked participants to think about—and list—up to five negative descriptors used to characterize their academic potential in the past. Instead of providing specific examples about the types of descriptors I was interested in yielding, I included examples of descriptors I was *not* interested in "*general adjectives commonly used to describe students based on personal attributes—such as "motivated", "lazy", or "dumb*". Once participants clicked the '*continue*' button to move on, they were presented with a list of the labels they had provided in their response to the previous prompt, and two additional prompts that asked them to *describe* '*the people who have used these labels to describe you*' and '*the contexts in which the labels were communicated to you'*. Once participants clicked the '*continue*' button to move on, they were again presented with the list of labels they provided in the first prompt, and one final prompt which asked them to describe how '*these experiences of being categorized by these labels made them feel'* (see Appendix C for full wording of all prompts).

Section 2: Close-ended items assessing labeling experiences

***Frequency of labeling experiences*.** The frequency with which participants experienced being categorized as an *at-risk student* [*underprepared; disadvantaged; underrepresented*] was measured using the item "How often (if ever) have you heard yourself or other students like you described as an *at-risk* [*underprepared; disadvantaged; underrepresented*] student?" Participants rated the frequency of their labeling experiences using the following 5-point Likert-type scale: 0= "Never"; 1= "Rarely"; 2= "Occasionally"; 3= "Sometimes"; 4= "Often". Participants who answered all three frequency items with *"Never"*, were <u>not</u> presented with the subsequent items

that measured affective and potential motivational consequences of labeling; instead, they were redirected to the *Section 2* of the study.

Context of labeling experience. The closed-ended measures of the context in which participants experienced being labeled by each of the four *deficit-oriented* labels were assessed using two closed-ended items. For each label, one item asked participants about the individual who used the label to characterize them: "Who referred to you or students like you with the label *at-risk* [*underprepared; disadvantaged; underrepresented*] student?" Participants were presented with eight response categories (e.g., academic advisor; instructor; university or college staff or administrators; peers/classmates) and asked to check *all* that applied. Another item was used to assess the specific context these experiences occurred in: "Where have heard yourself or other students like you described as an 'at-risk student'?" Participants were presented with eight response categories (e.g., event; academic advising session; academic support services; in class) and asked to check *all* that applied.

Affective consequences of labeling experiences. The affective consequences of participants' experiences of being labeled were measured using one item for each of the four *deficit-oriented* labels (e.g., "How negatively did it make you feel to hear yourself or other students like you described as an *at-risk* [*underprepared; disadvantaged; underrepresented*] student?"). Participants were asked to indicate the extent to which they experienced negative affect using a 6-point Likert-type scale (1= "Not at all negative" to 6= "Extremely negative"; or "Not Applicable").

Section 3: Hypothetical scenarios. In the third section of this study, participants were randomly assigned to one of two conditions—a *deficit-oriented label* condition or a *neutral label* condition and presented with one vignette that described a hypothetical scenario between a low-

income, first-generation student and their academic advisor. The scenario described a meeting between the advisor and a low-income, first-generation student. In each scenario, the advisor provides feedback to the student regarding his [her] poor academic performance that semester and characterizes the student using a *deficit-oriented* (i.e., *at-risk*) or *neutral label* (i.e., *first-generation* student*)*. The label used by the advisor was the only aspect of the scenario that varied across conditions (see Appendix C for full wording of both scenarios). The sex of the hypothetical student in the scenario was matched to participants' sex based on their response to the item on the pre-screening questionnaire that asked them to indicate their 'biological sex' (i.e., male or female), such that participants who indicated they were male, read a scenario about a student named "Aaron", and participants who indicated they were female, read a scenario about a student named "April".

Valence of beliefs item. Participants' perceptions about the extent to which the advisor's feedback communicated positive or negative beliefs about the student was measured with one *valence belief* item (i.e., "Please indicate the extent to which you believe that Aaron's [April's] advisor is communicating positive or negative beliefs about Aaron [April] with the statement below:…"), which included a 6-point Likert type scale, from 1 = "Very Negative Beliefs" to 6 = "Very Positive Beliefs". This item was reverse coded for subsequent analyses, such that higher values reflected participants' perceptions of the advisor communicating a greater degree of negative beliefs.

Intentionality of communicating beliefs items. One item (i.e., *intentionality*) was used to assess participants' perceptions about the extent to which the hypothetical student's advisor was *intentionally* using their feedback to communicate their positive or negative beliefs about the hypothetical student 'To what extent they believed that the hypothetical student's advisor was

intentionally communicating their positive or negative beliefs about the student with the feedback they provided them". This item initially designed to be used with a 6-point Likert type scale from 1 = "Not at All Intentionally" to 5 = "Very Intentionally", but was instead inadvertently presented with the same 6-point Likert type scale used for the *valence beliefs* item (i.e., 1= "Very Negative Beliefs" to 6= "Very Positive Beliefs"). To better gauge whether the issue with the response scale had led participants to interpret this item in the same manner as the *valence beliefs* item—particularly given that they were presented in consecutive order—I conducted two correlational analyses. The results of these tests indicated that the correlations between the *valence beliefs* item and *intentionality* item for both label conditions were strongly and positively associated ($r(56) = .56$, $p < .001$ for the *neutral label condition*; $r(59) = .68$, $p < .001$ for the *deficit-oriented label condition)*, which suggests that any results yielded through analyses that include these labels should be interpreted with caution.

Perceptions of the consequences associated with being labeled by a deficit-oriented or neutral label. Participants' perceptions of the affective and motivational consequences that the hypothetical student described in each scenario might experience as a result of their labeling experience were measured using seven items (e.g., "The advisor's feedback probably made Aaron [April] feel like doing well in school isn't really that important."; "The advisor's comments probably made Aaron feel like he can overcome his academic challenges if he works hard."; see Appendix C for full list of items). Students' were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale (1= "Strongly Disagree" to 6= "Strongly Agree"). *Item 5* was reverse coded to align with the remaining six items, because it was originally worded in positive terms ("The advisor's comments probably made Aaron feel like he can overcome his academic challenges if he works hard"). A new *aggregate effects*

variable was computed by computing the average of participants' responses to all seven items. This variable was used as a dependent measure in subsequent analyses of this data, in lieu of using each of the seven items individually (α= .85).[10]

Procedure

This study was administered as an online study via Mechanical Turk and Qualtrics Research Suite, and participants were able to access and complete it from any desktop or laptop computer. At the time that prospective participants indicated they are interested in completing the study, they were redirected to Qualtrics, where they were presented with an electronic version of the consent form for this study. Prospective participants were required to give their consent to participate (by checking the appropriate box under the consent form), before they were allowed to complete the pre-screening questionnaire. Those who indicated that they did *not* give their consent to participate were redirected out of the study. Next, prospective participants completed the 11-item pre-screening questionnaire. Those who did not meet the criteria to be categorized into either sub-sample were redirected out of the study, and those who were eligible to be categorized into one of the two sample groups were presented with a message informing them of their eligibility and allowed to continue on to the study survey. Participants completed *Sections 1, 2,* and *3* in order—followed by the Demographics Questionnaire (see Appendix D for Demographics Questionnaire). Lastly, participants were presented with an electronic version of the debriefing form for this study (see Appendix E for copy of de-briefing form), which included the completion code that used to redeem their compensation on Mechanical Turk.

[10] Preliminary analyses were conducted on each individual item, but given that the results were similar across all seven items the aggregate analysis is reported in lieu of these.

Results

The data yielded in this study was analyzed using both qualitative and quantitative methods. The sections below include descriptions of the analytical techniques used for specific items and the results associated with each analysis.

Open-Ended Descriptions of Labeling Experiences

In this study, participants were asked to describe an experience in which they were categorized by a deficit-oriented label in terms of when the experience took place, the label that was used, the context (e.g., situation) in which the experience occurred, and the effects they experienced as a result of being labeled. Students' responses to these prompts were coded and then evaluated through descriptive and statistical analyses. The following subsections describe the coding process, the subsequent analyses, and the results yielded from those analyses.

Coding and analyses of participants' open-ended responses.

Deficit-oriented labels. Participants' responses to the open-ended prompt asking them to list up to five *deficit-oriented* labels used to characterize them in the past, were used to examine (a) between-group differences in the percentage of participants that reported being labeled by at least *one deficit-oriented label* in the past and (b) between-group differences in the total number of deficit-oriented labels students reported being categorized by in the past. First, I converted participants' qualitative responses into quantitative data by coding each label provided by participants as either a *deficit-oriented label=1* or *not applicable=0*. Labels were categorized as *deficit-oriented* if they contained descriptors that (a) implied the student was considered a *minority* at their institution (e.g., underrepresented), or (b) implied the student was *lacking in privilege or academic preparedness* (e.g., disadvantaged; underserved; underprepared, ill-prepared, underachiever), or (c) implied the student's likelihood of failing in college was greater

than the norm (e.g., at-risk; high-risk). For participants who listed 'unprepared', this was not considered synonymous to 'underprepared' unless participants noted chronic factors that affected their preparedness. All other labels and responses that did not meet these criteria were coded as *not applicable*. Using this data, I then created two new variables—one which reflected the total sum of deficit-oriented labels provided by each participant, and a categorical variable that indicated whether a participant had reported at least one deficit-oriented label and was coded as *yes=1* and *no=0.*

A descriptive analysis of this data showed that 68.9 percent of participants (n= 82) indicated they had been labeled by at least one descriptor that met the criteria to be considered 'deficit-oriented'. Using the categorical variable, I conducted a Chi-Square test of independence to examine between-group differences in the proportion of students in each sub-sample who reported being categorized by at least one deficit-oriented label. The results of this analysis indicated that the difference between the proportion of LIFG and NLIFG students who reported being categorized by at least one *deficit-oriented* label (n= 46 or 78% for LIFG v. n= 36 or 60% for NLIFG) was significant, $\chi^2(1, N= 119) = 4.48$, p = .034. Next, I conducted a Poisson regression to examine sample group differences in the total number of deficit-oriented variables provided by participants. Poisson regression was used because it is appropriate when modeling count data. The results indicated that the mean number of deficit-oriented labels provided by LIFG students (M = 1.86, SD = 1.40) was marginally higher than the mean total for NLIFG students (M = 1.33, SD = 1.51), B = .34 (95% CI, –.01 to .68), p = .055.

The context and "the labeler." Participants' reports of the individual who labeled them and the context in which this experience occurred were examined through their responses to two open-ended prompts. For the context prompt, participants' responses were recoded into a new

context recoded variable by classifying each response into one of seven categories based on the participants' description (i.e., campus event; faculty meeting; advising session; class; high school; home; online). For the "labeler" prompt, participants' responses were recoded into a new *labeler recoded* variable by classifying each response into one of five categories based on the participants' descriptions (i.e., advisor; staff/administrator; instructor; peers; parent/siblings), which were also developed based on a review of students' responses. As *Tables 3.2* and *3.3 show*, participants who indicated they were labeled by at least one descriptor that met the criteria to be categorized as 'deficit-oriented', also indicated that those labels were most frequently communicated to them by instructors or advisors in the context of class, during one-on-one meetings with instructors, or advising sessions.[11] A series of uncorrected Chi-Square tests were conducted to examine sample group differences for the various context and labeler categories. These test indicated that, relative to NLIFG students, a significantly larger proportion of LIFG students reported being labeled by their advisors and staff/administrators, whereas there were no significant sample group differences in contexts.

Table 3.2. Summary of participants' open-ended responses related to the *contexts* in which their deficit-oriented labeling experience occurred by sample group (expressed in percentages).

Context (N= 82)	Campus Event	Faculty Meeting	Advising Session	Class	High School	Home	Online
LIFG	10.9	26.1	37.0	41.3	8.7	2.2	19.6
NLIFG	5.6	22.2	22.2	27.8	13.9	2.8	19.4

Note. Sample size for LIFG n= 46 and n= 36 for NLIFG.

[11] The percentages associated with each category represent the proportion of participants who indicated they experienced being categorized in that context or by individuals in the specified roles. Participants often indicated more than one context and/or individual.

Table 3.3. Summary of participants' open-ended responses related to the *individual* involved in their deficit-oriented labeling experience by sample group (expressed in percentages).

Labeler (*N*= 82)	Advisor	Instructor	Staff/Administrator	Peers	Parents/Siblings
LIFG	45.7[b]	47.8	30.4[a]	34.8	4.3
NLIFG	25.0	50.0	5.6	19.4	8.3

Note. Sample size for LIFG *n*= 46 and *n*= 36 for NLIFG.
[a] denotes a statistically significant sample group difference in proportion (i.e., $p < .05$).
[b] denotes a marginally significant sample group difference in proportion (i.e., $p < .10$).

Effects of a deficit-oriented labeling experience. The self-reported affective and motivational consequences of students' labeling experiences, were investigated by analyzing their responses to the prompt that asked them to describe the way these deficit-oriented labeling experiences made them feel.

First, I created four new variables, *positive affect, negative affect, motivational increase,* and *motivational decrease*, and coded participants' responses across each of the four categories quantitatively (*yes = 1; no = 0*), such that each response had a code for whether the participant indicated they experienced positive affect, negative affect, motivational decreases, and motivational benefits as a result of that experience. In cases where participants' responses indicated they experienced both *positive* and *negative affect (and/or motivational decreases* and *benefits)*, they received a score of *1* across both categories. One participant's response was not related to the consequences they experienced as a result of being labeled and was categorized as *not applicable* and removed from the Chi-Square analyses reported next.

Next, I conducted a series of Chi-Square tests to probe for sample group differences in the proportion of LIFG and NLIFG students who reported experiencing *negative affect*, *motivational decreases*, and *motivational benefits* as a result of their labeling experience. Participants who did not list any deficit-oriented labels in their response to the first prompt in this

section were excluded from the analyses. The results indicated that there was no significant difference in the proportion of NLIFG students and LIFG students that reported experiencing *negative affect* (82.9% *v.* 89.1%, respectively), $\chi^2(1, N= 81) = .67$, $p = .414$, *motivational decreases* (32.6% *v.* 37.1%, respectively), $\chi^2(1, N= 81) = .18$, $p = .671$, or *motivational benefits* (21.7% *v.* 20%, respectively), $\chi^2(1, N= 81) = .04$, $p = .849$. Positive affect was not included in the analysis because there were no participants who indicated they experienced this subsequent to their labeling experience.

Frequency, Contexts, & Affective Consequences of Deficit-Oriented Labeling Experiences

Frequency. The frequency of participants' deficit-oriented labeling experiences was explored through their responses to closed-ended items that asked them to indicate the extent to which they had been categorized by four deficit-oriented labels in the past (i.e., *at-risk; underprepared; disadvantaged; underrepresented*). In addition to using participants' responses to these items individually as dependent measures, they were also used to compute a composite variable that reflected the average frequency with which participants reported being categorized across the four labels. This aggregate was also used as a dependent variable in some analyses. Using participants' responses to the individual items for each label, I conducted a series of independent samples *t*-tests to examine sample group differences in the frequency of LIFG and NLIFG students' labeling experiences. These analyses revealed statistically significant sample group differences for three of the four labels, such that LIFG students reported being labeled as *at-risk* ($M = 2.83$, $SD = 1.26$), *disadvantaged* ($M = 2.90$, $SD = 1.19$), and *underrepresented* ($M = 2.92$, $SD = 1.22$) more frequently than NLIFG students ($M = 2.30$, $SD = .98$; $M = 2.27$, $SD = 1.07$; $M = 2.10$, $SD = 1.10$, respectively), $t(117)s > 2.56$, $ps \leq .012$, $ds > .46$.[12] However, this

[12] All *t* statistics and corresponding effect sizes are reported as positive values—the direction of an effect can be determined by the reported means and/or the text description.

difference was not statistically significant for the *underprepared* label, $t(117) = .32$, $p = .747$, $d = .06$.

Context. The contexts in which participants experienced being labeled by each of the four deficit-oriented labels were examined through one item that asked participants to indicate the circumstances in which these experiences occurred (if applicable), by selecting the contexts (e.g., class, orientation) and individuals (e.g., instructor, advisor) involved. As seen in *Tables 3.4* and *3.5*, a descriptive analysis of these responses showed that participants most frequently reported that they were labeled by an instructor during a class or during the instructor's office hours, or by their advisor during an advising session. For LIFG students only, the context most often mentioned was an on-campus event, such as an orientation.[13] As with the open-ended responses, another series of uncorrected Chi-Square tests were conducted to examine sample group differences in the contexts and labelers reported by students. These tests indicated that, relative to NLIFG students, a significantly larger proportion of LIFG students reported being labeled with deficit-oriented labels during on-campus events, at home, and online, as well as being labeled by advisors and staff.

Table 3.4. Summary of participants' closed-ended responses related to the *contexts* in which their deficit-oriented labeling experience occurred by sample group (expressed in percentages).

Context (N= 114)	Campus Event	Online	Advising Session	Class	Support Services	High School	Home
LIFG	75.9[a]	48.2[a]	69.0	74.1	55.2[a]	41.4	48.3[a]
NLIFG	39.3	28.6	64.3	66.1	35.7	39.3	25.0

Note. Sample size for LIFG n= 58 and n= 56 for NLIFG.
[a] denotes a statistically significant sample group difference in proportion (i.e., $p < .05$).

[13] The percentages associated with each category represent the proportion of participants who indicated they experienced being categorized in that context or by individuals in the specified roles. Participants were allowed to indicate more than one context and/or individual.

Table 3.5. Summary of participants' closed-ended responses related to the *individual* involved in their deficit-oriented labeling experience by sample group (expressed in percentages).

Labeler (N= 114)	Advisor	Instructor	Staff	Roommates	HS Counselor	Peers	Parents
LIFG	86.2[a]	74.1	70.7[a]	27.6	39.7	56.9	22.4
NLIFG	62.5	62.5	44.6	25.0	42.9	42.9	26.8

Note. Sample size for LIFG n= 58 and n= 56 for NLIFG.
[a] denotes a statistically significant sample group difference in proportion (i.e., $p < .05$).

Affective Consequences. The extent to which participants experienced negative affect as a result of being characterized by each of the four labels was analyzed in two ways, which were designed to probe different aspects of students' affective reactions to these labeling experiences. For each label, participants who indicated they had "*never*" been characterized by the deficit-oriented label were excluded from the analyses for these items, which resulted in varying sample sizes for the subsequent analyses (n= 84-97).[14]

First, I examined between-group differences in the extent to which participants reported feeling negative affect as a result of being categorized by deficit-oriented labels, through a series of independent samples *t*-tests. These analyses were conducted using the negative affect item for each of the four labels as dependent variables. The results of these analyses revealed that there were no statistically significant between-group differences in the extent to which LIFG and NLIFG students indicated they experienced negative affect as a result of being labeled as *underprepared* (M= 3.41, SD= 1.15; M= 3.23, SD= 1.04, respectively), $t(95) = .80$, p= .424, d= .16, *at-risk* (M= 3.33, SD= 1.23; M= 2.98, SD= .98, respectively), $t(89) = 1.61$, p= .112, d= .33, *underrepresented* (M= 2.38, SD= 1.27; M= 2.58, SD= 1.08, respectively), $t(82) = .79$, p = .429,

[14] Responses were excluded from these analyses for each label if participants chose *not applicable* for the effect item or indicated they had *never* been characterized the label, but responded to the corresponding effect item for that label with any response option besides *not applicable* (*underprepared* n= 22; *at-risk* n= 28; *disadvantaged* n= 29; *underrepresented* n= 35).

d= .17, or *disadvantaged* (*M*= 3.06, *SD*= 1.01; *M*= 3.03, *SD*= 1.27, respectively), *t*(88) = .14, *p*= .888, *d*= .03. Following this, I conducted a series of hierarchical regression analyses to explore associations between the frequency with which participants indicated they were labeled by each of four deficit-oriented labels and the negative affective consequences they subsequently experienced, as well as differences in these associations by sample group.

The dependent measure for each analysis was the negative affect item for one of the four deficit-oriented labels. In Block 1, I added the dummy-coded *sample group* (*G*) variable as a predictor, as well as the *frequency item* corresponding to the same label as the dependent variable; and, in Block 2, I added the *sample group* × *frequency*. As seen in *Table 3.6*, these analyses yielded statistically significant standardized regression coefficients (β) for *frequency* for all but the 'disadvantaged' label. However, none of the analyses yielded significant coefficients for *sample group* or the *sample group* by *frequency* interaction. Note that separate correlational analyses conducted within each group indicated a significant positive correlation between frequency of the 'disadvantaged' label and negative affect for NLIFG students. In addition, the correlation between frequency of the 'underrepresented' label and negative affect was not significant for LIFG students, despite the significant main effect.

Table 3.6. Parameter estimates for OLS regressions with frequency of labeling experiences and sample group predicting affective consequences.

	at-risk (*n*= 91)	*underprepared* (*n*= 97)	*underrepresented* (*n*= 84)	*disadvantaged* (*n*= 90)
Adjusted R^2 (Block 1)	.17	.19	.08	–.01
Adjusted R^2 (Block 2)	.17	.11	.09	<.01
F (Block 1)	10.45[a]	7.43[a]	4.38[a]	.42
F (Block 2)	7.20[a]	4.93[a]	3.58[a]	1.06

	at-risk (n= 91)	*underprepared* (n= 97)	*underrepresented* (n= 84)	*disadvantaged* (n= 90)
$Sample\ Group$ (G) β	–.02	–.12	.18	<.01
$Frequency\ (F)\ \beta$	.43[a]	.36[a]	.31[a]	.10
$G * F\ \beta$	–.29	–.10	.53	.63

Note. Main effects are reported from Block 1 of the analysis and the interaction term is reported from Block 2.
[a] denotes a statistically significant F or β ($p < .05$).

Interpretations of Hypothetical Students' Labeling Experiences

The results reported below are from a series of analyses designed to investigate participants' interpretations of hypothetical scenarios, each of which described an advisor characterizing a low-income, first-generation student as an *at-risk* or *first-generation* student. More specifically, these analyses were designed to test my primary hypotheses regarding differences in students' interpretations of a deficit-oriented label (*at-risk*) versus a neutral label *(first-generation),* and their perceptions of the potential effects of being labeled by each of these descriptors. The first analysis used the *valence belief* item as a dependent measure to examine the extent to which participants believed the advisors held *positive* or *negative* beliefs about the student they were advising. The second analysis used the *intentionality* item to examine the extent to which participants believed the advisors were *intentionally* trying to communicate their positive or negative beliefs about the student through their feedback. Lastly, for the third set of analyses, I used the *aggregate effects* variable as a dependent measure to examine participants' perceptions regarding the extent to which the hypothetical student might experience affective and motivational consequences as a result of the interaction with their advisor.

All of these analyses were conducted as *2*-way ANOVAs, with the *valence belief* item, the *intentionality* item, or the aggregate effects variable as the dependent measure, and *sample*

group (*LIFG; NLIFG*) and *label condition* (*deficit-oriented label*; *neutral label*) as between-subjects factors.[15] Further, all of the statistically significant interactions that were relevant to the hypotheses being tested in this study were probed through uncorrected pairwise comparisons based on estimated marginal means.

Valence and intentionality of beliefs communicated by a deficit-oriented label and a neutral label. The results of the analysis using *valence belief* item yielded a main effect of *label condition* on participants' valence beliefs rating, $F(1, 115) = 15.24$, $p < .001$, $\eta_p^2 = .12$, such that on average, participants in the *deficit-oriented label* condition rated the advisor's feedback as communicating more negative beliefs about the hypothetical student ($M = 3.94$, $SE = .15$), than did participants assigned to the *neutral label* condition ($M = 3.10$, $SE = .15$). The results of two, one-sample *t*-tests indicated that the mean rating for valence beliefs for the *deficit-oriented label* condition was significantly higher than the midpoint of the scale (i.e., 3.5), $t(57) = 2.84$, $p = .006$—whereas the mean rating for the *neutral label* condition was significantly lower than the midpoint of the scale, $t(60) = 2.42$, $p = .019$. These results indicate that participants generally perceived the 'at-risk' label as communicating relatively negative beliefs about the hypothetical student and the 'first-generation' label as communicating relatively positive beliefs. Lastly, both the main effect of *sample group* and the *2*-way interaction between *sample group* × *label condition* were not significant, $F(1, 115) = .85$, $p = .358$, $\eta_p^2 = .01$; $F(1, 115) = 2.27$, $p = .135$, $\eta_p^2 = .02$, respectively.

The results of the analysis with the *intentionality* item yielded a statistically significant main effect of *label condition* on participants' perceptions of the extent to which the advisor was

[15] Initial models were run prior to these that included the gender of the hypothetical student (male; female) as an additional between-subjects factor and participants' age as a covariate. However, there were no statistically significant main effects or interactions including gender or age in any of these models, so they were removed from the final analyses reported in this section.

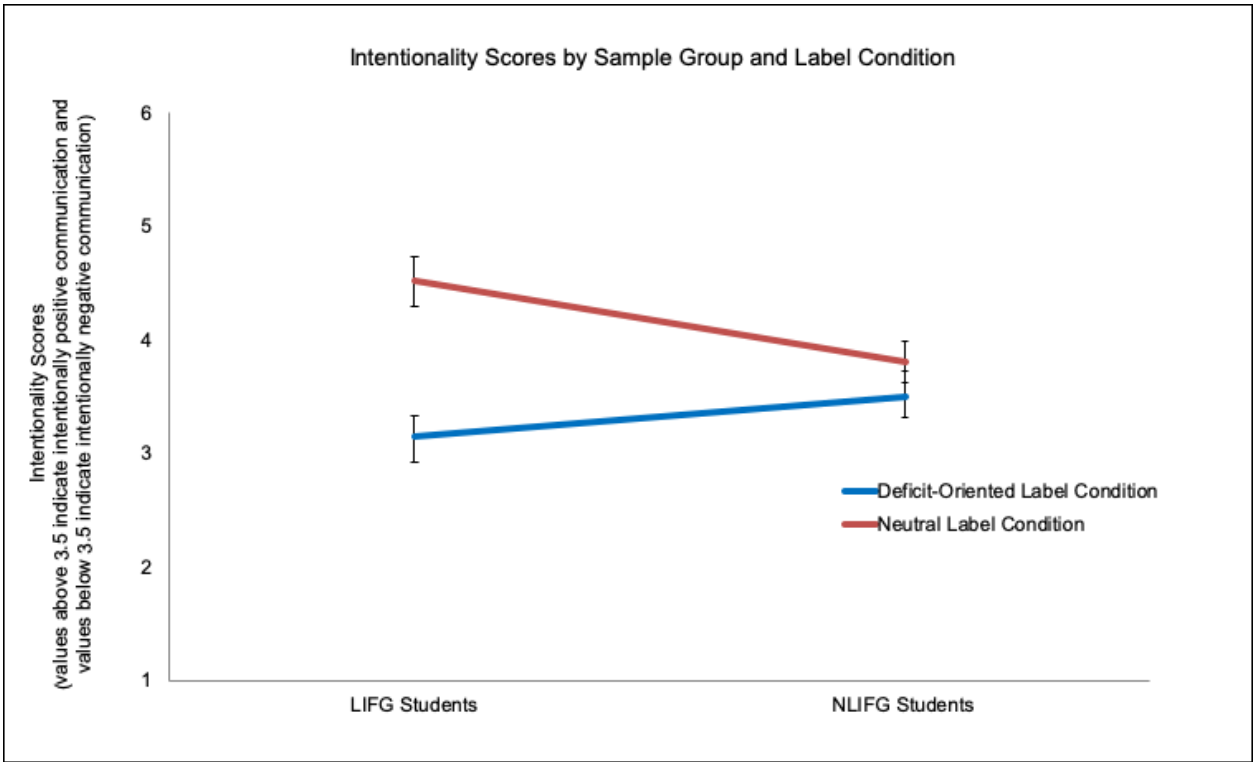
intentionally conveying their beliefs about the student with their feedback, $F(1, 115) = 16.65$, $p < .001$, $\eta_p^2 = .13$. More specifically, on average, participants in the *neutral label* condition perceived the advisor as being more intentionally positive in using their feedback ($M = 4.16$, $SE = .14$), relative to participants in the *deficit-oriented label* condition ($M = 3.32$, $SE = .15$). However, it is important to note that the response scale inadvertently referenced the valence of the advisor's beliefs and not the degree of intentionality, so these results should be interpreted cautiously.

Although the main effect of *sample group* was nonsignificant, $F(1, 115) = .77$, $p = .381$, the *2*-way interaction between *sample group* × *label condition* was statistically significant $F(1, 115) = 6.73$, $p = .011$, $\eta_p^2 = .06$. Pairwise comparisons by *sample group* showed that for participants assigned to the *neutral label* condition, LIFG students' perceived the advisor as being more intentionally positive in using their feedback to communicate their beliefs about the hypothetical student ($M = 4.52$, $SE = .22$) than NLIFG students' ($M = 3.81$, $SE = .18$), $F(1, 115) = 6.18$, $p = .014$, $\eta_p^2 = .05$. In contrast, participants' perceptions regarding the extent of the advisor's intention to communicate their beliefs about the student did not differ significantly between LIFG and NLIFG students in the *deficit-oriented* label condition ($M = 3.15$, $SE = .19$; $M = 3.50$, $SE = .23$), $F(1, 115) = 1.44$, $p = .233$, $\eta_p^2 = .01$ (see *Figure 3.1* below). However, the mean rating on the intentionality item for LIFG students in this condition was significantly lower than the midpoint of the scale, $t(57) = 2.20$, $p = .031$, whereas the mean rating for NLIFG students was equal to the midpoint of the scale. The comparisons by *label condition* showed that LIFG students' ratings on the intentionality item differed significantly between those who were assigned to the *deficit-oriented label* condition ($M = 3.15$, $SE = .19$) and those assigned to the *neutral label* condition ($M = 4.52$, $SE = .22$), $F(1, 115) = 22.28$, $p < .001$, $\eta_p^2 = .16$, whereas this

difference was not statistically significant for NLIFG students' (M= 3.50, SE= .23; M= 3.81, SE= .18, respectively), $F(1, 115) = 1.10$, p= .296, η_p^2 = .01.

Perceptions of the negative affective and motivational consequences of a deficit-oriented label and a neutral label. The results of the analysis using the *aggregate effects* variable as a dependent measure yielded a significant main effect of *label condition*, $F(1, 115) = 24.50$, $p < .001$, η_p^2 = .18, such that relative to participants in the *neutral label* condition (M= 3.13, SE= .12), participants' in the *deficit-oriented label* condition (M= 4.00, SE= .13) expected the student to experience a greater degree of negative affective and motivational consequences as a result of the interaction with their advisor. However, the analysis did not yield a significant main effect of *sample group*, $F(1, 115) = .68$, p= .411, η_p^2 = .01, or a significant *sample group* × *label condition*, $F(1, 115) = .78$, p= .379, η_p^2 = .01.

Figure 3.1. The significant two-way interaction between sample group and label condition for participants' intentionality scores in Study 1.



Discussion

The analyses of students' responses to both open-ended and closed-ended items about their deficit-oriented labeling experiences yielded several interesting findings. Although the majority of participants indicated they had previously experienced being categorized by a *deficit-oriented* label, LIFG students were significantly more likely than NLIFG students to report having been categorized by these labels in the past and being labeled as *at-risk*, *disadvantaged*, and *underrepresented* more often. Participants also indicated these experiences had taken place during interactions with instructors in class or with academic advisors during advising sessions. Moreover, relative to NLIFG students, LIFG students also indicated these experiences took place more often at home and online. Another interesting finding was that for three out of four of the deficit-oriented labels, there was a significant correlation between the frequency of participants' labeling experiences and the negative affect associated with these experiences. However, the associations between frequency of labeling experiences and negative affect extended to all participants, and thus, did not support my prediction that LIFG students would be more likely than NLIFG students to experience an accumulation of negative effects as a result of a greater number of deficit-oriented labeling experiences.

With respect to the hypothetical scenarios, both LIFG and NLIFG students in the *deficit-oriented label* condition perceived the advisor as communicating negative beliefs about the hypothetical student with their feedback, whereas participants in the *neutral* label condition perceived the advisor as communicating more positive beliefs about the hypothetical student with their feedback. Interestingly, LIFG students in the *neutral label* condition rated the advisor as being more intentionally positive in using their feedback to communicate their beliefs about the student than NLIFG students assigned to the same condition and LIFG students assigned to

the *deficit-oriented label* condition. Finally, both LIFG and NLIFG students in the *deficit-oriented label* condition expected the hypothetical student to experience a greater degree of negative consequences as a result of the interaction with the advisor, relative to participants in the *neutral label* condition.

Overall, the findings from this study provided valuable insights into several aspects of students' deficit-oriented labeling experiences. First and foremost, they provide empirical support for the anecdotal claims made by anti-deficit scholars; specifically, deficit-oriented labels do appear to be communicated to college students. These findings also extend those of Castro (2014), such that interactions involving deficit-oriented labels generally seem to occur with instructors, academic advisors, and other individuals who work closely with students (e.g., university staff). The findings from this study also suggest that deficit-oriented labeling experiences are associated with some degree of negative affective consequences and that a greater number of these experiences are associated with a greater degree of negative affective consequences. Although there was no direct evidence to suggest that LIFG students experienced a higher level of accumulated effects compared to NLIFG students, LIFG students did report having these experiences more often, which suggests that this possibility should be investigated further.

Interestingly, although all participants perceived the 'first-generation' label more positively than the 'at-risk' label, LIFG students seemed to interpret this characterization as more *intentionally* positive than NLIFG students did. Although I expected LIFG students to perceive the 'first-generation' label less negatively than the 'at-risk' label, this finding was somewhat unexpected—given that many of the same assumptions typically made about low-income Black and Hispanic students are also often applied to first-generation college students (Gray, 2013). I

was not expecting LIFG students to perceive the 'first-generation' label as intentionally communicating positive beliefs. However, this finding also makes sense, considering that for many students, bearing the 'first-generation' label represents a source of pride and accomplishment for being the first in their family to go to college—an element that makes this descriptor uniquely different from the former. Therefore, it is possible that LIFG students might have interpreted the advisor's use of this label as a nod to the challenges and obstacles the student had likely already overcome to get into that institution, and possibly a way of communicating their confidence in the student's ability to overcome any current academic struggles they may be experiencing. From an applied perspective, this finding is encouraging because it suggests that in addition to this label, it may be possible to identify additional and more adaptive ways in which to characterize students in lieu of deficit-oriented labels.

Limitations

As is the case with all research, this study was not without its limitations. First, the sample used for this study was relatively small, consisting of less than sixty-five students per sample group, which were further reduced for the analyses using the data from the hypothetical scenarios. Further, all of the participants for this study were recruited through Mechanical Turk, which is problematic because college students completing surveys online for supplemental income may have unique characteristics that differ from the majority of the college student population. These limitations introduce some concerns about the extent to which my findings can be generalized to other LIFG and NLIFG students.

Another limitation of this study was related to the design of the survey itself, such that it may have been problematic to ask participants to describe and reflect on their own deficit-oriented labeling experiences and then immediately following that, asking them to interpret the

labeling experience of a hypothetical student. Specifically, for participants assigned to the *deficit-oriented label* condition for the hypothetical scenario task, the student in the scenario was characterized by one of the same labels that participants had been asked about in the first half of the study (i.e., *at-risk*). Therefore, it is possible that participants' own experiences of being labeled as 'at-risk' could have influenced the manner in which they interpreted the advisor's feedback in the hypothetical scenario. That said, this portion of the study was meant to pilot materials for a future study, which meant that I would have an opportunity to compare these findings against those of a subsequent study.

Implications for Study 2 and 3

Due to the exploratory nature of this study, the subsequent research conducted for this dissertation was designed to refine and extend the findings from this preliminary study, both by addressing various methodological limitations and exploring new research questions.

In order to address the potential limitations of using MTurk as the only source of recruitment, I included other methods of recruitment in one of the subsequent studies. In the subsequent research, I also aimed to address the possible problems associated with asking participants to describe and reflect on their own deficit-oriented labeling experiences and then immediately following that, asking them to interpret the labeling experience of a hypothetical student. That is, for participants in the deficit-oriented label condition, reflecting on the negative consequences of being labeled in a deficit-oriented manner during the first two sections of the study may have primed them to interpret these labels in a negative light in the last section of the study. However, reversing the order of the sections for the next study would have likely caused similar issues, such that participants assigned to the *deficit-oriented label* condition might have also been influenced by their interpretations of the hypothetical scenario when responding to

questions about their own experiences of being labeled as *at-risk*. Consequently, I made the decision to divide the subsequent research into two separate studies, focusing primarily on students' own labeling experiences in one, and their interpretations of a hypothetical student's experience in another.

Moreover, I felt it was important to examine other types of consequences students might experience as a result of being characterized by a deficit-oriented label, specifically related to their motivation. In the present study, when participants were asked to indicate the extent to which a hypothetical student might experience various consequences as a result of being labeled as 'at-risk', participants' responses seemed to suggest that these experiences could also adversely influence students in ways that would impact their academic self-perceptions and sense of belonging. There is some evidence to support this argument, such that students who are characterized as *learning disabled* have lower levels of academic self-concept and sense of belonging in that academic context, compared to their non-labeled peers (e.g., Winne, Woodlands, & Wong, 1982). There is also is evidence in the motivational literature to suggest that poor academic self-perceptions can negatively influence students' feelings of belonging in college, and vice versa—and that, in combination, they can potentially result in lower levels of academic engagement (e.g., Walton & Carr, 2012). As such, in Study 2, in addition to examining affective consequences, I more carefully examined students' academic self-perceptions and engagement, and their sense of belonging in college.

Additionally, I explored the possibility that students might suffer stereotype threat effects following an experience of being labeled by a deficit-oriented descriptor. Although there is no shortage of speculation in the literature regarding the negative stereotypes associated with labels like 'at-risk', there is some documented evidence regarding the negative assumptions about the

motivation and academic competence of students characterized by such labels (Castro, 2014; Gray, 2013). However, no prior research has explored the possibility that characterizing students by these deficit-oriented descriptors might trigger stereotype threat. Moreover, there is sufficient evidence to indicate that the *perception* of a negative stereotype is often sufficient to elicit the effects of stereotype threat, regardless of the extent to which that perception is accurate (Aronson & Steele, 2005). By this account, any student who is labeled as 'at-risk' could be vulnerable to experiencing effects of stereotype threat if they perceive the 'at-risk' label as being associated with negative stereotypes about their academic competence and/or potential—whether or not this association exists in the mind of the labeler. Moreover, considering the pervasive stereotypes associated with Black and Hispanic students—particularly those from low-income backgrounds—when labeled as 'at-risk,' these students might interpret the label as indicating racial stereotypes as well. This suggests that LIFG students could be particularly vulnerable to experiencing the effects of stereotype threat as a result of being characterized by a deficit-oriented descriptor. Thus, examining stereotype threat effects in this context (as I did in Study 3) could prove to be particularly beneficial for advancing our understanding of the unique challenges that hinder LIFG students' success in college.

Finally, based the substantial amount of evidence documenting the ways that students' motivational beliefs influence their academic outcomes (e.g., Aronson, Fried, & Good, 2002; Dweck & Leggett, 1988), one of the goals of the next two studies was to explore students' motivational beliefs as potential moderators of their interpretations of deficit-oriented labels. As discussed in Chapter 2, the existing research suggests that students' *academic mindsets*, their *racial identity beliefs*, and their *stereotype threat vulnerability* might be particularly likely to influence how students interpret labeling experiences in academic environments. Thus, in Study

3, I explored these beliefs as a means for understanding within- and between-group differences in the way students interpret and react to deficit-oriented labeling experiences, and—given the malleability of these beliefs—as a means to inform the development of programs focused on fostering students’ resiliency to stigmatizing experiences.

CHAPTER 4: STUDY 2

The purpose of Study 2 was to continue exploring LIFG and NLIFG students' deficit-oriented labeling experiences. Specifically, this study focused on research questions related to the frequency with which students experienced being labeled by a deficit-oriented descriptor, the contexts in which these experiences occurred, and the affective and motivational consequences students endured as a result of these experiences. As with the previous study, I probed students' deficit-oriented labeling experiences through both open-ended and closed-ended items. More specifically, I used an open-ended prompt that asked participants to describe an experience in which they were labeled by a deficit-oriented descriptor, including the context in which the experience occurred and the way they felt afterwards. Although this prompt was similar to the one used in Study 1, this one included more information about the types of details participants should include in their description.

For this study, I also made several changes to the closed-ended items, both as a result of the findings from the first study and in an attempt to streamline the survey and make it feasible for participants to complete it in under 10 minutes. First, with respect to the items that asked participants to indicate the frequency with which they had been labeled by specific deficit-oriented labels, I did not ask participants about the *underrepresented* label, instead, they were only asked about the *at-risk*, *underprepared*, and *disadvantaged* labels, due to time restraints. Given that White college students are rarely a minority group within their institution, in contrast to the other three labels, the *underrepresented* label was likely irrelevant to the majority of the students in the NLIFG sample. I also revised the response scale for the frequency item for each label, in order to make it more concrete, such that rather than using subjective adjectives to quantify frequency (e.g., 'never', 'sometimes', 'often'), I used objective frequencies (e.g., 'not at

all this past academic year', '1-2 times this past academic year', 'more than 5 times this past academic year'). Moreover, rather than ask participants about the extent to which they experienced negative affect as a result of being labeled for each individual label, I asked them to indicate the extent to which they had experienced various affective and motivational consequences as a result of being labeled by 'one or more' of the three labels they were questioned about (i.e., *at-risk*, *underprepared*, and *disadvantaged*). And lastly, I also incorporated measures of participants' *academic mindsets*, their *racial identity beliefs*, and their *stereotype vulnerability*, in order to examine associations between these factors and the extent to which participants reported negative consequences as a result of being labeled by one or more deficit-oriented descriptors.

Several of my hypotheses for Study 2 were the same as from the previous study. For instance, for the open-ended items, I expected that more LIFG students would report having been labeled by a deficit-oriented label than NLIFG students. With respect to frequency with which students would report being labeled by three specific deficit-oriented descriptors, I predicted (based on the findings from Study 1) that LIFG students would generally report having been labeled by the deficit-oriented descriptors more frequently than NLIFG students. I also expected that relative to NLIFG students, LIFG students would also report experiencing a greater amount of negative affective and motivational consequences as a result of these experiences. Moreover, I also expected associations between students' academic mindsets, their racial and ethnic identity beliefs, their stereotype vulnerability, and the extent to which they reported experiencing negative consequences as a result to being labeled by deficit-oriented descriptors. Because perceiving intelligence and academic ability as malleable should lead students to respond more constructively to being characterized by a descriptor that communicates negative information

about their academic ability, I expected students' endorsement of a growth mindset to be negatively associated with the perceived negative consequences of being labeled by deficit-oriented descriptors.

With respect to racial identity beliefs, given the evidence that suggests stronger identification with their race and/or ethnicity can protect Black and Hispanic students from the negative effects of stigmatization, my intuition was that stronger identification with one's race and/or ethnic group would be negatively associated with the extent to which LIFG students reported experiencing negative consequences as a result of being labeled by deficit-oriented descriptors. I also predicted this same association for the NLIFG students, given that strongly identifying as a White student—which is an identity not associated with any pervasive negative academic stereotypes—might foster a sense of resiliency that protects these students from the potential consequences associated with being labeled by a descriptor that conveys negative information about their academic competence.

Finally, I expected to find a positive association between students' stereotype vulnerability and the perceived negative consequences of being labeled by deficit-oriented descriptors. Given that deficit-oriented labels are likely associated with negative stereotypes about students' motivation and academic competence that could potentially apply to both LIFG and NLIFG students, I expected to find this association universally, across all participants. However, given that LIFG students are associated with multiple academically stigmatized identities, I expected that these labels would be interpreted as being particularly threatening by these students, such that the association between their stereotype vulnerability and the consequences they report experiencing as a result of deficit-oriented labeling experiences would be stronger than that of NLIFG students. Lastly, considering that I assessed students' racial and

ethnic identity beliefs in the context of students' identities as a *college student*—rather than their global identity—my intuition was that for Black and Hispanic students specifically, stronger identification with their race and/or ethnicity would also mean they perceived educational attainment as a central aspect of being Black and/or Hispanic. Although this hypothesis was not directly tested in this study, finding a negative association between LIFG students racial and ethnic identity beliefs and their stereotype vulnerability would provide some evidence support to my argument.

In the remainder of this chapter, I describe the methods used to collect the data for this study, including specific details regarding the recruitment of participants, the survey used (including all prompts, items, and measures), and the procedure participants followed to participate in the study. Following this, I describe the analyses conducted on the data and the results yielded from those analyses. Finally, I end with a summary and discussion of the findings.

Method

Inclusion/Exclusion Criteria

Prospective participants were deemed eligible to participate in this study if they had not participated in Study 1 and met the criteria required to be categorized into one of the two sample groups (i.e., LIFG; NLIFG), which was the same criteria used in Study 1.

Participants

Final sample. The final sample for Study 2 was N= 180 (LIFG, n= 88; NLIFG, n= 92) and consisted of undergraduate students between the ages of 18-32 (M=23, SD= 3.68), of which 57% were male. The majority of participants indicated they were native English speakers (97%) and upperclassmen (72%), majoring in the sciences (36%), social sciences (16%), humanities (15%), and education (12%). Seventy-three percent of students indicated that they were attending

a public university or college and 33% indicated that their institution was religiously affiliated. Sixty-eight percent of the LIFG sub-sample identified as Hispanic or Latino and 46% identified as Black or African American. One hundred percent of the participants in the NLIFG sub-sample identified as non-Hispanic White. Lastly, an independent *t*-test examining participants' responses to the subjective SES question on the demographics questionnaire (i.e., "How financially well off were you growing up?") indicated that on average, LIFG participants perceived themselves as being significantly less "well off" (M= 2.26; SD= .98) than NLIFG students (M= 3.30, SD= .84), $t(178) = 7.68$, $p < .001$, d= 1.14 (see *Table 4.1* for additional demographic information by sample groups).

Table 4.1. Summary of Study 2 demographic information for the full sample and by sample groups.

Sample	Age (*M*)	Sex *(Male)*	Hispanic	Black	White	Upperclassmen	Public Institution	Religious Affiliation
Full Sample (*n*= 180)	22.86	57.2	33.3	22.2	73.9	72.2	73.3	33.3
LIFG (*n*= 88)	22.80	53.4	68.2	45.5	46.6	63.7	71.6	46.6
NLIFG (*n*= 92)	23.16	60.9	0	0	100	80.5	75.0	20.7

Note: All numbers in columns 3-9 are expressed in percentages.

Exclusion of cases from final sample. A total of 1,570 Mechanical Turk workers consented to participate in this study, indicated they were full-time college students at a 4-year college or university, and completed the pre-screening questionnaire, from which 296 met the inclusion criteria for one of the two sample groups (143 LIFG; 153 NLIFG) and completed some portion of the study. Of those, 70 participants were subsequently removed from all subsequent

analyses due to incomplete data (n=58), because they completed the study in less than four minutes, which was one minute less than half of the estimated time for completion (n=12). One issue that arose in this study, which was not a concern in Study 1, was there appeared to be a number of survey responses that were produced by survey bot software (e.g., responses were identical to one another for many of the items; responses provided identical definitions of a label, rather than listing labels) Given that concerns over survey bot software use on MTurk has become increasingly problematic (Mason & Suri, 2012), data for 25 participants were flagged and removed from all subsequent analyses. In an attempt to determine the age cutoff in a methodical manner, rather than employing the same age cutoff that was used in Study 1 (i.e., 29), the maximum age threshold in this study was determined by computing the age that was equal to one standard deviation above the mean *age* of the semi-final sample (n=201; M=24.88; SD=6.84; i.e., 32). As a result, data from an additional 21 participants (10 LIFG; 11 NLIFG) were removed from the final sample for this study because they exceeded this age cutoff.

Recruitment & Compensation

Participants were recruited for this online study via Mechanical Turk and compensated using the same process and criteria used in Study 1.

Measures

Section 1: Close-ended items assessing labeling experiences.

***Frequency of labeling experiences*.** The frequency with which students had experienced being categorized as an *at-risk* [*underprepared; disadvantaged*] student, was measured using three items "How often (*if ever*) have you experienced being categorized as an *at-risk [underprepared; disadvantaged]* student?". Participants rated the frequency with which they had been labeled by these descriptors using the following 5-point Likert-type scale: 0= "Not at all

this past school/academic year"; 1= "Once in the past school/academic year"; 2= "A couple of times in the past school/academic year"; 3= "About 3-4 times in the past school/academic year"; 4= "More than 5 times in the past school/academic year". Students who answered all three frequency items with *"Never"*, were not presented with the subsequent items that measured affective and potential motivational consequences of labeling; instead, they were redirected to the section of the study that asked them to respond to an open-ended prompt about a prior labeling experience.

Affective and motivational consequences of labeling experiences. The affective and motivational consequences of students' labeling experiences were measured using 11 items that were either revised version of items used in Study 1 or adapted from other studies. The items measured the impact of students' labeling experiences on their sense of belonging, academic self-perceptions, affect, and engagement. (e.g., "Being labeled as at-risk, and/or underprepared, and/or disadvantaged made me feel like I'm not a valued member of my university's community."; see *Table 4.2* for the full list of items). Students' were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale from 1= "Strongly Disagree" to 6= "Strongly Agree". *Items 3* and *8*, were originally worded in positive terms (i.e., "Being labeled as an at-risk, and/or underprepared, and/or disadvantaged student made me feel…" *…like my university supports me and wants me to succeed*; *…motivated to work harder in my courses,* and were reverse coded prior to conducting any analyses. The 11 items were then used to create four aggregate subscale scores. Specifically, participants' responses for *items 1 and 2* were averaged to create a sense of belonging index (i.e., *SOB,* α=.77)[16], *items 4-7* were averaged to

[16] Item 3 was removed from the SOB subscale and was not used in subsequent analyses because it was poorly correlated with the other two items, $r(105)s < .049$, $p \geq .082$, and reduced the reliability of the subscale to α=.39.

create an *academic self-perceptions index* (*ASP,* α=.82), *items 10* and *11* were averaged to create an *affect index* (*AFF,* α=.69), and *items 8* and *9* were averaged to create an academic engagement index for each scenario (*AE;* though the alpha was low [α=.45], the items were significantly correlated, $r(105)= .291$, $p = .002$).

Section 2: Open-ended prompt about labeling experiences.

***Prompt for open-ended description of labeling experience*.** Participants were asked to recall the most recent experience in which they were categorized by a *deficit-oriented* label and then prompted to describe specific details about their labeling experience and provide some examples for each of these aspects. The primary difference between the prompt used in this study and the one used in Study 1 was that, in this study, participants were asked to describe only one labeling experience (see Appendix F for exact wording).

Table 4.2. Full list of affective and motivational effect items used in Studies 2 and 3.

Item	Subscale
"*The advisor's feedback probably made Aaron/Ryan [April/Casey] feel…*"	
Item 1: …like he [she] is not a valued member of his [her] university's community.	*SOB*
Item 2: …like he [she] don't belong at his [her] university.	*SOB*
Item 3: …like his [her] university supports him [her] and wants him [her] to succeed.	*SOB*
Item 4: …less confident in his [her] ability to do well in college	*ASP*
Item 5: …like he [she] needs more help than other students at his [her] university to pass his [her] classes.	*ASP*
Item 6: …like he [she] is not as smart as most of the other students at his [her] university.	*ASP*

Item 7: …like he [she] is not 'college material'.	*ASP*
Item 8: …motivated to work harder in his [her] classes.	*AE*
Item 9: …hesitant to take any challenging courses moving forward.	*AE*
Item 10: …discouraged about his [her] future in college.	*AFF*
Item 11: …embarrassed and/or ashamed.	*AFF*

Section 3: Individual differences measures.

***Academic mindsets*.** Students' academic mindsets were assessed using Dweck's (1999) 8-item Theories of Intelligence scale (e.g. "To be honest, you can't really change how intelligent you are"; see Appendix I for full list of items). Students were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale (1= "Strongly Disagree" to 6 = "Strongly Agree"). A single mindset index (M= 4.00, SD= 1.09, α = .91) was computed for each participant by reverse-coding the 4 fixed items (i.e., items 1, 2, 4, & 7), and computing the mean for each participant across all 8 items, such that higher scores on the index reflect more of a growth mindset (and less of a fixed mindset).

***Sensitivity to negative racial stereotypes about academic competence*.** The 8-item Stereotype Vulnerability Scale (Barnard et al., 2008) was used to assessed the extent to which students' feel threatened by negative stereotypes about the academic competence of students who belong to their racial or ethnic group (e.g., "Some people feel I have less academic success because of my race [ethnic background]."; see Appendix I for full list of items). Students were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale (1= "Strongly Disagree" to 6 = "Strongly Agree"). A single index (M= 3.25, SD= .85, α = .68) was computed for each participant by reverse-coding 4 items (items 2, 3, 5, & 7), and computing

the mean for each participant across all 8 items, such that higher scores on the index reflect a greater degree of vulnerability to negative stereotypes.

Racial identity beliefs. Students' racial identity beliefs were measured using seven items from a revised version of the 8-item *Centrality* subscale of the Multi-Dimensional Black Identity Scale (Sellers et al., 1997). The original measure was designed to assess the extent to which Black individuals feel that their race is a central aspect of their identity—however, for the purposes of this research, the wording of the items was revised to assess the extent to which Black [White; Hispanic/Latino; Black Hispanic] students feel that their race or ethnic background is a central aspect of their identity as college students (e.g., "Overall, being Black [White; Hispanic/Latino; a Black Hispanic] has very little to do with how I feel about myself as a college student."; see Appendix I for full list of items). Students were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale (1= "Strongly Disagree" to 6 = "Strongly Agree"). Participants were assigned to one of four possible versions of the scale (i.e., Black; White; Hispanic/Latino; Black Hispanic) based on their responses to the race and ethnicity items in the pre-screening questionnaire.[17] A single index (M= 3.44, SD= 1.10, α = .88) was computed for each participant by reverse-coding three items (items 1, 3, & 7), and computing the mean for each participant across all 7 items, such that higher scores on the index reflect a greater degree of importance of race or ethnic background for students' identity as a college student.

Procedure

Once prospective participants were directed to the study on Qualtrics research suite, participants were first presented with an electronic version of the consent form and were required

[17] Participants received the version for White students if they identified as White and non-Hispanic, they received the version for Black students if they identified as Black and non-Hispanic, they received the version for Black Hispanic students if they identified as Black and Hispanic, and they received the version for Hispanic students if they identified as White and Hispanic.

to give their consent to participate before they were allowed to complete the pre-screening questionnaire. Participants who did not give their consent to participate were redirected out of the study and presented with an "end of survey" message. Next, prospective participants completed the 11-item pre-screening questionnaire. Participants who were *not* eligible to be categorized into either of the sub-samples were redirected out of the study and presented with an "end of survey" message that indicated they were ineligible. Participants who were eligible to complete the survey were presented with a message informing them of their eligibility and were allowed to continue on to the study survey. Participants then completed *Sections 1-3* in order. In *Section 1*, participants were asked to indicate how often, if ever, they had been labeled by three deficit-oriented labels (i.e., *at-risk; underprepared; disadvantaged*). Once participants responded to the frequency items for each label, they were asked to respond to a series of 12 items that asked about the affective and motivational consequences they experienced as a result of being characterized by one or more of these labels. In *Section 2*, they were asked to respond to the open-ended prompt about a recent experience in which they were labeled by a deficit-oriented descriptor. In *Section 3* of the study, participants were asked to complete three different motivational measures—a measure of academic mindsets, a measure of sensitivity to negative racial [ethnic] stereotypes, and a measure of racial identity attitudes. This was followed by the demographic's questionnaire and an electronic version of the debriefing form, which was identical to the one used in Study 1 (see Appendix E), and included a completion code they could then use to receive their compensation through Mechanical Turk's website.

Results

Students' Open-Ended Descriptions of Labeling Experiences

In Study 2, students were asked to describe an experience in which they were categorized by a deficit-oriented label by responding to a series of open-ended prompts which asked them to describe when the experience took place, the label that was used, the context (e.g., situation) in which the experience occurred, and the effects they experienced as a result of being labeled. Students' responses to these prompts were first coded and then evaluated through descriptive and statistical analyses. The following subsections describe the coding process, the subsequent analyses, and the results yielded from those analyses.

Coding and analyses of students' open-ended responses.

The label. In order to analyze students' responses to the prompt that asked them to indicate which deficit-oriented label they were categorized by, I coded students' responses as *deficit-oriented label*=1 or *not applicable*=0, based on whether or not the label they described could be considered a deficit-oriented label. Labels were assessed and coded in a very similar manner as Study 1, with the exception of some added criteria. More specifically, for instances where students indicated they had been labeled as *unprepared*, the label was coded as deficit-oriented only if the context in which the labeling experience occurred met the one or more of the requirements to be categorized as a deficit-oriented label. If the student described being labeled as *unprepared* for factors under their control and/or factors that were temporary, the label was coded as *not a deficit-oriented label* (e.g., because they were late to class; failed to study for an exam). All other labels and responses that did not meet these criteria such as adjectives (e.g., stupid; dumb; lazy) or irrelevant responses were coded as *not applicable*. This added criteria for

the *unprepared* label was used as a result of a large proportion of NLIFG students who seemed to be confounding the *underprepared* label with *unprepared*.

Using this new categorical variable, I conducted both a descriptive analysis and a Chi-Square test of independence to examine between-group differences in the proportion of students in each sub-sample who reported being categorized by a deficit-oriented label. The results of the descriptive analysis showed that 39.7% of LIFG students reported being categorized by a deficit-oriented label (n=35), whereas 29.3% of NLIFG students indicated being labeled by a deficit-oriented descriptor (n=27). Although this indicated a between-group difference of approximately 10%, the results of the Chi-Square test indicated that this difference in proportion was not statistically significant, $\chi^2(1, N= 180) = 2.17, p= .141$.

The context. Students' open-ended responses to the prompt that asked them to describe the context in which the labeling experience occurred were categorized into one of five categories based on the students' description (i.e., class; faculty meeting; orientation; advising; extra-curricular) and coded as such into a new variable named *context*. These categories were based on both the categories used in Study 1, but were also tailored to align with participants' responses in this study. Some students indicated these experiences took place in contexts such as "a club meeting", "marching band practice", or "during practice" for a sport; these responses were coded as *social*, meaning they took place in a non-academic setting on campus. It is important to note that the only responses that were coded for the *context* variable were those that were also coded as being *deficit-oriented relevant* for the categorical variable discussed in the prior subsection; thus, the sample for this variable was N=62 (LIFG n= 35; NLIFG n= 27). As seen in *Table 4.3*, for both subsamples, the contexts listed most frequently were the classroom and advising sessions. Moreover, a series of uncorrected Chi-Square tests were conducted to

examine sample group differences the contexts and labelers reported by LIFG and NLIFG students, which indicated that relative to NLIFG students, a marginally larger proportion of LIFG students reported being labeled by advisors, whereas a significantly larger proportion of NLIFG students reported being labeled by instructors.

Table 4.3. Summary of participants' open-ended responses related to the *contexts* in which their deficit-oriented labeling experience occurred by sample group (expressed in percentages).

Context (*N*= 62)	Class	Faculty Meeting	Advising Session	Orientation	Extra-Curricular
LIFG	28.6	5.7	37.1	17.1	11.4
NLIFG	44.4	7.4	22.2	14.8	11.1

Note. Sample size for LIFG *n*= 35 and *n*= 27 for NLIFG.

Table 4.4. Summary of participants' open-ended responses related to the *individual* involved in their deficit-oriented labeling experience by sample group (expressed in percentages).

Labeler (*N*= 62)	Advisor	Instructor	Peers	Staff
LIFG	57.1[b]	25.7	14.3	2.9
NLIFG	33.3	55.6[a]	7.4	3.7

Note. Sample size for LIFG *n*= 35 and *n*= 27 for NLIFG.
[a] denotes a statistically significant Chi-Square statistic for the comparison of sample groups ($p < .05$).
[b] denotes a marginally significant Chi-Square statistic for the comparison of sample groups ($p < .10$).

The "labeler". Students' open-ended responses to the prompt that asked them to indicate who had used the deficit-oriented label to categorize them were also coded into one of four categories based on students' descriptions (i.e., instructor; advisor; peer; staff). As with the *context* variable discussed above, the categories for this new *labeler* variable were also developed based on the categories from Study 1 and a review of students' responses; and, again, it was limited to the subsample of students who provided a *deficit-oriented relevant* label. As seen in *Table 4.4* (above), for both subsamples, the individuals listed most frequently were advisors and instructors.

Effects of deficit-oriented labeling experience. Students' open-ended responses to the prompt that asked them to detail the effects they experienced as a result of the labeling experience they described were first coded into one of five categories based on students' descriptions. Responses that indicated a negative effect was experienced as a result of the labeling experience (e.g., "*I felt demoralized and degraded*") were coded as *negative*, responses that indicated a positive effect was experienced as a result of the labeling experience (e.g., "*I was motivated to work harder*") were coded as *positive*, responses that indicated the effect of the labeling experience was neither positive or negative (e.g., "*I didn't care*") were coded as *neutral,* responses that indicated that both positive and negative effects were experienced as a result of the labeling experience (e.g., "*I felt sad but then later on more motivated to work harder*") were coded as *both positive and negative,* and lastly, responses that were irrelevant to the prompt and/or uninterpretable (e.g., "*I felt learned*") were coded as *not applicable/irrelevant*. As with the previous variables, the categories for this new *effects* variable were developed based on a review of students' responses and was also limited to the subsample of students who provided a *deficit-oriented relevant* label. The results of the descriptive analysis using the new categorical *effects* variable indicated that 82% of students (n=51) reported feeling negative effects as a result of their deficit-oriented labeling experience, whereas 13% of students (n=8) reported feeling positive effects as a result. Each of the remaining three categories had an n=1 and were not included in any subsequent analyses. In contrast to Study 1, these categories were mutually exclusive and did not distinguish between affective and motivational consequences. Because I was concerned that coding process in the first study was overly subjective, I made these changes in an effort to develop a process that was objective, straightforward, and minimized the likelihood of inconsistencies in coding. Following the coding process, I conducted a Chi-Square

test to determine if the proportion of LIFG students who reported experiencing negative effects as a result of being labeled by a deficit-oriented descriptor was significantly higher than the proportion of NLIFG students. However, the results of that analysis indicated that the difference was not significant, $\chi^2(1, N= 59) = .13, p = .716$.

Frequency of Labeling Experiences

The frequency with which students experienced being categorized by deficit-oriented labels were evaluated through their responses to three closed ended questions, each of which asked them to indicate how often (*if ever*) they had experienced being categorized by three different deficit-oriented labels (i.e., at-risk; underprepared; disadvantaged), from '0= "Not at all this past school/academic year" to 4= "More than 5 times in the past school/academic year.'

Students' responses to the three label items were evaluated and compared through two different approaches. First, I evaluated the frequency students reported experiencing being labeled at a broad level, by conducting a series of Chi-Square tests to examine differences in the proportion of LIFG students who reported having experienced being labeled at least once this past academic year, to the proportion of NLIFG students. To do this, I computed a series of dichotomous, categorical variables—one for each of the three label items that reflected whether students had indicated they experienced being categorized by that label *at least once in the past academic year* (coded as 1) or *not at all in the past academic year* (coded as 0). The results of the Chi-Square tests indicated there were statistically significant differences across all analyses, such that the proportion of LIFG students who indicated being labeled at least once in the past academic year was significantly greater than that of NLIFG students for the *at-risk* (LIFG 55.7%; NLIFG 28.2%), *disadvantaged* (LIFG 63.6%; NLIFG 17.3%), and *underprepared* labels (LIFG 53.4%; NLIFG 33.6%), $\chi^2(1, N= 180)s > 7.12, ps < .007$.

Next, I compared between-group differences in the frequency with which students reported being categorized by deficit-oriented labels this past academic year, by conducting a series of independent t-tests to compare mean differences for each of the three deficit-oriented labels individually and in aggregate form. For these analyses, I used the original frequency rating variables for each of the labels. The results of these analyses revealed statistically significant between-group differences for two of the three labels, such that on average, LIFG students reported that in the past academic year, they experienced being labeled as *at-risk* and *disadvantaged* significantly more frequently (M=.95, SD=1.06; M=1.22, SD=1.11, respectively) than NLIFG students (M=.50, SD=.88; M=.39, SD=.92, respectively), $t(178)s > 5.41$, $ps < .002$, $ds > .82$. However, this difference was only marginally significant for the *underprepared* label (LIFG M=.92, SD=1.06; NLIFG M=.63, SD=1.04, respectively), $t(178) = 1.85$, p=.065, d= .28.

Affective and Motivational Consequences of Labeling Experiences

The affective and motivational consequences of students' experiences of being labeled as *at-risk*, and/or *underprepared*, and/or *disadvantaged*, were assessed for the aggregate *SOB*, *AF*F, *ASP*, and *AE* variables. Given that these items were only presented to students who indicated they had experienced being labeled by *one or more* of the three deficit-oriented labels *at least once in the past academic year*, the sample size for these items decreased from 180 to 107 (LIFG n= 67; NLIFG n= 40).

Between-group differences in effects of labeling experiences. In order to examine between-group differences in the extent to which students experienced negative motivational and affective consequences as a result of their deficit-oriented labeling experiences, I conducted a series of independent t-tests using the four aggregate subscale scores as dependent variables. The results of these analyses indicated that there were no statistically significant between-group

differences in the mean effects reported by LIFG and NLIFG students for any of the aggregate subscale scores, $t(105)s < .63$, $ps > .530$, $ds < .10$.

Association between frequency of labeling experiences and affective and motivational effects. The associations between the frequency with which students in each group reported experiencing being categorized as at-risk, disadvantaged, and/or underprepared and the extent to which they experienced motivational effects were assessed through a series of hierarchical regression analyses. The dependent measure for each analysis was one of the four effect subscale scores as the outcome variable (i.e., *ASP; SOB; AFF; AE*). In Block 1, I added the dummy-coded *sample group* (*G*) variable and the aggregate *frequency item* as predictors; and, in Block 2, I added the *sample group* × *frequency*. Given that students who indicated that they had not experienced being labeled by at least one of the three labels in the past academic year were not presented with the effect items, the samples for each group were $n = 67$ for LIFG students and $n = 40$ for NLIFG students. As seen in *Table 4.5*, these analyses yielded statistically significant Beta coefficients for *frequency* for all but one of the subscale scores (i.e., academic engagement). However, none of the analyses yielded significant coefficients for *sample group* or the *sample group* by *frequency* interaction. Note that separate correlational analyses conducted within each group indicated that frequency was not significantly associated with ASP or AFF for NLIFG students, despite the significant main effects.

Table 4.5. Results of OLS regression analyses with frequency of deficit-oriented labeling experiences and sample group predicting affective and motivational outcomes.

	SOB	ASP	*AFF*	*AE*
Adjusted R^2 (Block 1)	.12	.04	.05	<.01
Adjusted R^2 (Block 2)	.11	.04	.05	–.01
F (Block 1)	8.12[a]	3.38 [a]	3.70[a]	1.08

	SOB	ASP	*AFF*	*AE*
F (Block 2)	5.36[a]	2.59[b]	2.83[a]	.71
Sample Group (G) β	–.05	–.04	.04	<.001
$Frequency\ (F)\ \beta$	.37[a]	.24[a]	.26[a]	.14
$G * F\ \beta$	.02	–.18	–.19	.03

Note. N= 107 for all analyses. Main effects are reported from Block 1 of the analysis and the interaction term is reported from Block 2.
[a] denotes a statistically significant $F\ or\ \beta\ (p < .05)$.
[b] denotes a marginally significant $F\ or\ \beta\ (p < .10)$.

Students' Motivational Beliefs

Associations between motivational beliefs. Two sets of correlational analyses were conducted to examine associations between students' academic mindsets, stereotype vulnerability beliefs, and racial identity beliefs, for each sample subgroup. As shown in *Table 4.6*, the analysis using LIFG students' scores yielded statistically significant correlations between their *mindset and stereotype vulnerability scores* and their *stereotype vulnerability and racial identity belief scores,* but not their *mindset and identity belief scores*. In contrast, the results of the analysis using NLIFG students' scores showed that these students' *mindset and stereotype vulnerability scores* were not significantly correlated—however, the association between their *stereotype vulnerability and racial identity belief scores* and *mindset and racial identity belief scores* were both statistically significant.

Between-group differences in students' motivational beliefs. Potential between-group differences in students' mean scores on all three motivational belief measures were examined through a series of independent *t*-tests. The results of these analyses yielded statistically significant between-group differences in students' mean *stereotype vulnerability scores* and their *racial identity belief scores*, but not in their *mindset scores*.

Table 4.6. Results of correlational analyses by sample group examining associations between students' motivational beliefs.

Motivational Beliefs	*1*	*2*	*3*
1. Racial Identity Beliefs	—	.26*	–.29*
2. Stereotype Vulnerability	.47*	—	.05
3. Academic Mindsets	.15	.35*	—

Note. Correlations for LIFG students (n= 88) are presented to the left of the diagonal and correlations for NLIFG students (n= 92) are presented to the right of the diagonal.
* Denotes a statistically significant r ($p < .05$).

More specifically, the results indicated that on average, LIFG students' were significantly more vulnerable to negative racial stereotypes about academic competence (M=3.77, SD=.60) than NLIFG students (M=2.76, SD=.76), $t(178) = 9.80$, $p<.001$, d= 1.47, and that their racial identity beliefs played a more central role in their identity as college students than that of NLIFG students (M=3.86, SD=.99; M=3.04, SD=1.06, respectively), $t(178) = 5.40$, $p<.001$, d= .80. There was only a marginally significant difference between LIFG and NLIFG students' *mindset scores* (M=4.13, SD=1.03; M=3.86, SD=1.13, respectively), $t(178) = 1.69$, p=.093, d= .25. Moreover, it is important to note that the mean *mindset scores* for LIFG and NLIFG students were both significantly higher than the midpoint of the scale (i.e., 3.5), $t(87\text{-}91)s > 3.06$, $ps <$ .004—which reflected these students' greater endorsement of a growth mindset relative to a fixed mindset.

Association between students' motivational beliefs and effects of labeling experiences. A series of hierarchical regression analyses were conducted to explore the associations between participants' motivational beliefs and the extent to which they reported experiencing negative affective and motivational consequences as a result of their deficit-oriented labeling experiences, as well as differences in these associations by sample group. The

dependent measure for each analysis was one of the four effect subscale scores as the outcome variable (i.e., *ASP; SOB; AFF; AE*). In Block 1, I added the dummy-coded *sample group* (*G*) variable as a predictor, as well as participants' standardized *mindset* (*M*) *scores*, *stereotype vulnerability* (*SV*) *scores*, and *racial identity belief* (*ID*) *scores*. In Block 2, I added the G × MS, G × SV, and G × ID interaction terms.

As seen in *Table 4.7*, the results of the analyses yielded statistically significant or marginally significant effects of *academic mindsets* and *stereotype vulnerability* for three out of the four outcomes (*SOB*, *ASP*, and *AFF*). The fourth outcome (*AE*) was significantly predicted by racial identity beliefs. These results indicate that, across sample groups, students with stronger academic growth mindsets were less negatively affected by their deficit-oriented labeling experiences, particularly when it came to their sense of belonging, affect, and academic self-perceptions. In contrast, participants with greater stereotype vulnerability were more negatively affected by their labeling experiences, and with respect to the same outcomes. Finally, students with stronger racial identity beliefs reported being more disengaged due to their labeling experiences.

These main effects were qualified by two marginally significant interactions. First, there was a marginal mindset × sample group interaction for the model predicting academic self-perceptions, such that there was a significant negative association between mindsets and ASP for LIFG students ($\beta = -.41$), $t(99)$, $p = .003$, but not for NLIFG students ($\beta = -.06$), $t(99) = .40$, $p = .693$. And, there was also a marginal racial identity beliefs × sample group interaction for the model predicting negative affect, such that there was a non-significant negative association between ID beliefs and AFF for LIFG students ($\beta = -.21$), $t(99) = 1.52$, $p = .132$, but a non-

significant positive association between these variables for NLIFG students ($\beta = .22$), $t(99) = 1.29$, $p = .201$.[18]

Table 4.7. Model estimates for analyses examining associations between motivational beliefs and consequences of deficit-oriented labeling experiences.

Parameter Estimates	*SOB*	*ASP*	*AFF*	*AE*
Adjusted R^2 (Block 1)	.20	.04	.11	.04
Adjusted R^2 (Block 2)	.20	.07	.15	.02
F (Block 1)	7.76[a]	2.04[b]	4.29[a]	2.23[†]
F (Block 2)	4.77[a]	2.19[a]	3.67[a]	1.27
$Sample\ Group\ (G)\ \beta$	.17	.05	.17	.13
$\boldsymbol{Mindsets\ (M)\ \beta}$	**–.30[a]**	**–.18[b]**	**–.27[a]**	**–.04**
$\boldsymbol{Stereotype\ Vulnerability\ (SV)\ \beta}$	**.47[a]**	**.23[b]**	**.31[a]**	**.11**
$\boldsymbol{Racial\ Identity\ Beliefs\ (ID)\ \beta}$	**–.03**	**.01**	**.03**	**.24[a]**
$\boldsymbol{G * M\ \beta}$	**.18**	**.24[b]**	**.05**	**–.01**
$\boldsymbol{G * SV\ \beta}$	**–.02**	**–.21**	**–.14**	**–.05**
$\boldsymbol{G * ID\ \beta}$	**–.07**	**–.09**	**–.28[b]**	**.07**

Note. Primary effects of interest are in bold. Main effects are reported from Block 1 of the analysis and the interaction term is reported from Block 2.
[a] Denotes a statistically significant $F\ or\ \beta\ (p < .05)$.
[b] Denotes a marginally significant $F\ or\ \beta\ (p < .10)$.

Discussion

The results of this study yielded several interesting findings and replicated several findings from the previous study. With respect to participants' open-ended responses, approximately 32 percent indicated they had been labeled by a deficit-oriented descriptor in the

[18] Simple slopes were computed using procedures described by Aiken and West (1991).

past and that these experiences had most frequently occurred during interactions with instructors or academic advisors. As expected, the majority of LIFG and NLIFG students reported experiencing negative affective consequences as a result of these experiences. Participants' responses to the closed-ended items showed that LIFG students were significantly more likely than NLIFG students to have experienced being labeled as *at-risk, disadvantaged, and underprepared* in the past academic year. LIFG students also reported significantly or marginally more of these experiences in that time period for each label.

Although there were no significant between-group differences in the negative consequences that students reported experiencing, a greater number of deficit-oriented labeling experiences was associated with a greater degree of negative effects on students' sense of belonging in college, academic self-perceptions, and affect (across sample groups). The results indicated that participants' motivational beliefs also influenced the negative effects students reported experiencing as a result of their deficit-oriented labeling experiences. Across sample groups, students with stronger academic growth mindsets were less negatively affected by their deficit-oriented labeling experiences than students with weaker growth mindsets, particularly when it came to their sense of belonging, affect, and academic self-perceptions. In contrast, students with greater stereotype vulnerability were more negatively affected by their labeling experiences than students with less vulnerability. Further, students with stronger racial identity beliefs reported being more disengaged due to their labeling experiences. The findings also indicated there were marginal sample group differences in the influence of students' motivational beliefs on their academic self-perceptions. Specifically, LIFG students with stronger growth mindsets had academic self-perceptions that were less negatively affected than LIFG students with weaker growth mindsets, whereas this was not the case for NLIFG students. Lastly, for

LIFG students, stronger racial identification was negatively associated with affective consequences, whereas the reverse was true for NLIFG students—however, neither of these associations were statistically significant.

Overall, these findings both compliment and extend the findings from Study 1. Across both studies, participants' close-ended responses indicated that LIFG students experienced being labeled as *at-risk* and *disadvantaged* more often than NLIFG students, which is consistent with predictions posited by several scholars (Castro, 2014; Pearl, 1991; Valencia, 1997; 2010). However, in contrast to Study 1, the proportion of Study 2 participants whose responses to the open-ended prompt indicated they had been labeled by a deficit-oriented descriptor in the past, was considerably lower (32% *v.* 69%), and did not differ by group. One possible reason for this decrease in proportion across studies is that in this study, the open-ended prompt asked students to recall only one experience in which they were labeled, whereas the prompt for the first study asked them to recall up to five deficit-oriented labels they had been characterized by in the past. It is also possible that by presenting the closed-items first in this study (the order was reversed in Study 1), the specific deficit-oriented labels included in those items served as concrete examples of the types of labels the open-ended prompt was referring to as 'deficit-oriented', which could have inhibited some participants from elaborating on their experiences being labeled by other descriptors that possibly could have been considered deficit-oriented as well. Although these labels were also embedded in the open-ended prompt, participants may have skimmed the prompt and missed these examples.

Another important finding from this study was with respect to the associations between the frequency with which students experienced being labeled by deficit-oriented descriptors and the extent to which students reported experiencing specific types of affective and motivational

consequences. Specifically, there were significant positive associations between the frequency of participants' labeling experiences and the extent to which their sense of belonging, academic self-perceptions, and affect were negatively influenced by their labeling experiences. However, this effect of frequency extended to both LIFG and NLIFG students and there were no significant sample group differences in these associations. The findings also suggest that stronger growth mindsets may shield students from some of the negative effects of academically stigmatizing experiences, which is consistent with findings from prior research (Aronson, Fried, & Good, 2002).

Surprisingly, for both LIFG and NLIFG students, greater stereotype vulnerability and racial identification were positively associated with some of the negative affective and motivational consequences of labeling experiences. Given the evidence from prior work linking greater stereotype vulnerability with negative affective and motivational outcomes for marginalized college students (e.g., Aronson & Steele, 2005), the negative influence of greater stereotype vulnerability was only expected for LIFG students. Moreover, there was some evidence to support my hypotheses that stronger racial identification may act as a protective factor against the effects of stigmatization for LIFG students.

Limitations. The limitations associated with this study were very similar to that of Study 1, thus, this discussion is brief. A major limitation of Study 2 was that participants' responses to the closed-ended items that asked them about specific deficit-oriented labels likely influenced their responses to the open-ended prompt. However, given my concerns that participants were being primed to think about these labels as being negative because the open-ended prompt asked them to recall an experience in which they were characterized by a 'negative label', reversing the order in which these tasks were presented was the only feasible solution. However, considering

that the findings were generally consistent across both studies, despite the fact that the order in which those sections were presented was reversed from one study to the next, only serves to validate them further.

CHAPTER 5: STUDY 3

The purpose of Study 3 was twofold. One objective of this study was to continue exploring LIFG and NLIFG students' interpretations of *deficit-oriented* labels versus a more *neutral* label. A second objective was to begin to explore the possibility that students' deficit-oriented labeling experiences might be associated with subsequent stereotype threat effects. Below, I discuss the manner in which these objectives were addressed in this study and my respective predictions. It is important to note that Study 2 and 3 were run concurrently; thus, the results and findings from Study 2 did not influence the design of Study 3.

Examining students' interpretations of deficit-oriented versus neutral labels. Similar to Study 1, this portion of the study relied on hypothetical scenarios about a low-income, first-generation college student receiving feedback from their advisor, and included subsequent items measuring the extent to which participants interpreted the advisor's feedback as communicating positive or negative beliefs about the student, the extent to which they believed the advisor was intentionally communicating those beliefs, and their perceptions about the types of affective and motivational consequences that the student in the scenario might experience as a result of the advisor's feedback. However, based on the findings from Study 1, I made a few changes in order to improve on the methodology and expand on these findings.

First, considering that the *first-generation* label used in the *neutral label condition* might not apply to some 'at-risk' or 'underprepared' students, I felt it was important to explore other ways of characterizing students that might reduce the risk of stigmatization. Drawing on Castro's (2014) findings, I reasoned that it might be possible to reduce the damaging implications of *at-risk* and *underprepared* by adding context to such labels that would lead participants to attribute the perceived discrepancies in academic performance or achievement to broad social inequities

rather than some internal 'deficiency' or inherent limitation. Fostering a less threatening interpretation of 'at-risk' or 'underprepared' in this way might help mitigate at least some of the negative affective and motivational consequences of these labeling experiences. Therefore, in addition to the *deficit-oriented label* and *neutral label* conditions, I added a *deficit-oriented label + context* condition to Study 3. In this new condition, the advisor characterizes the student using the same deficit-oriented label used by the advisor in the *deficit-oriented label* condition, but provides additional context that attributes that characterization to 'a lack of access to the types of educational opportunities and resources that helped their peers prepare for college'.

Moreover, I also added an additional scenario to each condition, such that participants were presented with two scenarios in total. For participants in the *deficit-oriented label* and *deficit-oriented label + context* conditions, this allowed me to explore their interpretations of the *underprepared* label, in addition to the *at-risk* label. However, I was not able to identify another label that could be used in the second scenario for the *neutral label* condition, so participants in this condition read two scenarios in which the advisor characterizes the student as a *first-generation* student. Lastly, I continued to explore the influence of students' motivational beliefs by examining the potential influence of students' academic mindsets, racial and ethnic identity beliefs, and stereotype vulnerability on their interpretations of both deficit-oriented and neutral labels, and the perceived consequences of being referred to with these labels.

Predictions. Given the previous findings, the addition of an extra condition, and the inclusion of students' motivational beliefs as potential moderators of the labeling manipulation, I had several new hypotheses for this portion of the study. First, I expected to find a general main effect of *label condition* for the valence of the advisor's beliefs, such that I expected that LIFG and NLIFG students in the *neutral label* and the *deficit-oriented label + context* conditions

would interpret the advisor's feedback across both scenarios as communicating less negative beliefs than the participants in the *deficit-oriented* label condition. I also expected that LIFG and NLIFG students in the *deficit-oriented* label condition would perceive the student in those scenarios as reporting negative affective and motivational consequences to a greater extent than participants in the other two conditions. I did not expect a main effect of label condition on participants' interpretations of the advisor's intention to communicate their beliefs about the student when providing feedback. However, based on the findings from Study 1, I did expect to find an interaction between *sample group* and *label condition* for several of the dependent measures. More specifically, I expected that for participants in the *neutral label* and *deficit-oriented + contexts* conditions, relative to NLIFG students, LIFG students would interpret the advisor's feedback as (a) more positive, (b) more intentional, and (c) less negative in its consequences for the hypothetical student. These predictions were based on my intuition that because LIFG students would be more likely to have experienced being academically stigmatized in the past, it is possible that they might also be more likely to pick up on subtle differences between negative characterizations of their academic competence that center on their personal inadequacies, versus characterizations that attribute academic struggles to structural factors beyond their control. In contrast, because NLIFG students are less likely to have experienced chronic academic stigmatization, these students may not necessarily perceive the advisor's elaboration of why the hypothetical student is considered as 'at-risk' or 'underprepared' as an attempt to shift the focus towards social injustices and away from the student. With respect to the *deficit-oriented* condition, based on the findings from Study 1, I did not expect to find any sample group differences between LIFG and NLIFG students'

interpretations of the advisor's feedback or in their perceptions of the effects the hypothetical students in those scenarios might experience as a result of their interaction with their advisor.

Lastly, I also had several predictions with respect to the moderating role of participants' motivational beliefs. First, although it is possible that academic mindsets might influence participants' *interpretations* of the labels themselves, there is evidence to suggest that these beliefs influence students' *responses* to stigmatizing experiences, but not their perception and interpretation of those experiences (Aronson, Fried, & Good, 2002). Therefore, although I did not expect to find differences in participants' interpretations of the labels based on their academic mindsets, I did expect that these beliefs would moderate the extent to which participants indicated the hypothetical students would experience negative consequences as a result of the advisor's feedback. Given that students who endorse a growth mindset perceive their academic outcomes as being within their control to a greater extent than those with weaker growth mindsets, I expected that participants' perceptions of the effects the hypothetical student would experience would be moderated by students' academic mindsets, but only in the *deficit-oriented labeling* condition. Specifically, I expected that participants with stronger growth mindsets would indicate that the hypothetical student would experience a lesser degree of negative affective and motivational consequences, relative to those with weaker growth mindsets.

In contrast to my predictions for academic mindsets, I expected that the influence of participants' stereotype vulnerability and racial and ethnic identity beliefs would differ between sample groups and label condition, such that I only expected these beliefs would influence LIFG students' perceptions of the effects the hypothetical student would experience in the *deficit-oriented label* condition. Specifically, I expected the LIFG students who showed a lesser degree

of vulnerability to negative stereotypes and those whose race and/or ethnicity was more central to their identity as college student, would also be more likely to indicate that the hypothetical students would experience a lesser degree of negative consequences, relative to LIFG students who showed greater vulnerability to stereotypes and weaker racial and ethnic identity beliefs. However, I did not expect these beliefs to influence LIFG students in the *neutral label* and *deficit-oriented label + context* conditions. Moreover, given the lack of alignment in the background characteristics between the NLIFG students and the hypothetical students described in the scenarios, I did not expect NLIFG students' stereotype vulnerability or their racial identity beliefs to show this moderation effect in any of the label conditions. Had the background characteristics of the hypothetical students been more consistent with those of NLIFG participants (allowing these participants to more easily take the perspective of the students, as they were instructed to do), I would have expected their stereotype vulnerability to moderate these differences, but perhaps to a lesser degree than LIFG students.

Exploring potential stereotype threat effects. Another portion of Study 3 was designed to explore potential stereotype threat effects associated with experiences of being labeled by a deficit-oriented descriptor, as well as the extent to which these effects might be moderated by their motivational beliefs. This possibility was examined by randomly assigning participants to spend three minutes describing an actual or hypothetical experience in which they were labeled as an *at-risk student*, or as a *first-year student*, and then directing participants to complete a verbal test immediately following the manipulation task. These specific labels were chosen because the *at-risk* label seems to be more pervasive than other deficit-oriented labels (e.g., *underprepared*; *disadvantaged*), and the *first-year student* label would apply to all college students at one point or another. Further, given that I could not ensure that all participants had

experienced being labeled as *at-risk* at some point in the past—nor was this a likely possibility—participants who had not experienced this in the past were instructed to describe a hypothetical experience instead. This was certainly not the ideal approach, particularly given that if a disproportionate number of participants were to describe a hypothetical experience, it could reduce the overall effectiveness of the manipulation. However, it was the best option I was able to identify in lieu of excluding participants from this task altogether—which would have likely reduced my sample size dramatically.

Predictions. Based on previous findings within the stereotype threat literature, I expected to find a main effect of *label condition* on participants' verbal scores, such that relative to participants assigned to describe an experience in which they were labeled as a *first-year student*, those asked to do the same but with respect to being labeled as an *at-risk student* would experience a greater degree of stereotype threat and score lower on the verbal test. I expected this would be the case due to the negative connotations of the *at-risk* label, which might be enough to induce stereotype threat in all participants, regardless of their backgrounds or motivational beliefs. However—given the negative stereotypes associated with students of color, low-income students, and first-generation students—I also expected to find an interaction between *sample group* and *label condition*, such that LIFG students in the *at-risk* label condition would experience a greater degree of stereotype threat compared to NLIFG students in the same condition.

Moreover, I also expected students' academic mindsets and stereotype vulnerability to moderate these effects for both LIFG and NLIFG students, whereas I expected that racial and ethnic identity beliefs would only moderate these effects for LIFG students. More specifically, I predicted that for participants in the *at-risk* label condition, those with stronger growth mindsets

and those with lower stereotype vulnerability would experience a lesser degree of stereotype threat, and thus, have higher scores on the verbal test, relative to those with weaker growth mindsets and greater vulnerability to stereotypes. Lastly, I expected that LIFG students whose race and or ethnicity are more central to their identities as college students, would experience a lesser degree of stereotype threat and thus, have higher scores on the verbal test, relative to LIFG students whose race and ethnicity are less central to their identities as college students.

In the remainder of this chapter, I describe the methods used to collect the data for *Study 3*, including specific details regarding the recruitment of participants, the survey used (including all prompts, items, and measures), and the procedure participants followed to participate in the study. Following this, I describe the analyses conducted on the data and the results yielded from those analyses. Finally, I end with a summary and discussion of the findings.

Methods

Inclusion/Exclusion Criteria

Prospective participants were deemed eligible to participate in Study 3 if they had not participated in Studies 1 and 2, and met the criteria required to be categorized into the LIFG or NLIFG sub-sample, which was the same criteria used in the two previous studies.

Participants

Final Sample. The final sample for Study 3 was N= 274 (LIFG, n=108; NLIFG, n= 166) and consisted of undergraduate students recruited from Mechanical Turk (n=113; 82 LIFG; 31 NLIFG) or emails to undergraduate course listservs at Boston College (n=161; 26 LIFG; 135 NLIFG). The participants in the final sample were 54% female, between the ages of 18-31 (M= 21, SD= 2.59), primarily native English speakers (94%) and a little over half of the sample indicated they were lowerclassmen (59%). Participants reported they were majoring primarily in

the sciences (30%), social sciences (25%), humanities (16%), and education (12%). Eighty-two percent of students indicated that they were attending a private university or college and 79 percent indicated that their institution was religiously affiliated.[19] Approximately 76 percent of the LIFG sub-sample identified as Hispanic or Latino and 40 percent identified as solely or partly Black or African American. One hundred percent of the participants in the NLIFG sub-sample identified as non-Hispanic White. Lastly, based on the differences in how the majority of students from each sample group were recruited (see below), I conducted an independent *t*-test to examine group differences in age. In addition to this, I also examined sample differences in participants' responses to the subjective SES question on the demographic's questionnaire (i.e., "How financially well off were you growing up?"), as was done in the prior two studies. The results of these analyses indicated that (a) on average, participants in the LIFG sample group were older than participants in the NLIFG sample group (M= 22.04, SD= 2.83 ; M= 20.39, SD= 2.20, respectively), $t(272) = 5.40$, p= .020, and (b) on average, LIFG participants perceived themselves as being significantly less "well off" (M=2.03; SD=.81) than NLIFG students (M= 3.14, SD= 1.30), $t(272) = 8.75$, p< .001 (see *Table 5.1* for additional demographic data for each sample group by recruitment method).

Exclusion of cases from final sample. A total of 2,052 prospective participants consented to participate in this study (Mechanical Turk workers n=1,113; Boston College undergraduate students n=939) and completed the pre-screening questionnaire, from which 473 met the inclusion criteria for one of the two sample groups (152 LIFG; 321 NLIFG) and completed some portion of the study. Of those, 199 participants were removed from all subsequent analyses due to incomplete data (n=43), because they completed the study in less

[19] It is important to note that these percentages were higher than in the two previous studies, but that it was likely due to the fact that a large proportion of the sample was recruited from Boston College.

than four minutes (n=145), or because they exceeded the age cutoff for the study of 32 (n=11), which was the same as the threshold used in Study 2.[20] Interestingly, in this study, there were no blatant issues concerning survey bot software use—therefore, no responses were flagged and/or removed for this reason.

Table 5.1. Summary of Study 3 demographic information for the full sample and sample groups by recruitment method.

Sample	Age (*M*)	Sex *(Male)*	Hispanic	Black	White	Upperclassmen	Public Institution	Religious Affiliation
Full Sample (*n*= 274)	21.04	46.4	29.9	15.7	77.7	41.3	29.9	79.2
LIFG *MTurk* (*n*= 82)	22.85	51.2	75.6	39.0	46.3	47.6	72.0	61.0
LIFG *BC* (*n*= *26*)	19.46	50.0	76.9	42.3	34.6	26.9	0[21]	100
LIFG Total (*n*= 108)	22.04	50.9	75.9	39.8	43.5	42.6	56.5	70.4
NLIFG *MTurk* (*n*= 31)	23.23	74.2	0	0	100	74.2	67.7	22.6
NLIFG *BC* (*n*= 133)	19.74	36.3	0	0	100	32.6	0	100[22]
NLIFG Total (*n*= 166)	20.39	43.4	0	0	100	40.4	12.7	84.9

Note: All values in columns 4-9 are expressed in percentages.

[20] The age cutoff was 32 but there were no participants who reported that age, therefore the maximum age for the final sample was 31.

[21] 7.7% of participants from this group indicated that they attended a public institution, which likely reflects a mistake on participants' part when completing the survey given that they were all recruited from Boston College, which is a private institution.

[22] Only 99.3% of participants from this group indicated that they attended a religiously-affiliated institution, which likely reflects a mistake on participants' part when completing the survey given that they were all recruited from Boston College, which is a religiously affiliated institution.

Recruitment

Participants for Study 3 were initially recruited from Boston College (BC) via emails to course listservs and the remainder of the sample was subsequently recruited through Mechanical Turk.[23] Given that I was only able to recruit a limited number of LIFG participants from BC, the majority of the LIFG sample were recruited from Mechanical Turk, whereas the majority of the NLIFG sample were recruited from BC. Although the quotas for both sample on Qualtrics were the same, a considerable number of NLIFG students were completing the survey simultaneously and Qualtrics does not adjust the quotas until participants complete the survey. Therefore, it was difficult to track the number of NLIFG students that had completed the survey in order to keep the sample groups proportionate. However, NLIFG students were also recruited from MTurk, but funding and timing constraints limited my ability to continue the recruitment process in order to achieve proportionate sample groups per recruitment method.

Course/Organization Listservs. The recruitment emails sent to course listservs contained general information about the study—including the pre-screening process, the study survey, and compensation—as well as a link that students used to access the study on Qualtrics.

Mechanical Turk. Participants who were recruited for this online study via Mechanical Turk were recruited in the same manner as participants in the two previous studies.

Compensation

Participants' were compensated for their time depending on how they were recruited. Participants who were recruited through emails to course listservs were given the opportunity to enter in a raffle to win one of two $75 Amazon gift cards. Participants were asked to provide

[23] Recruitment emails were also sent to faculty and instructor-managed course listservs at Stanford University and to listservs associated with social organizations at over 50 colleges and universities in the US, but these strategies were not successful.

their Boston College affiliated email address that could be used to contact them if they won of the raffles. Participants who are recruited through Mechanical Turk were compensated in the same manner as the two previous studies.[24]

Materials & Measures

Hypothetical scenarios. Study 3 used hypothetical scenarios similar to the those used in Study 1, such that each one described an interaction between a hypothetical low-income, first-generation student and their advisor. However, in this study, participants were randomly assigned to one of three conditions—a *deficit-oriented label* condition; a *neutral label* condition; or a *deficit-oriented label + context* condition. For all three conditions, participants were presented with two hypothetical scenarios, each of which described an interaction between a hypothetical low-income, first-generation student and their advisor. The order in which the scenarios were presented did not vary across conditions or participants, and the sex of the hypothetical student in each scenario was matched to participants' sex based on their response to the item on the pre-screening questionnaire that asked them to indicate their 'biological sex' (i.e., male; female).

In each scenario, the advisor provided feedback to the student regarding his [her] poor academic performance that semester, and uses a label to characterize the student, which is the portion of the scenarios that differed across conditions. More specifically, in the *deficit-oriented label* condition, participants were presented with the same scenario used in the *deficit-oriented label* condition in Study 1 and one additional scenario, in which the advisors use a deficit-oriented label to characterize the student (*at-risk student* in *Scenario 1*; *underprepared student* in

[24] To accommodate the different methods of compensation, this study was administered using two different copies of the Qualtrics survey, one for participants who were recruited through Mechanical Turk and one for participants who were recruited through emails to course listservs.

Scenario 2) and provide no clarification or elaboration to the student as to the reasons why they would be characterized in that manner. In the *deficit-oriented label + context condition*, participants were presented with two scenarios that were identical to the scenarios in the *deficit-oriented label* condition with one exception—after the advisor characterized the student by a deficit-oriented descriptor (*at-risk student* in *Scenario 1* and *underprepared student* in *Scenario 2*) they provided additional context for the label by explaining to the student (e.g., 'that *at-risk* [underprepared] students sometimes struggle because they haven't had access to the same types of educational resources/opportunities as his [her] peers'; see Appendix H for wording of context in both scenarios). Lastly, in the *neutral label* condition, participants were presented with the same scenario used in the *neutral label* scenario in Study 1 and one additional scenario in which the advisor characterized the student with a neutral label when providing the student feedback on his [her] performance. For this condition, the label *first-generation student* was used as the neutral label in both scenarios (see Appendix H for the exact wording of all scenarios).

Interpretations of deficit-oriented labels versus a neutral label. Students' interpretations of the advisor's feedback, including the type of label used in the feedback were assessed using items that measured students' perceptions of the valence of beliefs held by the advisor about the hypothetical student and the intention on the part of the advisor in communicating those positive or negative beliefs. These items were the same across all scenarios and label conditions, and only varied in terms of the hypothetical student referenced in the items, which matched the name of the student in the scenario.

Valence of advisor's beliefs. Participants' perceptions about the extent to which the advisor's feedback communicated positive or negative beliefs about the student was measured with one item in each scenario (*S1_valence*, *S2_valence*; i.e., "To what extent to do you believe

that the feedback from [hypothetical students' name] advisor is communicating positive or negative beliefs about [hypothetical students' name]?"), which included a 6-point Likert-type response scale that ranged from 1 = "Very Negative Beliefs" to 6 = "Very Positive Beliefs." Participants' responses to both *S1_valence* and *S2_valence* were reverse coded, such that higher values on each item indicated a greater degree of negative beliefs and a lesser degree of positive beliefs (*S1_valence*: *M*=3.59, *SD*= 1.37; *S2_valence*: *M*=3.47, *SD*= 1.48).

Advisor's intentionality in communicating beliefs. In each scenario, two items were used to assess participants' perceptions about the extent to which the hypothetical student's advisor was using their feedback (which included the label) to intentionally communicate their *positive* or *negative* beliefs about the hypothetical student. The first intentionality item for each scenario (*S1_intentionality_1; S2_intentionality_1)* asked participants to what extent they "believed that the hypothetical student's advisor was intentionally communicating their positive or negative beliefs about the student with the feedback they provided them." This item used a 5-point Likert type scale from 1 = "Not at All Intentionally" to 5 = "Very Intentionally". The second intentionality item (*S1_intentionality_2; S2_intentionality_2),* was presented as a statement (e.g., "Aaron's [April's] advisor didn't think much about how he [she] would work his [her] feedback to Aaron [April]"), to which participants were asked to indicate their level of agreement with, on a 6-point Likert type scale from 1= "Strongly Disagree" to 6= "Strongly Agree". Participants' responses to the second intentionality item were reverse coded to align with the first one, such that higher values for *intentionality_1* and *intentionality_2* across both items reflected a greater degree of intentionality. Given that the items were on different scales, they were both items standardized and then averaged to form one intentionality score per scenario.

The intentionality score for each scenario had a mean of zero. Values for the intentionality score for *Scenario 1* ranged from –1.39 to 1.89 and scores for *Scenario 2* ranged from –1.69 to 1.69.

Perceptions of the affective and motivational consequences of being labeled by a deficit-oriented or neutral label. Students' perceptions of the affective and motivational consequences that the hypothetical student described in each scenario might experience as a result of their labeling experience were measured using the same 11 effect items used to measure the affective and motivational consequences experienced by students' themselves in Study 2. However, given that in Study 3 students were presented with two scenarios and were asked to respond to these items for each scenario, each student had data for two sets of 11 effect items. The only difference between the items used in this study and those used in Study 2 were that the items were framed in terms of the hypothetical students described in the scenarios (e.g., "Being labeled as at-risk, and/or underprepared, and/or disadvantaged student probably made Aaron [April] feel like he [she] is not a valued member of his [her] university's community."; see *Table 4.2* for full list of items). Students' were asked to indicate the extent to which they agreed with each item on a 6-point Likert-type scale (1= "Strongly Disagree" to 6= "Strongly Agree"). Participants' responses on these items were then used to create the same subscale scores for sense of belonging, academic self-perceptions, and affect, for each scenario (i.e., *S1_SOB, α=.67*; *S2_SOB, α=.74*; *S1_ASP, α=.83*; *S2_ASP, α=.87*; *S1_AFF, α=.78*; *S2_AFF, α=.87*), that were created in *Study 2*. However, unlike the previous study, the items used to create the academic engagement (*AE*) subscale score (i.e., *item 8* and *9)* were only correlated in *Scenario 2*, $r(274) = .18$, $p = .002$, but not for *Scenario 1*, $r(274) = .01$, $p = .862$. Therefore, these items were used individually in subsequent analyses.

Stereotype threat manipulation. In order to explore potential stereotype threat effects of students' deficit-oriented labeling experiences, participants completed a task that required them to describe an experience in which they were labeled by a deficit-oriented label or a neutral label. In the *deficit-oriented label* condition, participants were asked to recall and briefly describe an experience in which they were labeled as an *at-risk student*, whereas participants assigned to the *neutral label* condition were asked to recall and describe an experience in which they were labeled as a *first-year student*. In both prompts, participants were instructed to imagine and describe a hypothetical experience if they had not experienced being labeled as an *at-risk* [*first-year*] student in the past. Participants were asked to take approximately 2-3 minutes to briefly describe this experience in writing (see Appendix H for wording of prompt).

Verbal test. Once participants completed the threat manipulation, they were asked to complete a brief analytical task that consisted of 15 GRE-type analogy questions (see Appendix H for sample test). This task was similar in nature to the types of tests used to assess performance in prior stereotype threat research conducted with college students (Steele & Aronson, 1995). Participants were given three minutes to complete the test prior to being automatically redirected to the next section of the study. Correct responses were coded as *1* and incorrect or missing responses were coded as *0*. A single *verbal_score* was computed by summing across all 15 items, such that scores could range from 0-15, with higher scores indicating better performance on the test (M= 4.82, SD= 2.72, α=.68).

Personal goals for academic performance in college. The extent to which participants valued doing well in college was assessed with the following item: "Doing well in college is very important to me" (1= "Strongly Disagree" to 6= "Strongly Agree"). This item was embedded within the demographics questionnaire that was presented at the end of the study and was used as

a covariate in analyses testing for stereotype threat effects (M= 5.46, SD = .75). The results of an independent samples t-test also showed there were statistically significant between-group differences in participants' mean response to this item, $t(272) = 2.01$, p = .046, d= .26, such that on average, LIFG students agreed with the statement "Doing well in college is very important to me" to a significantly lesser degree than NLIFG students (M= 5.35, SD= .70 ; M= 5.54, SD= .77, respectively).

Students' academic mindsets. Students' academic mindsets were assessed using the same 8-item *Implicit Theories of Intelligence Scale* by Dweck (1999; see Appendix I for list of items) that was used in Study 2. A single mindset index (M = 4.06, SD = .96, α = .86) was computed for each participant in the same manner as the previous study. It is important to note that for the analyses described in the *Results* section, participant's standardized *mindset scores* were used in lieu of the original variable, such that the index was centered at the mean, with a standard deviation of 1.

Students' sensitivity to negative racial stereotypes about academic competence. The same 8-item *Stereotype Vulnerability Scale* (Barnard et al., 2008) used in Study 2 was used to assess the extent to which students' feel threatened by negative stereotypes about the academic competence of students who belong to their racial or ethnic group (see Appendix I for list of items). A single index stereotype vulnerability index (M = 2.93, SD = .94, α = .76) was computed for each participant in the same manner as the previous study. As with participants' *mindset scores*, for the purposes of the primary analyses, participant's standardized *stereotype vulnerability scores* were used in lieu of the original variable.

Students' racial identity attitudes. Students' racial identity beliefs were measured using the same seven items used in Study 2, which were adapted from the 8-item *Centrality* subscale of

the *Multi-Dimensional Black Identity Scale* (Sellers et al., 1997; see Appendix I for list of items). A single index (M = 3.24, SD = .96, α = .80) was computed for each participant in the same manner as the previous study. As with both the *mindset* and *stereotype vulnerability scores*, participant's standardized *identity belief scores* were used in lieu of the original variable for the primary analyses.

Procedure

The procedure for Study 3 was similar to the procedure for both Study 1 and 2. Once prospective participants were redirected to the study on Qualtrics research suite, the procedure for participating in the study was the same for all participants, regardless of the recruitment method. First, they were presented with an electronic version of the consent form and were required to give their consent to participate before being re-directed to pre-screening questionnaire. Participants who indicated that they did not give their consent to participate, were redirected out of the study. Next, prospective participants completed the 11-item pre-screening questionnaire. Participants who were not eligible to be categorized into either of the sub-samples were redirected out of the study at that point. Participants who were eligible to be categorized into one of the two samples groups were presented with a message informing them of their eligibility and were given the opportunity to continue on to the study survey. Participants then completed *Sections 1-3* in order—followed by the same demographic's questionnaire used in the two previous studies. Lastly, participants were presented with the same electronic version of the debriefing form used in the two previous studies (see Appendix E). Participants who were recruited via Mechanical Turk were also presented with a completion code on this page, whereas participants recruited directly from their institution were redirected to a separate page and asked

to enter their name and university affiliated email address in order to be entered into the raffle for the gift cards.

Results

Students' Motivational Beliefs

Associations between students' academic mindsets, stereotype vulnerability, and racial identity beliefs. Two sets of correlational analyses were conducted to examine associations between students' academic mindsets, stereotype vulnerability beliefs, and racial identity beliefs, by sample group. As seen in *Table 5.2*, the results of the analysis using LIFG students' scores yielded statistically significant correlations between their *mindset* and *stereotype vulnerability scores* and their *stereotype vulnerability* and *racial identity belief scores*, but not their *mindset and racial identity belief scores*. The results of the analysis using NLIFG students' scores yielded statistically significant correlations between their *mindset* and *stereotype vulnerability scores* and their *stereotype vulnerability* and *racial identity belief scores*, but not their *mindset* and *racial identity belief scores*.

Table 5.2. Results of correlational analyses by sample group examining associations between students' motivational beliefs.

Motivational Beliefs	*1*	*2*	*3*
1. Racial Identity Beliefs	—	.23*	–.07
2. Stereotype Vulnerability	.54*	—	–.17*
3. Academic Mindsets	.18	.22*	—

Note. Correlations for LIFG students (n= 108) are presented to the left of the diagonal and correlations for NLIFG students (n= 166) are presented to the right of the diagonal.
* denotes a statistically significant r $(p < .05)$.

Between-group differences in students' motivational beliefs. Potential between-group differences in students' mean scores on all three motivational belief measures were examined

through a series of independent *t*-tests. The results of these analyses yielded statistically significant between-group differences in students' mean *stereotype vulnerability scores* and their *racial identity belief scores*, but not in their *mindset scores*. More specifically, the results indicated that on average, LIFG students' were significantly more vulnerable to negative racial stereotypes about academic competence (M=3.75, SD=.66) than NLIFG students (M=2.40, SD=.68), $t(272) = 16.39$, $p < .001$, d= 2.02, and that their racial identity beliefs played a more central role in their identity as college students than that of NLIFG students (M=3.84, SD=.89; M=2.84, SD=.78, respectively), $t(272) = 9.81$, $p < .001$, d= 1.20. There difference between LIFG and NLIFG students' mindset scores was not statistically significant (M=4.02, SD=.82; M=4.09, SD=1.05, respectively), $t(272) = .57$, p=.567, d= .07. Moreover, it is important to note that the mean mindset scores for LIFG and NLIFG students were significantly higher than the midpoint of the scale (i.e., 3.5), $t(<165)s > 6.60$, $ps < .001$—which reflected these students' greater endorsement of a growth mindset and lesser endorsement of a fixed mindset.

Students' Interpretations of Hypothetical Students' Labeling Experiences

These analyses were designed to test my primary hypotheses regarding differences in students' interpretations of *deficit-oriented* labels (*at-risk student*; *underprepared student*) versus a *neutral* label *(first-generation student),* and their perceptions of the potential effects of being labeled by each of these descriptors. The first series of analyses used the *valence belief items* for each scenario as dependent measures to examine the extent to which participants believed the advisors held *positive* or *negative* beliefs about the students they were advising. The second series of analyses used the *intentionality scores* for each scenario to examine the extent to which participants believed the advisors were *intentionally* trying to communicate their positive or negative beliefs about the student through their feedback. Lastly, for the third set of analyses, I

used the *SOB*, *ASP*, and *AFF* subscale scores (along with the two *academic engagement items*) for each scenario as dependent measures, to examine participants' perceptions of the types of affective and motivational consequences the hypothetical students might experience as a result of the interaction with their advisor.

Given that this portion of the study relied on a repeated measures design and incorporated several continuous covariate measures, all of these analyses were conducted as mixed-measures ANCOVAs. The first series of models that were estimated (referred to as *Model 1* for each dependent measure) included *scenario* (first; second) as a within-subjects factor[25], and *sample group* (*LIFG*; *NLIFG*), *label condition* (deficit-oriented label; deficit-oriented label + context; neutral label), *gender* (male; female), and *recruitment method* (BC; MTurk) as between-subject factors. A standardized version of the variable for participants' *age* was included as a continuous between-subjects covariate, and all two and three-way interactions between scenario, *sample group*, *recruitment method*, *label condition*, *age*, and *gender* were also included. Following this, another series of models were estimated (referred to as *Model 2* for each dependent measure) that added participants' mean-centered *mindset scores*, *stereotype vulnerability scores*, and *racial identity belief scores* as continuous between-subjects covariates to *Model 1*, as well as all two and three-way interactions that were relevant to my hypotheses. The results of these models were only used to report any statistically significant main effects of participants' motivational beliefs and any significant interactions that included these beliefs *and* were relevant to my hypotheses. All of the statistically significant interactions that were relevant to the hypotheses being tested in this study were probed further through pairwise comparisons based on estimated marginal

[25] For all analyses, the same neutral scenario was always paired with the first *deficit-oriented label* or *deficit-oriented label + context* scenario (i.e., *at-risk* label); and the other neutral scenario was always paired with the second *deficit-oriented label* or *deficit-oriented label + context* scenario (i.e., *underprepared* label).

means. It is also important to note that all pairwise comparisons for the *label condition* variable were conducted with a SIDAK adjustment to account for the three different comparisons between the three conditions.

The results reported in the sub-sections below are those relevant to the hypotheses being tested for this dissertation. However, the full results for both *Model 1* and *2* for each dependent measure are reported in the following tables: *Table 5.3* for *valence belief ratings* and *intentionality scores; Table 5.4* for effects on *academic self-perceptions* (*ASP*)*, sense of belonging in college* (*SOB*), and *affect* (*AFF*); *Table 5.5* for effects on *academic engagement* (*AE*) *item 8* and *item 9*.

Perceived valence of advisor's beliefs.

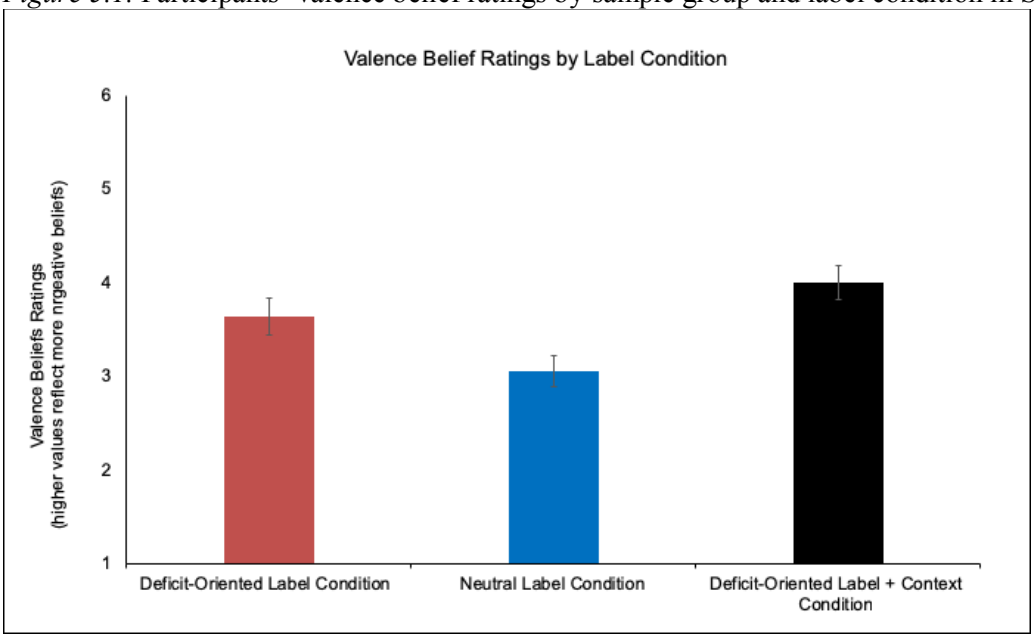
Model 1. The results of the initial analysis using the *S1_valence* and *S2_valence* items as dependent measures did not yield a significant main effect of *scenario*, but did yield significant main effects of *recruitment method* and *gender*, such that on average, participants recruited through MTurk had significantly less negative belief ratings than participants recruited from BC (M= 2.99, SE=.13; M= 3.86, SE= .12, respectively), and male participants had significantly less negative belief ratings than female participants (M= 3.25, SE=.10; M= 3.60, SE= .10, respectively). There was also a marginally significant main effect of *age*, that was probed through marginal means estimated at 1 *SD* above and below the mean age of the sample; on average, older participants had significantly less negative belief ratings than younger participants (M= 3.29, SE=.11; M= 3.56, SE= .12, respectively).

Importantly, the analysis also yielded significant main effects for *sample group* and *label condition*. With respect to *sample group,* LIFG students' mean valence rating across scenarios and label conditions (M= 3.16, SE=.11) was significantly less negative than NLIFG students'

mean rating *M*= 3.69, *SE*= .10). Additionally, the results from a series of single sample *t*-tests revealed that LIFG students' ratings for four of the six scenarios were significantly lower than the midpoint of the scale, which indicated that LIFG students generally perceived the advisor as communicating relatively positive beliefs, *ts* > 1.97, *ps* < .012. In contrast, NLIFG students' ratings for five of the six scenarios were significantly higher than the midpoint of the scale, *ts* > 2.51, *ps* < . 016, indicating that they perceived the advisor as communicating relatively negative beliefs across the majority of the scenarios.

As depicted in *Figure 5.1*, the pairwise comparisons by *label condition* showed that on average, for participants in the *neutral label* condition, their mean valence belief rating (*M*= 3.05, *SE*=.17) was significantly lower than the mean ratings for participants' in the *deficit-oriented label* and *deficit-oriented label + context* conditions (*M*= 3.64, *SE*=.20; *M*= 4.00, *SE*= .18, respectively, *ps* < .05), whereas the mean ratings between the *deficit-oriented label* and *deficit-oriented label + context* conditions did not significantly differ (*p*= .380).

Figure 5.1. Participants' valence belief ratings by sample group and label condition in Study 3.



Although the interactions between *sample group* × *label condition* and *scenario* × *sample group* were nonsignificant, this analysis did yield several statistically significant or marginally significant interactions. First, there was a significant *label condition* × *scenario* interaction. Pairwise comparisons by *label condition* showed that for *Scenario 2*, the mean *valence belief* rating for participants in the *neutral label* condition (M= 2.76, SE=.13) was significantly lower than the mean rating for participants in the *deficit-oriented label* and *deficit-oriented label + context* conditions (M= 3.77, SE=.15; M= 3.53, SE=.14, respectively, $ps < .001$), whereas the mean ratings between the *deficit-oriented label* and *deficit-oriented label + context* conditions did not significantly differ (p= .505). In contrast, the differences in mean ratings between the *neutral label*, *deficit-oriented label* and *deficit-oriented label + context* conditions for *Scenario 1* were not statistically significant (M= 3.30, SE=.13; M= 3.54, SE=.15; M= 3.65, SE= .14, respectively, $ps > .150$). The pairwise comparisons by *scenario* showed that the mean ratings for *Scenario 1* and *2* were significantly different for participants in the *neutral label* ($p < .001$), but not for the *deficit-oriented label* and *deficit-oriented label + context* conditions ($p > .130$).

Lastly, there was a statistically significant *label condition* × *age* interaction, which was further probed through pairwise comparisons based on marginal means estimated at 1 *SD* above and below the mean age of the sample. These analyses indicated that for participants who were 1 *SD* below the mean age of the sample, those assigned to the *neutral label* condition (M= 3.06, SE=.17) gave lower ratings than those in the *deficit-oriented label* and *deficit-oriented label + context* conditions (M= 3.64, SE=.20; M= 4.00, SE=.18, respectively; $ps < .047$), whereas the difference between the latter two conditions was not statistically significant ($p = .380$). For participants who were 1 *SD* above the mean age of the sample, participants in the *neutral label condition* (M= 3.00, SE=.16) had a mean rating that was significantly lower than the mean rating for participants in the *deficit-oriented label* condition (M= 3.67, SE=.16; $p = .005$) but not for those assigned to the *deficit-oriented label + context* (M= 3.19, SE=.18; $p = .802$). The difference in mean rating between those in the *deficit-oriented label* and *deficit-oriented label + context* conditions was marginally significant ($p = .093$).

Model 2. The results for the analysis that included participants' motivational beliefs yielded a marginally significant main effect of participants' *racial identity beliefs scores*. Marginal means were estimated at 1.5 *SD* above and below the mean racial identity beliefs score, which indicated that on average, participants with stronger racial identity beliefs had marginally less negative belief ratings (M= 2.83, SE= .16) than those with weaker racial identity beliefs (M= 3.24, SE= .16).

This analysis also yielded statistically significant interactions between *sample group* × *mindset score*, *label condition* × *mindset score*, and *sample group* × *stereotype vulnerability score*. The latter interaction was further qualified by a significant *3*-way interaction between *sample group* × *label condition* × *stereotype vulnerability score*. These interactions were probed

further by estimating marginal means and conducting pairwise comparisons. For the interactions between *sample group* × *mindset score* and *label condition* × *mindset score,* I estimated marginal means for mean valence belief ratings by *sample group* or *label condition*—each at 1.5 *SD* above and below the mean mindset score. For the interactions between *sample group* × *stereotype vulnerability score* and *sample group* × *label condition* × *stereotype vulnerability score,* I estimated marginal means at .5 *SD* above and below the mean stereotype vulnerability score. Given that the lowest stereotype vulnerability score for the LIFG sample was 2.13, which was less than 1 *SD* below the mean score, it would not have been sensible to estimate mean ratings at 1.5 *SD* below the mean vulnerability score (as with the mindset analyses).

With respect to the *sample group* × *mindset score* interaction (see *Figure 5.2*), for participants with weaker growth mindsets (i.e., 1.5 *SD* below the mean), LIFG students' belief ratings (M= 2.27, SE= .24) were significantly less negative than NLIFG students' ratings (M= 3.57, SE= .15, $p < .001$). For participants with mindset scores 1.5 *SD* above the mean (i.e., stronger growth mindsets), mean valence belief ratings between LIFG and NLIFG students were not significantly different (M= 3.01, SE= .30; M= 3.30, SE= .16, p= .387)).

For the *label condition* × *mindset score* interaction, the estimated means indicated that for participants with mindset scores 1.5 *SD* above the mean mindset score, the mean belief rating did not significantly differ between the *deficit-oriented label* (M= 3.66, SE= .29), *neutral label* (M= 3.00, SE= .22), or *deficit-oriented label + context* conditions (M= 2.80, SE= .34; $ps > .142$). For participants with mindset scores 1.5 *SD* below the mean score, the mean belief rating for participants in the *neutral label condition* was marginally lower (M= 2.65, SE= .21) than it was for participants in the *deficit-oriented label + context condition* (M= 3.36, SE= .26, p= .073), whereas the contrasts between *neutral label* and *deficit-oriented label conditions* (M= 2.75, SE=

.21; $p = .977$) and the *deficit-oriented label* and *deficit-oriented label + context conditions* ($p = .170$) were not statistically different (see *Figure 5.3*).

For the *sample group* × *stereotype vulnerability* score interaction (see *Figure 5.4*), the pairwise comparisons showed that for participants with stereotype vulnerability scores .5 *SD* above the mean, mean valence belief ratings across all scenarios and label conditions did not significantly differ between LIFG and NLIFG students ($M = 2.96$, $SE = .16$; $M = 3.19$, $SE = .14$, $p = .272$). In contrast, for participants with stereotype vulnerability scores .5 *SD* below the mean, LIFG students' mean valence beliefs rating was significantly lower than NLIFG students ($M = 2.31$, $SE = .27$; $M = 3.67$, $SE = .10$, respectively, $p < .001$). Two additional correlational analyses were conducted to examine the association between participants' stereotype vulnerability and their mean valence belief ratings for each scenario, by sample group. These analyses showed that for LIFG students, higher stereotype vulnerability scores were strongly associated with higher valence belief ratings for both scenarios, $r(106)s > .31$, $ps < .001$, whereas the associations for NLIFG students were equally strong but in the reverse direction, $r(164)s > -.31$, $ps < .001$.

Figure 5.2. Participants' valence belief ratings by sample group and academic mindsets.

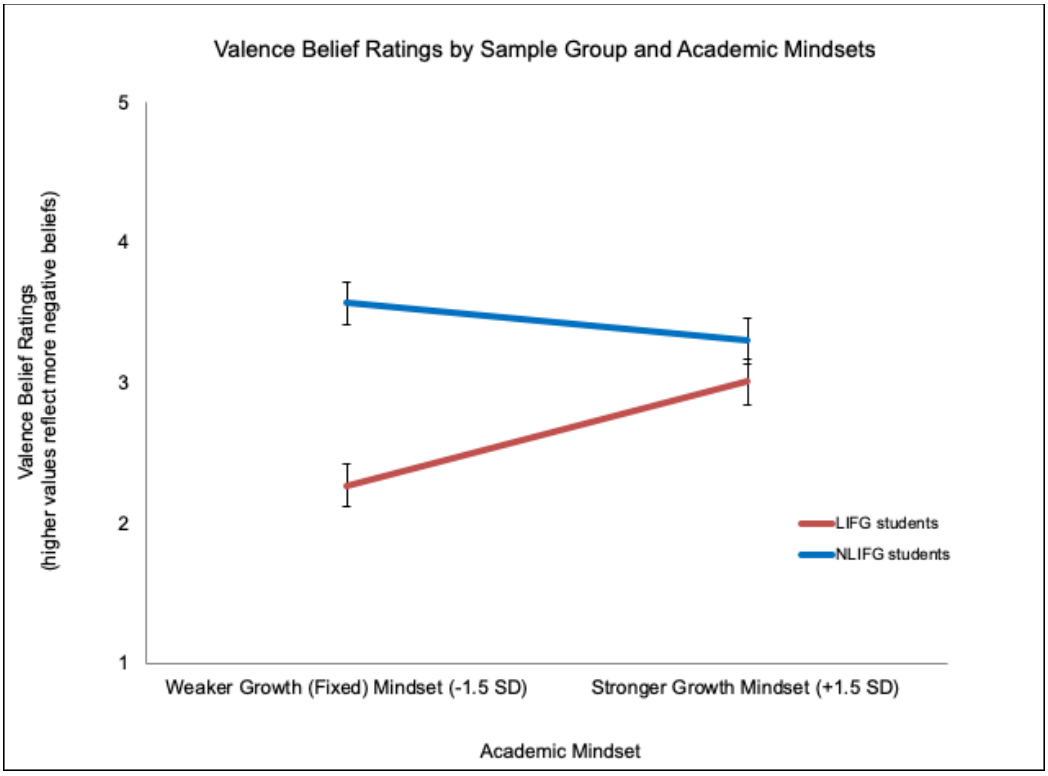


Figure 5.3. Participants' valence belief ratings by label condition and academic mindsets.

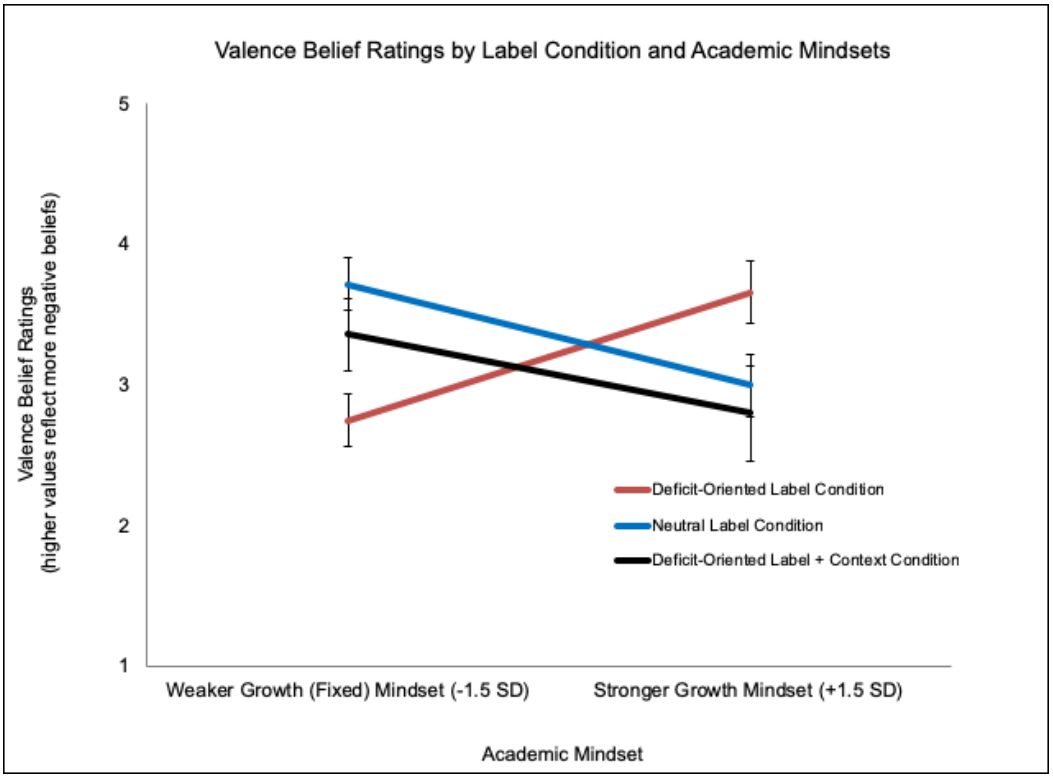
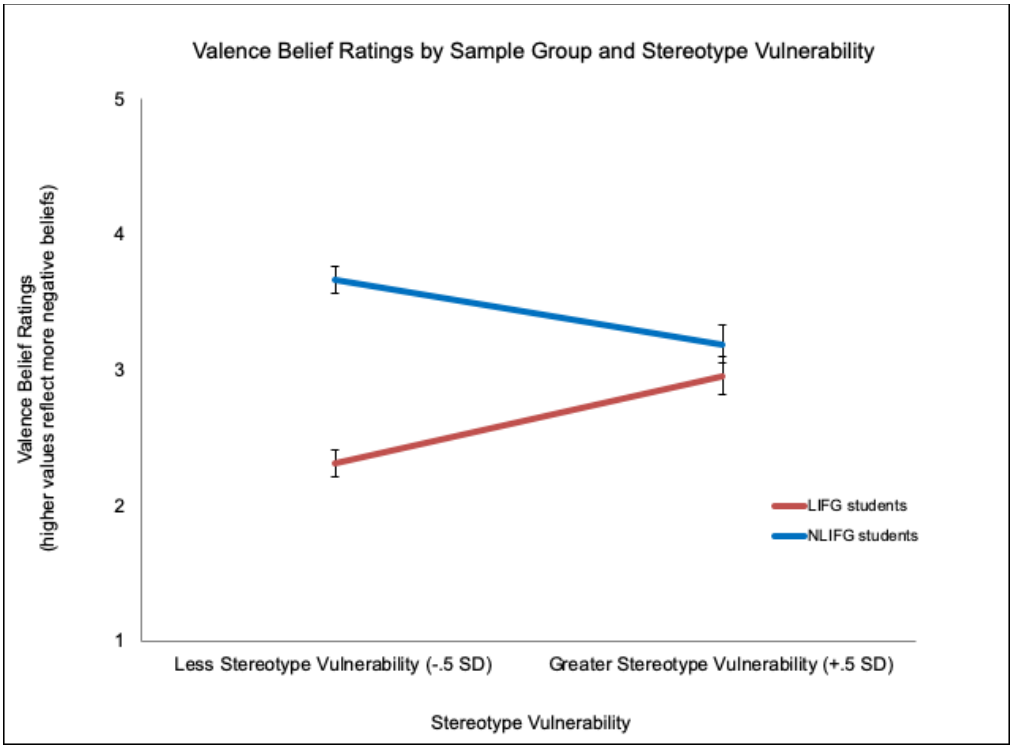


Figure 5.4. Participants' valence belief ratings by sample group and stereotype vulnerability.



Lastly, for the *sample group* × *label condition* × *stereotype vulnerability* score interaction, the pairwise comparisons for *sample group* indicated that relative to LIFG students, NLIFG students with stereotype vulnerability scores .5 *SD* above the mean (i.e., more vulnerable to stereotypes) had marginally higher mean valence belief ratings in the *deficit-oriented label* + *context condition* (M= 2.76, SE= .29; M= 3.48, SE= .24, respectively, p= .056). However, there were no significant differences between sample groups for the *deficit-oriented label* (M= 3.37, SE= .26; M= 3.35, SE= .25, respectively) or *neutral label* conditions (M= 2.76, SE= .18; M= 2.76, SE= .23, ps > .948). For participants with stereotype scores that were .5 *SD* below the mean, LIFG students' mean belief ratings were significantly lower than NLIFG students, when they were assigned to the *deficit-oriented label* (M= 2.14, SE= .38; M= 3.96, SE= .17) and the *deficit-oriented label* + *context* (M= 2.25, SE= .55; M= 3.85, SE= .16, ps < .006), and

marginally lower for those in the *neutral label* condition (M= 2.55, SE= .35; M= 3.21, SE= .15, respectively, p = .085). The pairwise comparisons for *label condition* indicated that LIFG students with stereotype vulnerability scores .5 *SD* above the mean, there were no significant within-group differences in participants' mean valence belief ratings between LIFG students in the *deficit-oriented label* condition (M= 3.37, SE= .26), *deficit-oriented label + context* (M= 2.76, SE= .29), or *neutral label* conditions (M= 2.76, SE= .18; *ps* > .127). The comparisons for NLIFG students were nonsignificant between those in the *deficit-oriented label* and *deficit-oriented label + context conditions* (M= 3.35, SE= .25; M= 3.48, SE= .24, respectively) and those in the *deficit-oriented label* and *neutral label* condition (M= 3.35, SE= .25; M= 2.76, SE= .23, respectively; *ps* > .220), but marginally significant for those in the *neutral label condition* compared to those in the *deficit-oriented label + context* condition (M= 2.76, SE= .23; M= 3.48, SE= .24, respectively; p = .086). For LIFG students with stereotype vulnerability scores .5 *SD* below the mean, their mean valence belief ratings did not differ significantly between those in the *deficit-oriented label* condition (M= 2.14, SE= .38) compared to those in the *deficit-oriented label + context* (M= 2.25, SE= .55) and *neutral label* conditions (M= 2.55, SE= .35; *ps* > .791). However, NLIFG students in the *neutral label* condition (M= 3.21, SE= .14) had mean valence belief ratings that were significantly lower than those in the *deficit-oriented label* (M= 3.96, SE= .17) and *deficit-oriented label + context* conditions (M= 3.85, SE= .16; *ps* $<$ .007), but the mean rating did not differ between those in the *deficit-oriented label* and *deficit-oriented label + context* conditions ($p = .928$; see *Figure 5.5*).

Figure 5.5. Significant *3*-way interaction between sample group, label condition, and stereotype vulnerability on valence belief ratings.

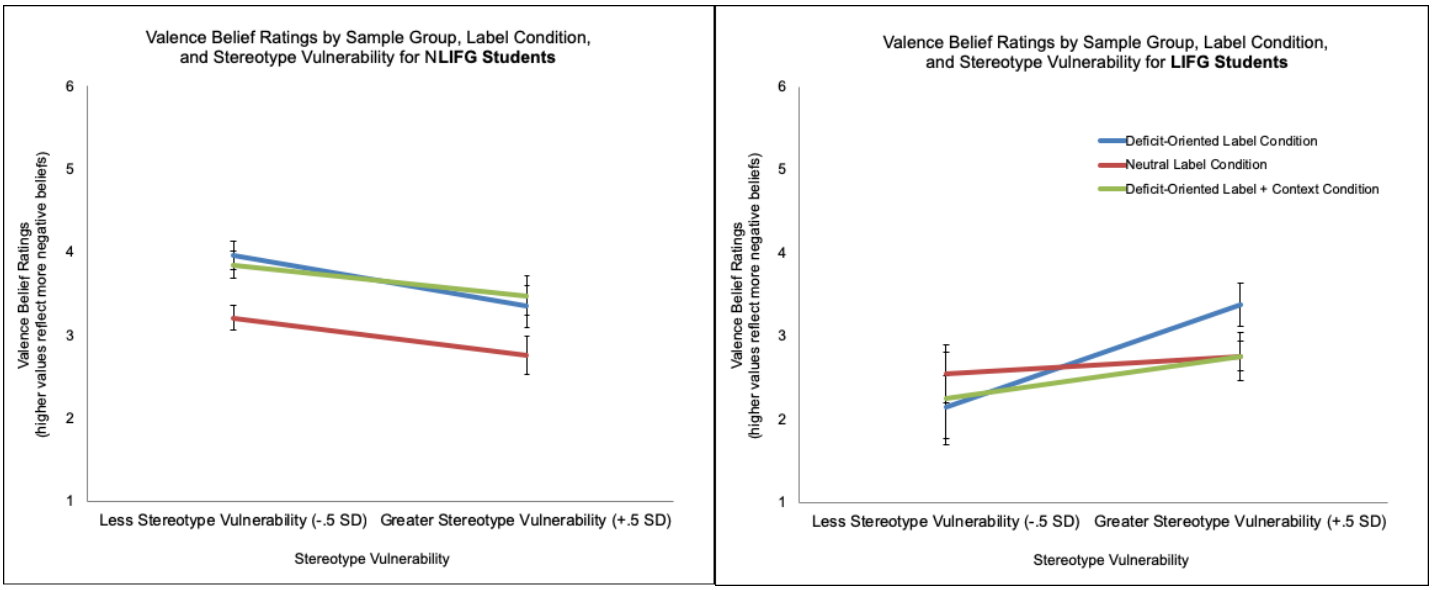


Table 5.3. Summary of ANCOVA results for participants' valence belief ratings and intentionality scores.[26]

	Valence Belief Ratings			*Intentionality Scores*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Scenario (SC)[1]	2.32	.129	.01	.59	.445	<.01
Sample (S)[1]	11.73	.001	.05	.83	.363	<.01
Recruitment (R)[1]	21.67	<.001	.08	22.05	<.001	.08
Label Condition (L)[1]	9.58	<.001	.07	1.47	.232	.01
Age (A)[1]	2.86	.092	.01	.01	.919	<.001
Gender (G)[1]	7.21	.008	.03	1.09	.297	<.01
Mindset (M)	1.23	.269	.01	1.26	.262	.01

[26] Important notes with respect to the results reported in the table: (a) although SPSS automatically computed interaction effects between the within-subjects factor and each between-subject factors (including each between-subjects interaction terms), the results reported for *Scenario* here are limited to *2* and *3*-way interactions from Model 1; (b) all four-way interaction terms from both models were omitted; (c) all other interaction terms included in Models 1 and 2 are reported; and (d) only one interaction involving R was included in these analyses.

	Valence Belief Ratings			*Intentionality Scores*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Stereotype Vulnerability (SV)	.65	.422	<.01	.92	.339	<.01
Racial Identity (ID)	3.17	.076	.01	1.52	.218	.01
SC × S[1]	.23	.632	<.01	<.01	.974	<.001
SC × L[1]	7.79	.001	.06	2.75	.066	.02
SC × A[1]	.59	.444	<.01	<.001	.991	<.001
SC × G[1]	.25	.616	<.01	.01	.916	<.001
S × L[1]	.64	.529	.01	2.73	.067	.02
S × A[1]	1.03	.311	<.01	.34	.563	<.01
S × G[1]	2.13	.146	.01	2.31	.130	.01
S × R[1]	.18	.673	<.01	.05	.823	<.001
L × A[1]	3.71	.026	.03	1.45	.236	.01
L × G[1]	1.16	.317	.01	.53	.592	<.01
S × M	5.65	.018	.02	1.31	.254	.01
S × SV	28.98	<.001	.11	1.67	.198	.01
S × ID	.01	.936	<.001	.94	.332	<.01
L × M	3.58	.029	.03	1.63	.197	.01
L × SV	1.81	.167	.02	.35	.705	<.01
L × ID	2.29	.103	.02	1.16	.314	.01
S × L × A[1]	2.60	.076	.02	.11	.899	<.01
S × L × G[1]	.86	.426	.01	.41	.666	<.01
SC × S × L[1]	3.92	.021	.03	.71	.493	<.01
SC × S × A[1]	.14	.709	<.01	<.01	.970	<.01

	Valence Belief Ratings			*Intentionality Scores*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
SC × S × G[1]	1.67	.198	.01	.01	.941	<.001
S × L × M	1.43	.242	.01	1.61	.202	.01
S × L × SV	3.49	.032	.03	1.27	.283	.01
S × L × ID	1.16	.314	.01	.90	.409	.01

[1] Denotes the results reported from *Model 1* (*dfs = 254*). All other reported results were reported from *Model 2* (*dfs* = 238).

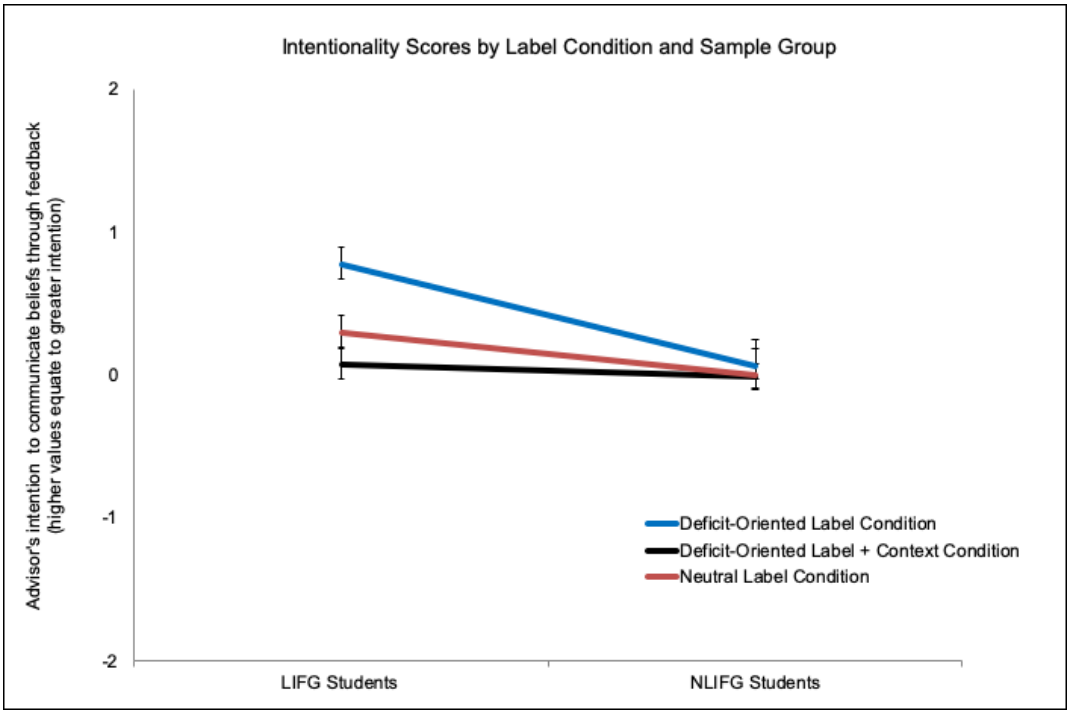
Intention of communicating beliefs.

Model 1. The results of the analysis with participants' *S1_intentionality_score* and *S2_intentionality_score* as dependent measures yielded a significant main effect of *recruitment method*, such that participants' mean intentionality score was significantly lower for participants recruited through BC (*M*= –.20, *SE*= .07), relative to those recruited from MTurk (*M*= .32, *SE*= .07). Although, there were no significant main effects of *sample group* or *label condition*, there was a marginally significant interaction between *sample group* × *label condition* (see *Figure 5.6*). The pairwise comparisons by *label condition* indicated that for LIFG students, the mean intentionality score was significantly higher for those assigned to the *neutral label condition*, relative to those assigned to the *deficit-oriented label condition* (*M*= .30, *SE*= .10; *M*= –.78, *SE*= .12, respectively; *p* =.038). However, there were no significant differences in mean intentionality scores between those in the *neutral label* and *deficit-oriented label + context conditions* (*M*= .08, *SE*= .11, *p* = .329) or the *deficit-oriented label* and *deficit-oriented label + context conditions* (*p*= .674). In contrast, none of the comparisons were significant for NLIFG students across the *deficit-oriented label, neutral label,* and *deficit-oriented label + context conditions* (*M*= .06, *SE*= .19; *M*= –.003, *SE*= .09; *M*= –.01, *SE*= .09, respectively; *ps* > .890). The

comparisons by *sample group* indicated that LIFG and NLIFG students' mean intentionality scores significantly differed in the *neutral label condition* (p = .025) but not in the *deficit-oriented label* or *deficit-oriented label + context conditions* (*ps* > .531).

Model 2. The results of the analysis that included participants' motivational beliefs did not yield any statistically significant main effects of *academic mindset*, *stereotype vulnerability*, or *racial identity beliefs*. Moreover, none of the *two* and *three*-way interactions with students' motivational beliefs were statistically significant (see *Table 5.3* above for full report of ANCOVA results).

Figure 5.6. Intentionality scores by sample group and label condition in Study 3.



Students' perceptions of the effects of being labeled. A total of five ANCOVAs were conducted using participants' subscale scores for *SOB*, *ASP*, and *AFF*, as well as their responses to two items related to academic engagement, *AE Item 8* (reverse-coded; "Being labeled as at-

risk, and/or underprepared, and/or disadvantaged probably made Aaron/Ryan [April/Casey] feel motivated to work harder in his [her] classes.") and *AE item 9* ("Being labeled as at-risk, and/or underprepared, and/or disadvantaged probably made Aaron/Ryan [April/Casey] feel hesitant to take any challenging courses moving forward."). The results reported in the next two subsections are limited to those that were consistent across analyses and/or those that were most relevant to the hypotheses being tested. In cases where a main effect or interaction was statistically significant for more than one subscale score, means and follow-up tests are provided for the dependent measures that were of primary interest to this study, which were academic self-perceptions (*ASP*) and sense of belonging (*SOB*), or for the dependent measure that yielded the strongest effect. A full report of the main effects and relevant interactions from these analyses can be viewed in *Tables 5.4* and *5.5*.

Model 1. The results of the analyses using participants' *ASP* and *AFF* subscale scores yielded a marginally significant main effect of *recruitment method.* Specifically, participants recruited through MTurk generally perceived the advisor's feedback as negatively influencing the student's academic self-perceptions to a significantly lesser extent than did participants recruited through BC. Several of the analyses also yielded significant main effects of *scenario* and *label condition*. More specifically, participants' *ASP* scores were significantly more negative in *Scenario 1* (M= 4.47, SE= .07) compared to *Scenario 2* (M= 4.21, SE= .08), and significantly less negative in the *neutral label condition* (M= 3.97, SD= .10) than in the *deficit-oriented label* (M= 4.63, SE= .11, $p < .001$) and *deficit-oriented label + context conditions* (M= 4.46, SE= .10, $p = .001$). However, the difference in mean scores between those in the *deficit-oriented label* and *deficit-oriented label + context conditions* was not significant ($p = .490$; see *Tables 5.4* and *5.5* for detailed results). The analysis for *AE Item 8* also yielded a significant main effect of

sample, relative to NLIFG students (*M*= 3.11, *SE*= .09), LIFG students generally disagreed less that the advisor's feedback would make the student feel 'motivated to work harder in his [her] classes' (*M*= 2.77, *SE*= .11).

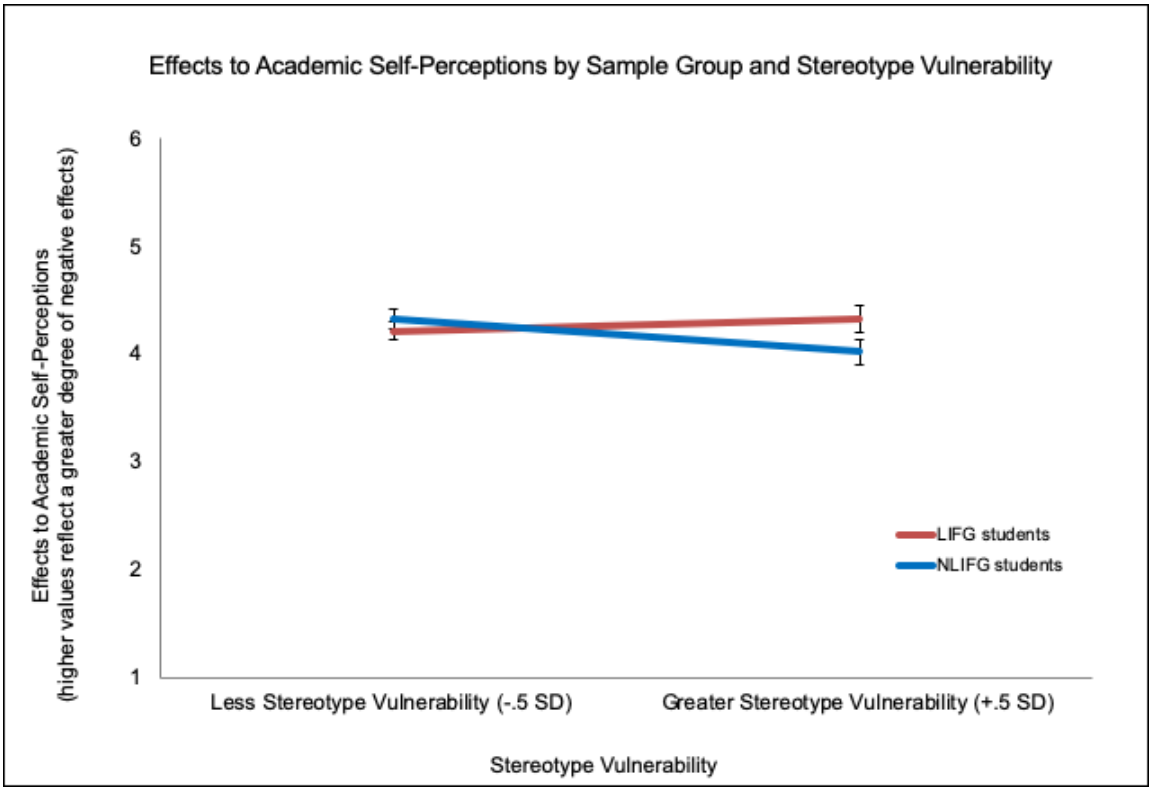
Additionally, the *scenario* × *label condition* was statistically significant or marginally significant for all subscale scores except *AE Item 9*. Pairwise comparisons for participants' *ASP* scores by *label condition*, showed that on average, *ASP* scores for *Scenario 1* were significantly less negative for participants in the *neutral label* condition (i.e., *first-generation student* label; *M*= 4.25, *SE*= .11) than those in the *deficit-oriented label condition* (i.e., *at-risk* label; *M*= 4.66, *SE*= .12, *p*= .023), but not compared to those in the *deficit-oriented label + context condition* (i.e., *at-risk* label; *M*= 4.48, *SE*= .11, *p*= .327). The remaining comparison for *Scenario 1* was nonsignificant ($p = .572$). For *Scenario 2*, participants in the *neutral label* condition (i.e., *first-generation student* label; *M*= 3.68, *SE*= .11) had *ASP* scores significantly less negative than participants in the *deficit-oriented label* condition (i.e., *underprepared* label; *M*= 4.59, *SE*= .13), and *deficit-oriented label + context conditions* (i.e., *underprepared* label; *M*= 4.43, *SE*= .12, *ps* < .001). However, the difference between *ASP* scores of those in the *deficit-oriented label* and *deficit-oriented label + context conditions* was not statistically significant (*p*= .672). The comparisons by *scenario* indicated that the difference in *ASP* scores between *Scenarios 1* and *2* was significant in the *neutral label* condition ($p < .001$), but not in the *deficit-oriented label* (*p*=.577) or the *deficit-oriented label + context condition* (*p*=.639). Lastly, none of the *Model 1* analyses yielded significant *scenario* × *sample group*, $F(1, 254)s < 1.12$, $ps > .148$, $\eta_p^2 s < .01$, or *sample group* × *label condition* interactions, $F(2, 254)s \leq 9.41$, $ps \geq .895$, $\eta_p^2 s \leq .01$ (see *Tables 5.4* and *5.5* for detailed results).

Model 2. The results of these analyses yielded a marginally significant main effect of *mindset scores* for the analyses including *AE Item 8* and *AE Item 9*. Marginal means were estimated at 1.5 *SD* above and below the mean mindset score for *AE Item 8*, which yielded the strongest effect. On average, participants with stronger growth mindsets believed that the hypothetical student in the scenarios would be 'motivated to work hard in their classes' to a lesser extent (*M*= 2.39, *SE*= .16) than those with weaker growth mindsets (*M*= 2.77, *SE*= .13). The analysis with *AE Item 8* also yielded a significant main effect of participants' *stereotype vulnerability scores*, $F(1, 238) = 5.53$, $p = .019$, $\eta_p^2 = .01$. Marginal means were estimated for *AE Item 8* at .5 *SD* above and below the mean *stereotype vulnerability score.* On average, participants with greater stereotype vulnerability had significantly higher ratings for this item (*M*= 2.69, *SE*= .10) than those with a lower stereotype vulnerability (*M*= 2.46, *SE*= .14; see *Tables 5.4* and *5.5* for detailed results).

These analyses also yielded several statistically significant interactions with participants' motivational beliefs. First, the analyses of *ASP* and *AE Item 8* yielded statistically significant *sample* × *stereotype vulnerability* interactions. Pairwise comparisons were conducted based on marginal means that were estimated for *ASP scores* at .5 *SD* above and below the mean *stereotype vulnerability score*. On average, participants' *ASP* scores did not significantly differ between LIFG and NLIFG students with stereotype vulnerability scores that were .5 *SD* below the mean score (*M*= 4.22, *SE*= .23; *M*= 4.33, *SE*= .09, respectively; *p*= .664). For participants with scores .5 *SD* above the mean, LIFG students' ratings were marginally higher than NLIFG students (*M*= 4.33, *SE*= .14; *M*= 4.02, *SE*= .12, respectively; *p*= .094; see *Figure 5.7*). Within-group differences were probed through correlational analyses examining the associations between participants' *stereotype vulnerability scores* and their *ASP* scores for each scenario, by

sample group. These analyses showed that for LIFG students, higher stereotype vulnerability scores were associated with higher *ASP* scores for *Scenario 1*, $r(106) = .28$, $p = .004$, but not for *Scenario 2*, $r(106) = -.03$, $p = .794$. In contrast, the analysis for NLIFG students showed that the associations between stereotype vulnerability and *ASP* scores for *Scenario 1* and *2* were both statistically significant and in the reverse direction, $r(164) = -.37$, $p < .001$ and $r(164) = -.23$, $p = .003$, respectively.

Figure 5.7. Significant 2-way interaction between sample group and stereotype vulnerability on participants' perceptions of effects on academic self-perceptions.



Second, the models including *SOB* scores and *AE Item 8* yielded a statistically significant interaction between *label condition* × *mindset scores*. Pairwise comparisons were conducted based on marginal means that were estimated using *SOB* scores at 1.5 *SD* above and below the

mean mindset score, which indicated that for participants with stronger mindsets (i.e., scores 1.5 *SD* above the mean), those in the *deficit-oriented label condition* believed that the hypothetical student in the scenarios would experience a greater degree of negative effects on their sense of belonging (SOB; M= 3.92, SE= .25) than those in the *deficit-oriented label + context condition* (M= 2.90, SE= .30; p= .026) and marginally higher than those in the *neutral label condition* (M= 3.24, SE= .20; p= .092). The comparison between the *deficit-oriented label + context* and *neutral label condition* was not significant (p= .724). For participants with weaker growth mindsets, those in the *deficit-oriented label + context condition* had *SOB* scores that were marginally higher (M= 3.80, SE= .23) than those in the *neutral label condition* (M= 3.18, SE= .19; p= .078). However, the comparison between those in the *neutral label condition* and those in the *deficit-oriented label condition* (M= 3.40, SE= .19), as well as the comparison between those in the *deficit-oriented label* condition and *deficit-oriented label + context condition* were not significant ($ps > .406$; see *Figure 5.8*).[27]

Lastly, the model including the *AE Item 8* yielded a statistically significant interaction between *sample* × *label condition* × *stereotype vulnerability scores*. Pairwise comparisons were conducted based on marginal means that were estimated at .5 *SD* above or below the mean stereotype vulnerability score. The comparisons by *sample group* indicated that for participants with greater stereotype vulnerability (i.e., scores .5 *SD* above or below the mean score), LIFG students had marginally lower ratings than NLIFG students for *AE Item 8* when they were in the *deficit-oriented label + context condition* (M= 1.11, SE= .52; M= 3.24, SE= .15; $p \leq .001$) and *deficit-oriented label condition* (M= 1.86, SE= .34; M= 3.38, SE= .16; $p \leq .001$). However, the comparison between LIFG and NLFG students in the *neutral label condition* was nonsignificant

[27] Given both *SOB* scores and *AE Item 8* yielded interactions of equal significance, the marginal means were estimated using *SOB* scores because this dependent measure was of particular interest to the current research.

(*M*= 2.33, *SE*= .33; *M*= 2.85, *SE*= .14; *p*= .149). For those with lower stereotype vulnerability, LIFG students had significantly lower ratings than NLIFG students for *AE Item 8* when they were in the *deficit-oriented label + context condition* (*M*= 2.35, *SE*= .28*; M*= 2.99, *SE*= .23; *p* < .001), but not the *neutral label condition* (*M*= 2.57, *SE*= .17; *M*= 2.81, *SE*= .21; *p*= .367) or the *deficit-oriented label condition* (*M*= 2.55, *SE*= .24; *M*= 2.88, *SE*= .44; *p* = .331).

Figure 5.8. Significant 2-way interaction between label condition and academic mindsets on participants' perceptions of effects on sense of belonging.

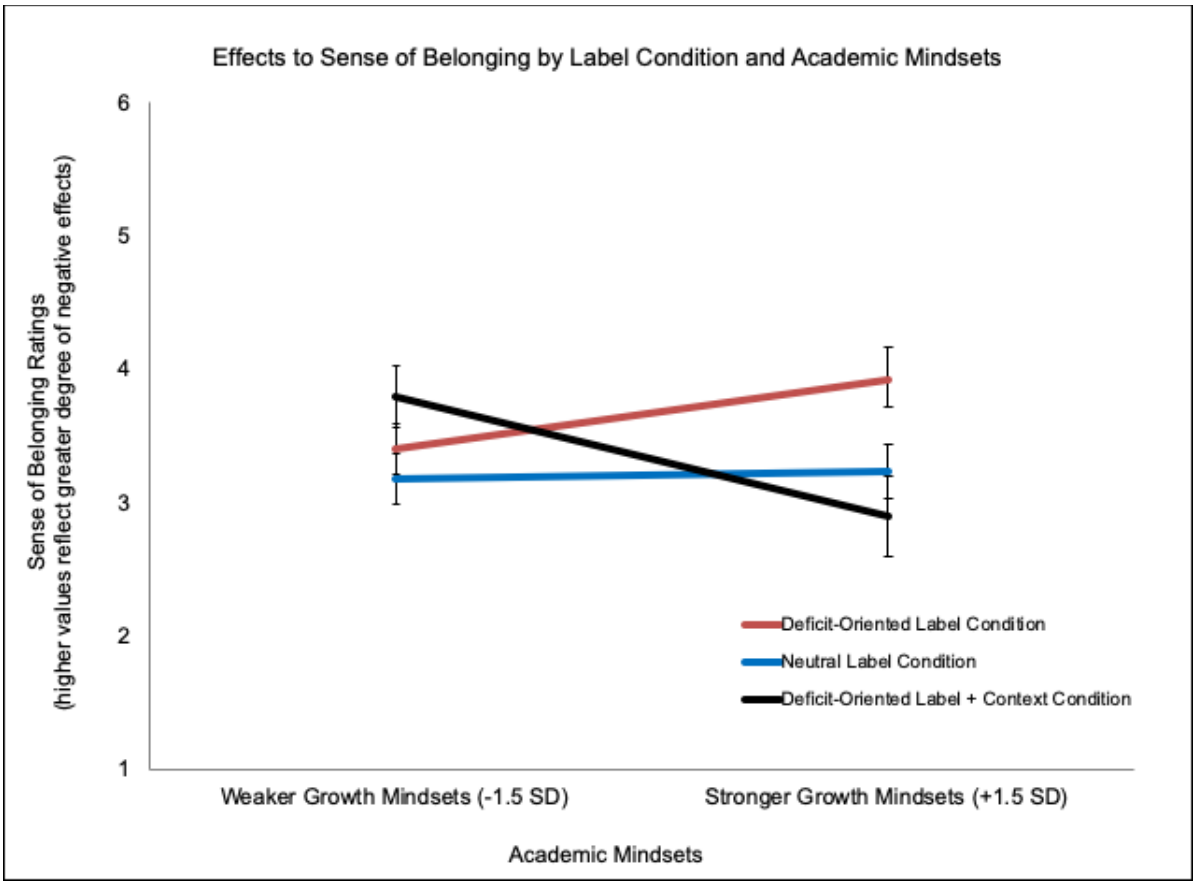


Table 5.4. Summary of ANCOVA results for participants' perceptions of effects of being labeled by a deficit-oriented or neutral label on academic self-perceptions (ASP), sense of belonging in college (SOB), and affect (AFF).[28]

	ASP			*SOB*			*AFF*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Scenario (SC)[1]	8.96	.003	.03	.16	.693	<.01	1.50	.222	.01
Sample (S) [1]	.51	.476	<.01	<.01	.937	<.001	.48	.491	<.01
Recruitment (R) [1]	3.39	.067	.01	.11	.739	<.001	3.08	.081	<.01
Label Condition (L) [1]	13.76	<.001	.10	10.42	<.001	.08	12.23	<.001	.09
Age (A) [1]	.06	.812	<.001	.12	.727	<.001	.06	.812	<.001
Gender (G) [1]	.82	.366	<.01	.48	.488	<.01	.68	.412	<.01
Mindset (M)	.99	.321	<.01	.33	.567	.001	.52	.470	<.01
Stereotype Vul. (SV)	1.11	.293	.01	.46	.496	<.01	1.69	.195	.01
Racial Id. (ID)	.24	.626	<.01	1.19	.276	.01	.58	.445	<.01
SC × S[1]	1.60	.206	<.01	2.10	.149	<.01	1.58	.210	.01
SC × L[1]	6.59	.002	.05	9.41	<.001	.07	6.81	.001	.05
SC × A[1]	.11	.740	<.001	.01	.912	<.001	.024	.878	<.001
SC × G[1]	1.99	.160	.01	1.44	.231	.01	2.53	.113	.01
S × L[1]	.85	.427	<.01	.70	.497	<.01	.66	.520	.01
S × A[1]	4.82	.029	.02	.77	.38	<.01	1.91	.168	.01
S × G[1]	7.56	.006	.03	12.05	.001	.05	2.97	.086	.01
S × R[1]	1.23	.268	<.01	.11	.741	<.001	.64	.425	<.01

[28] Important notes with respect to the results reported in the table: (a) although SPSS automatically computed interaction effects between the within-subjects factor and each between-subject factors (including each between-subjects interaction terms), the results reported for *Scenario* here are limited to *2* and *3*-way interactions from Model 1; (b) all four-way interaction terms from both models were omitted; (c) all other interaction terms included in Models 1 and 2 are reported; and (d) only one interaction involving R was included in these analyses.

	ASP			*SOB*			*AFF*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p	*F*	*p*	η^2_p
L × A[1]	1.60	.205	.01	.38	.686	<.01	1.12	.327	.01
L × G[1]	1.77	.172	.01	.28	.755	<.01	1.50	.225	.01
S × M	.02	.903	<.001	2.24	.136	.01	.05	.827	<.001
S × SV	5.22	.023	.02	1.97	.162	.01	1.02	.315	<.01
S × ID	.11	.740	<.001	.41	.522	<.01	.01	.931	<.001
L × M	.36	.699	<.01	4.38	.014	.04	.30	.739	<.01
L × SV	.45	.639	<.01	.124	.884	<.01	.71	.493	.01
L × ID	2.37	.096	.02	.74	.476	.01	2.05	.131	.02
S × L × A[1]	.08	.923	<.01	.92	.400	.01	4.05	.019	.03
S × L × G[1]	.41	.667	<.01	.54	.584	<.01	.26	.775	<.01
SC × S × L[1]	1.94	.145	.02	.74	.479	<.01	.75	.473	.01
SC × S × A[1]	.14	.709	<.01	.61	.437	<.01	.96	.385	.10
SC × S × G[1]	8.97	.003	.03	1.27	.261	.01	4.26	.040	.02
S × L × M	2.36	.096	.02	.89	.414	.01	1.65	.195	.01
S × L × SV	2.15	.119	.02	2.80	.063	.02	2.28	.104	.02
S × L × ID	.96	.386	.01	5.00	.007	.01	1.53	.220	.01

[1] Denotes the results yielded by *Model 1* (*dfs = 254)*. All other reported results were yielded by *Model 2* (*dfs* = 238).

Table 5.5. Summary of ANCOVA results for participants' perceptions of effects of being labeled by a deficit-oriented or neutral label on two items measuring aspects of academic engagement (AE).[29]

	AE Item 8			*AE Item 9*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Scenario (SC)[1]	.01	.942	<.001	5.41	.021	.02
Sample (S) [1]	6.00	.015	.02	.13	.715	<.01
Recruitment (R) [1]	2.50	.115	.01	1.28	.259	.01
Label Condition (L) [1]	3.57	.029	.03	7.36	.001	.06
Age (A) [1]	3.77	.053	.02	.15	.700	<.01
Gender (G) [1]	2.75	.098	.01	4.00	.047	.02
Mindset (M)	3.63	.058	.02	3.11	.079	.01
Stereotype Vul. (SV)	5.53	.019	.02	1.34	.248	.01
Racial Id. (ID)	2.69	.102	.01	.53	.468	<.01
SC × S[1]	1.11	.293	<.01	1.22	.271	.01
SC × L[1]	2.94	.054	.02	1.92	.149	.015
SC × A[1]	.70	.405	<.01	.83	.364	<.01
SC × G[1]	.06	.811	<.001	.55	.458	<.01
S × L[1]	1.77	.173	.01	.11	.895	<.01
S × A[1]	.08	.778	<.001	.58	.446	<.01
S × G[1]	.24	.623	.001	.19	.731	<.001
S × R[1]	.31	.577	.001	1.08	.299	<.01
L × A[1]	1.84	.16	.01	2.25	.108	.02
L × G[1]	.11	.895	.001	.32	.728	<.01

[29] Important notes with respect to the results reported in the table: (a) although SPSS automatically computed interaction effects between the within-subjects factor and each between-subject factors (including each between-subjects interaction terms), the results reported for *Scenario* here are limited to *2* and *3*-way interactions from Model 1; (b) all four-way interaction terms from both models were omitted; (c) all other interaction terms included in Models 1 and 2 are reported; and (d) only one interaction involving R was included in these analyses.

	AE Item 8			*AE Item 9*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
S × M	<.01	.954	<.001	1.25	.264	.01
S × SV	24.78	<.001	.09	2.17	.142	.01
S × ID	2.72	.101	.01	.49	.487	<.01
L × M	4.48	.012	.04	.21	.808	<.01
L × SV	3.79	.024	.03	.74	.476	.01
L × ID	.89	.414	.01	.89	.411	.01
S × L × A[1]	1.26	.286	.01	.88	.416	.01
S × L × G[1]	.20	.980	<.001	1.02	.363	.01
SC × S × L[1]	4.19	.016	.03	1.00	.369	.01
SC × S × A[1]	1.55	.214	.01	.04	.837	<.001
SC × S × G[1]	.60	.438	<.01	.06	.814	<.001
S × L × M	4.78	.009	.04	.21	.808	<.01
S × L × SV	1.59	.206	.01	.95	.387	.01
S × L × ID	1.12	.328	.01	2.27	.105	.02

[1] Denotes the results yielded by *Model 1* (*dfs = 254*). All other reported results were yielded by *Model 2* (*dfs* = 238).

Examining Potential Stereotype Threat Effects

***Deficit-oriented* versus *neutral labeling* experience manipulation check.** The extent to which the manipulation of labeling experiences was effective was assessed by first creating four new dichotomous variables (*positive; negative; neutral / both positive and negative; irrelevant*) and coding students' descriptions of the way they felt (or imagined they would have felt) after being labeled as an 'at-risk student' (*deficit-oriented label condition*; DOL for short) or as a 'first-year student' (*neutral label condition*; NL for short). Responses were coded using the same

codes developed for the analysis of students' open-ended responses in Study 2, such that responses that indicated a negative effect was experienced as a result of the labeling experience (e.g., "*I felt demoralized and degraded*") were coded as a 1 for the *negative* variable (*DOL*= 84; *NL*= 45), responses that indicated a positive effect was experienced as a result of the labeling experience (e.g., "*I was motivated to work harder*") were coded as a 1 for the *positive* variable (*DOL*= 27; *NL*= 34), responses that indicated the effect of the labeling experience was neither positive or negative or both (e.g., "*I didn't care*") were coded as a 1 for the *neutral* variable (*DOL*= 11; *NL*= 48), and lastly, responses that were irrelevant to the prompt or uninterpretable (e.g., "*a*"; "*NA*", etc.) were coded as 1 for the *not applicable/irrelevant* (*DOL*= 16; *NL*= 9). Each category was mutually exclusive, and responses were only coded as relevant to one of the four categories. The twenty-five responses that were coded as *not applicable/irrelevant* across both conditions were removed from all subsequent analyses that included the *manipulation condition* as a variable, because I had no way of assessing whether these participants had recalled a labeling experience that matched the condition they were assigned to, and therefore would not be able to make claims about the influence of that manipulation on their verbal scores. Therefore, the sample for these analyses was *N*= 249 (LIFG *n*= 92; NLIFG *n*= 157).

Following this, a series of Chi-Square tests to was conducted using the recoded quantitative *effect* variables to examine differences in the proportion of participants who reported positive, negative, and neutral effects, by condition. The results of these analyses indicated that there was a statistically significant between-group difference in the proportion of participants who reported neutral, $X^2(1, N= 249) = 28.50$, $p < .001$ and negative effects, $X^2(1, N= 249) = 27.84$, $p < .001$, but not positive effects, $X^2(1, N= 249) = .724$, $p= .395$. More specifically, participants assigned to describe an experience in which they were labeled as an *at-risk student*

reported feeling negative effects more often—but neutral effects less often—than participants assigned to describe an experience in which they were labeled as a *first-year student*. In contrast, there was no difference in the proportion of participants who reported positive effects between conditions. These results suggested that the manipulation of labeling experience was generally effective.

Primary analyses. The results reported below were yielded through an analysis designed to investigate my hypotheses regarding the effects of recalling an experience of being labeled as an *at-risk student* versus a *first-year student*, on participants' subsequent performance on a verbal test. Moreover, the analysis tested several hypotheses regarding the potential moderating effects of participants' personal goals for their academic performance in college, their stereotype vulnerability, their racial identity beliefs, and their academic mindsets. Given that participants' *verbal scores*—the dependent measure of interest for these hypotheses—was a one-time measure, and that several hypotheses were associated with continuous covariate measures, I conducted two ANCOVAs. The first was a preliminary model that included participants' *verbal scores* as the dependent measure, with *sample group* (*LIFG*; *NLIFG*), *manipulation condition* (*DOL; NL), gender,* and *recruitment method* as between-subject factors, and participants' *age* and the *personal goals* item as continuous between-subjects covariates. This model also tested a number of interactions, including (but not limited to): *sample group* × *recruitment method; sample group* × *condition*; *sample group* × *age; sample group* × *gender; condition* × *gender; condition* × *age; sample group* × *condition* × *personal goals.* This analysis did not yield a significant main effect of *gender*, *recruitment method*, or significant interactions between these factors and *sample group*, $F(1, 248)s \leq 1.58$, $ps \geq .210$, $\eta_p^2 s \leq .001$. Given that *gender* and

recruitment method were not relevant to any of the hypotheses being tested in this study, they were removed from the subsequent analysis.

The final analysis was computed in two steps. *Model 1* (i.e., the first step) included participants' *verbal scores* as the dependent measure, with *sample group* (*LIFG*; *NLIFG*), *manipulation condition* (*DOL; NL)* as between-subject factors, as well as participants' *age* and the *personal goals* item as continuous between-subjects covariates. This model tested the same interactions as the previous model, with the exception of those involving *gender* and *recruitment method* (which were not included in this analysis). *Model 2* (i.e., the second step) included the same factor and covariates as *Model 1*, but included standardized variables for participants' *mindset scores*, *stereotype vulnerability scores*, and *racial identity beliefs scores* as additional covariates. These factors were also included in two and three-way interactions with *sample group* and *condition* (see Table 5.6 for full list of interactions included in the analyses).

Model 1. The results of the first model did not yield a significant main effect of *condition, age,* or *personal goals* on participants' verbal scores (see *Table 5.6* for a full report of the main effects and interactions), but did yield a significant main effect of *sample group*, such that on average, LIFG students scored significantly lower on the verbal test (M= 3.73, SE= .28) than NLIFG students (M= 5.42, SE= .21). There was also a marginally significant *sample group* × *age* interaction, $F(1, 237) = 3.17$, $p = .077$, $\eta_p^2 = .01$. Pairwise comparisons were conducted based on marginal means estimated at 1 *SD* above and below the mean age of the sample, which indicated that for participants 1 *SD* younger than the mean age of the sample, LIFG students had significantly lower scores (M= 2.85, SE= .65) than NLIFG students (M= 5.02, SE= .44, p= .006), whereas there was no significant difference in scores between LIFG and NLIFG students who were 1 *SD* older the mean age of the sample (M= 3.60, SE= .62; M= 4.29, SE= .42, p=

.353). Within-group differences were probed through correlational analyses examining the associations between participants' *age* and their *verbal test scores*, by sample group. These analyses showed that for LIFG students, age was positively—but not significantly—associated with their *verbal test scores*, $r(92) = .11$, $p = .291$—whereas for NLIFG students, this association was negative and marginally significant, $r(157) = -.14$, $p = .072$.

Model 2. The results of the second analysis that included participants' motivational beliefs yielded a statistically significant main effect of participants' *racial identity beliefs scores*, but not for *mindset scores* or *stereotype vulnerability scores*. Marginal means estimated at 1.5 *SD* above and below the mean *racial identity beliefs score*, indicated that on average, participants who identified more strongly with their race and/or ethnicity (i.e., racial identity beliefs scores 1.5 *SD* above the mean score) had significantly lower verbal scores ($M = 3.26$, $SE = .44$) compared to participants who had weaker racial identity beliefs ($M = 4.64$, $SE = .42$).

This analysis also yielded significant *2*-way interactions between *sample group* × *mindset scores* and *sample group* × *stereotype vulnerability scores*. Pairwise comparisons were conducted for both interactions based on marginal means estimated at 1.5 *SD* above and below the mean mindset score and .5 *SD* above and below the mean stereotype vulnerability score. These analyses indicated that for participants with stronger growth mindsets (i.e., mindset scores 1.5 *SD* above the mean score), the difference in verbal scores between LIFG and NLIFG students was nonsignificant ($M = 3.70$, $SE = .77$; $M = 4.08$, $SE = .42$, respectively, $p = .664$), whereas this difference was significant for LIFG and NLIFG students with weaker growth mindsets ($M = 2.25$, $SE = .66$; $M = 5.49$, $SE = .38$, respectively, $p < .001$).

The comparisons for stereotype vulnerability scores indicated that for participants with greater stereotype vulnerability, the difference in verbal scores between LIFG and NLIFG

students was non-significant (M= 3.59, SE= .38; M= 4.43, SE= .37, respectively, p= .115), whereas this difference was significant for LIFG and NLIFG students with less stereotype vulnerability (M= 2.70, SE= .69; M= 5.01, SE= .23, respectively, $p < .001$). Two additional ANCOVA models were computed to probe within-group differences. These models were identical to the original *Model* 2 analysis, with the exception that each included *sample group* as a dummy-coded covariate rather than as a between-subjects factor. In one analysis, sample group was coded to specify LIFG as the reference group and in the other analysis it was coded to specify NLIFG as the reference group. The results of these analyses indicated that increases in stereotype vulnerability scores were significantly associated with increases in verbal test scores for LIFG students, $F(1, 225) = 3.09$, $p = .080$, $\eta_p^2 = .01$, and decreases in verbal scores for NLIFG students, $F(1, 225) = 4.77$, $p = .030$, $\eta_p^2 = .02$.

Table 5.6. Summary of ANCOVA results for analysis examining potential stereotype threat effects.

	Model 1			*Model 2*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
Sample (S)	22.15	<.001	.74	–	–	–
Condition (C)	1.89	.170	.01	–	–	–
Age (A)	.28	.600	<.01	–	–	–
Mindset (M)	–	–	–	.50	.830	<.001
Stereotype Vulnerability (SV)	–	–	–	.18	.675	<.01
Racial Identity (ID)	–	–	–	4.75	.030	.02
Personal Goals (P)	1.25	.264	.01	–	–	–

	Model 1			*Model 2*		
Variable	*F*	*p*	η^2_p	*F*	*p*	η^2_p
S × C	2.12	.146	.01	–	–	–
S × A	3.17	.077	.01	–	–	–
C × A	.71	.399	<.01	–	–	–
S × P	.13	.715	<.01	–	–	–
C × P	1.25	.264	.01	–	–	–
S × M	–	–	–	4.71	.031	.02
S × SV	–	–	–	7.58	.006	.03
S × ID	–	–	–	.06	.811	<.001
C × M	–	–	–	.17	.679	<.01
C × SV	–	–	–	.75	.387	<.01
C × ID	–	–	–	.04	.853	<.001
S × C × A	1.19	.276	.01	–	–	–
S × C × M	–	–	–	1.80	.181	.01
S × C × SV	–	–	–	.83	.364	<.01
S × C × ID	–	–	–	.001	.976	<.001
S × C × P	.31	.816	<.01	–	–	–

Note: Sample size is 249; *Model 1 dfs = 237*; *Model 2 dfs* = 225.

Discussion

Study 3 yielded several particularly interesting results pertaining to the hypotheses explored in this dissertation. First, the results of the analyses examining participants' responses to the hypothetical scenarios indicated that across all label conditions, LIFG students generally interpreted the advisor's feedback in the scenarios as communicating more positive beliefs about

the hypothetical students, whereas NLIFG students interpreted the advisor's feedback as communicating generally more negative beliefs. Across groups, students in the *neutral label condition* interpreted the advisor's feedback more positively than those in the *deficit-oriented label* and *deficit-oriented label + context conditions*. The analyses with participants' motivational beliefs indicated that for participants with a lesser degree of stereotype vulnerability, on average, LIFG students seemed to interpret the advisor's feedback more positively than NLIFG students, whereas this was not the case for LIFG students with greater stereotype vulnerability. The results also indicated that for participants with stronger growth mindsets, the group difference in the perceived valence of the advisor's feedback was reduced to nonsignificance, whereas this difference was significant for participants with weaker growth mindsets. With respect to the perceived intentionality of the advisor's feedback, LIFG students assigned to the *neutral label condition* interpreted the advisor as being significantly more intentional in communicating their more positive beliefs about the hypothetical student than those assigned to the *deficit-oriented label* condition; but this was not the case for NLIFG students. This finding was consistent with the findings from Study 1.

The analyses examining participants' perceptions of how the advisor's feedback would affect motivation and affective state of the hypothetical student indicated that the only significant difference by sample group emerged in relation to effects on one of the academic engagement items. More specifically, relative to NLIFG students, LIFG students generally agreed more that advisor's feedback would make the student feel 'motivated to work harder in his [her] classes'. Moreover, participants assigned to the *neutral label* condition believed that the hypothetical students would experience a lesser degree of negative effects on their academic self-perceptions, sense of belonging in college, affect, and academic engagement as a result of their interaction

with the advisor, relative to those assigned to the *deficit-oriented label* condition and/or the *deficit-oriented label + content* condition. The analysis with participants' motivational beliefs indicated that, for participants with greater stereotype vulnerability, LIFG students generally expected the hypothetical student in the scenario to experience a greater degree of negative consequences to their academic self-perceptions and academic engagement than did the NLIFG students; but, this was not the case for participants with lower stereotype vulnerability. Moreover, across all label conditions and scenarios, participants with stronger growth mindsets expected that the hypothetical student would experience a marginally lower degree of negative effects on their academic engagement, relative to those with weaker growth mindsets. Oddly, the results also showed that, for participants with stronger growth mindsets, those assigned to the *deficit-oriented label* condition expected the hypothetical student to experience a significantly greater degree of negative effects to sense of belonging than those assigned to the *neutral label* or *deficit-oriented label + context* conditions, whereas these differences between label conditions were not significant for participants with weaker growth mindsets.

Lastly, the results of the analyses exploring potential stereotype threat effects of students' deficit-oriented labeling experiences, failed to yield a significant effect of manipulation on participants' verbal test scores, but indicated that, in general, LIFG students' verbal test scores were significantly lower than NLIFG students' scores. However, the analysis with participants' motivational beliefs also suggested that endorsing certain beliefs could potentially ameliorate these disparities, such that the mean difference between LIFG and NLIFG students' verbal test scores was reduced to nonsignificance for students with stronger growth mindsets, but not for those with weaker growth mindsets. The results of this analysis also indicated that for both LIFG

and NLIFG students, stronger racial identity beliefs were associated with lower verbal test scores.

Overall, the findings from this study both replicated and expanded on the findings from *Study 1*. With respect to the hypothetical scenarios, both LIFG and NLIFG students interpreted the 'first-generation' label as communicating more positive (for LIFG students) or less negative beliefs (for NLIFG students) than the 'at-risk' or 'underprepared' labels (when used ambiguously or with added context). Interestingly, LIFG students generally seemed to interpret the advisor's feedback to the hypothetical student as communicating somewhat positive beliefs, whereas NLIFG students seemed to show the reverse trend. This finding was somewhat surprising, given that in Study 1, the corresponding means were not significantly different between groups and were above or very close to the midpoint of the scale. One possible explanation for these findings may be that because LIFG students have likely experienced being characterized by these types of labels in the past, they may also become somewhat desensitized or even primed to expect these experiences to occur. As a result, their immediate reactions may seem subdued, compared to that of students who rarely experience being characterized by these labels—such as NLIFG students. Despite LIFG students' more positive interpretations, in general, all participants seemed to believe that characterizing students as *at-risk* or *underprepared* would be more motivationally damaging than labeling them as a *first-generation student*.

Additionally, LIFG students perceived the advisor as being more intentional in communicating his or her positive beliefs about the hypothetical student when he or she characterized that student as a 'first-generation student', relative to when he or she ambiguously characterized the student in terms of a deficit-oriented descriptor. In contrast, NLIFG students did not seem to make this distinction in perceived intentionality to the same extent. These

findings suggest that—in contrast to NLIFG students—LIFG students may have picked up on the advisor's use of additional context as a way of as a way of reducing the stigma associated with the label. In regards to the neutral *first-generation* label, LIFG students may have interpreted the advisor's choice of descriptor as a nod to the student's background and as their way of communicating their recognition for what the student has already accomplished. With respect to NLIFG students, again, these findings suggest that NLIFG students' lack of exposure to these labels may lead them to interpret the labels as being intentionally harsh and negative.

The findings from this study suggest there are individual differences in the manner in which students interpret the *first-generation* label and *deficit-oriented* descriptors. For instance, for LIFG students, greater stereotype vulnerability was associated with more negative interpretations of the advisor's feedback across conditions, whereas the opposite association was found for NLIFG students. The results also suggest that to the extent that LIFG students endorsed a stronger growth mindset, the more their interpretations of the valence of the advisor's feedback were similar to NLIFG students' interpretations. However, both LIFG and NLIFG students with stronger academic mindsets expected that the hypothetical student would experience a marginally lesser degree of negative effects to their academic engagement. Overall, these findings were consistent with prior work showing that stronger growth mindsets shield marginalized students from the negative effects of stigmatizing experiences through their influence on students' response to stigmatization, but not necessarily students' perceptions of it (Aronson, Fried, & Good, 2002).

Finally, the results of this study failed to provide compelling evidence that students might experience stereotype threat effects as a result of being labeled as *at-risk*. However, there were a few caveats associated with the methods and results of this portion of the study that are worth

noting. First, this study relied on students' written recollections and hypothetical recollections of experiences in which they were labeled by a deficit-oriented descriptor, as a mechanism for activating the saliency of negative stereotypes. This could have been problematic if students failed to carefully think about these experiences and how they felt when they happened (or how they imagined they would have felt). Second, although the manipulation check indicated that the proportion of participants who reported experiencing negative effects following their labeling experience was greater for those in the *deficit-oriented label condition*, participants were not directly asked about their *current* affective state. It is possible that asking participants to indicate their affective state at the time that they were completing the manipulation task might have allowed for a more sensitive assessment of task's effectiveness. Lastly, I did not have a baseline for verbal skills, which would have allowed me to control for differences in students' pre-test verbal skills, which might have been helpful, given the sample differences in scores. Altogether, these limitations suggest that these findings should be interpreted with caution and the possibility that students may experience stereotype threat as a result of being labeled by a deficit-oriented descriptor should not be dismissed without further research.

Limitations. One major limitation of this study was that LIFG students were overrepresented among the participants recruited through MTurk but underrepresented among the participants recruited from Boston College, whereas the reverse was the case for NLIFG students. These disparities are concerning because they represent a confound that was not accounted for in the design of the study and could potentially underlie some of the sample group differences that were observed. The most notable difference between participants recruited from MTurk and BC was that those recruited from MTurk were generally older than students recruited from BC. However, there were also many similarities between recruitment subsamples within the

larger sample groups. For instance, with respect to LIFG students, in addition to sharing first-generation status and eligibility for both the Pell Grant and the Federal Work Study Program, the proportion of male students and Hispanic students were similar across the samples recruited from MTurk and Boston College. For the NLIFG sample, although there were fairly large differences in the proportion of male students, students attending religiously affiliated institutions, and upperclassmen between those recruited from MTurk and Boston College, all NLIFG students identified as non-Hispanic, White, were continuing-generation college students and were not eligible for either the Pell Grant and the Federal Work Study Program (see *Table 5.1* for additional sample demographics). Also note that, to account for variation in recruitment method across the two sample groups, recruitment was included as factor in the various ANCOVA analyses. Another major limitation was the disproportionate sizes of the two sample groups. Although many of the analyses had sufficient sample sizes, for some of the analyses, the sample sizes per condition for LIFG students were reduced to approximately n=33; the results of those analyses should be interpreted with caution.

CHAPTER 6: CONCLUSIONS

The overall aim of this dissertation was to explore the potential consequences associated with using deficit-oriented labels to characterize college students. There is an extensive amount of prior research documenting the adverse psychological, affective, and cognitive effects of applying labels to individuals that are associated with negative stereotypes (e.g., Angermeyer & Matschinger 1997; Martin et al., 2000). Findings from this work indicate that these effects function through social and internal processes of stigmatization (e.g., Link et al., 1999). Moreover, within the educational literature, there is evidence linking college students' academically stigmatizing experiences with lower academic self-perceptions, diminished sense of belonging in college, and stereotype threat effects (e.g., Aronson & Steele, 2005; Walton & Cohen, 2004). There is also a substantial amount of evidence indicating that marginalized college students may be disproportionately affected by these experiences, relative to their White peers from more privileged backgrounds (e.g., Walton & Cohen, 2007; Winograd & Rust, 2014).

Also relevant to the topic of my dissertation is a separate literature focused on advancing awareness of the deficit-thinking that underlies interpretations of marginalized students' academic experiences within higher education (e.g., Castro, 2014; Pearl, 1991; Valencia, 1997; 2010). The scholars contributing to this literature argue that, characterizing these students as 'at-risk' and 'underprepared' while earmarking them for 'interventions,' places the onus of underachievement on the students and deemphasizes the role of structural inequities that contribute these disparities in achievement to begin with. They also argue that framing students in terms of deficit-oriented labels leads to academic stigmatization, as it encourages negative assumptions and expectations about students' motivation and academic potential (Castro, 2014; Pearl, 1991; Valencia, 1997; 2010).

Characterizing marginalized college students by deficit-oriented descriptors may not only pathologize and stigmatize these students, but may also lead students to interpret academic support programs (i.e., interventions) as attempts to 'fix' them. In addition, there is reason to believe that enduring deficit-oriented labeling experiences at a chronic level may diminish students' academic self-perceptions and sense of belonging in college. Considering that both of these outcomes are associated with motivation and persistence in college, the consequences of these labeling experiences may ultimately exacerbate existing disparities in degree attainment. However, despite these important implications, no prior research has attempted to quantify the effects of labeling students as 'at-risk' or 'underprepared' in an empirical manner. This dissertation has focused on filling that gap in the literature.

A total of three studies were conducted—each of which included samples of Black and/or Hispanic first-generation college students from low-income backgrounds (i.e., LIFG students) and White students who were neither first-generation students nor from low-income backgrounds (i.e., NLIFG students). Individually, each of the studies were designed to explore various aspects of this topic—that when combined—would enhance our overall understanding of college students' deficit-oriented labeling experiences.

In Study 1, I conducted a preliminary exploration of several descriptive aspects of students' deficit-oriented labeling experiences—including the frequency with which students endured these experiences, the contexts in which they occurred, and the effects they experienced as a result. This study was also used to pilot materials designed to examine students' interpretations of hypothetical scenarios in which a student was depicted as being labeled as an 'at-risk' versus a 'first-generation' student. In Study 2, I continued to examine descriptive aspects of students' deficit-oriented labeling experiences. This study also explored associations

between students' motivational beliefs (i.e., academic mindsets; racial and ethnic identity beliefs; stereotype vulnerability) and the extent to which they reported experiencing negative affective and motivational consequences as a result of being labeled as 'at-risk', 'underprepared', and/or 'disadvantaged'. Finally, in Study 3, I used hypothetical scenarios—similar to those tested in Study 1—to continue examining students' interpretations of deficit-oriented labels versus alternative options that could potentially be applied in practical settings. I also examined how students' motivational beliefs influenced these interpretations. In addition, Study 3 explored the possibility the students might experience stereotype threat effects as a result of being labeled as 'at-risk'.

With that said, I begin this chapter by addressing the major findings across all three studies and discussing them in the context of both the objectives of this dissertation and prior research. Next, I discuss the limitations of this research and suggest recommendations for addressing them in future work. Finally, I conclude by addressing the implications of this work for students, particularly those from marginalized backgrounds, and for practitioners in higher education working with these students.

Conclusions

Frequency & Contexts of Deficit-Oriented Labeling Experiences

Findings from Studies 1 and 2 showed that compared to NLIFG students, LIFG students were more likely to report that they had experienced being characterized by specific deficit-oriented labels. More specifically, they were more likely to indicate that they experienced being labeled as 'at-risk' and 'disadvantaged' in the past, and reported enduring a greater number of these experiences in the past academic year, on average; the findings for the 'underprepared' label are less consistent across the studies. These findings support what many scholars and

practitioners already presumed to be true—which is, that labels like 'at-risk' are commonly and disproportionately used to characterize low-income students of color and/or first-generation students of color, within higher education (Marger, 1996; Pearl, 1991; Valencia, 1997; 2010). Although this finding may seem rather intuitive, it is also an important one, because prior to this research, no other study had attempted to yield quantitative data to corroborate the anecdotal evidence. Knowing which types of students are most likely to endure this particular type of stigmatization is potentially useful because it can influence the subsequent discussions surrounding the consequences of these stigma experiences and the types of strategies that might be most effective at mitigating them.

Across the same two studies, both LIFG and NLIFG students were generally consistent in their reports that these stigmatizing descriptors were communicated to them by instructors or academic advisors, whereas LIFG students reported that these labels were also communicated during on-campus events (e.g., orientations). These findings are consistent with existing research, in that they suggest that the individuals most likely to characterize students in this manner, are those that work closest with them (Castro, 2014; Gray, 2013). Given that both instructors and advisors play a critical role in promoting student success, these findings have important implications for practice (Cuseo, 2003). For instance, the links between student satisfaction with academic advising and higher retention rates have been well documented by prior research, such that students who are more satisfied with the quality of advising they receive are more likely to meet with their advisors in a consistent manner—and in turn, students who meet with their academic advisors more often tend to show higher retention rates (e.g., Drake, 2011; Tinto, 1999). These associations make sense, because relative to students who only meet with their advisors sporadically (or not at all), those who meet with their advisors more

frequently are more likely to develop trusting relationships with these individuals, which may not only motivate them to continue meeting with them in the future, but also make them more receptive to incorporating any feedback they receive from their advisors (e.g., new study strategies, information about tutoring services, etc.).

Moreover, the associations between advising and retention have been found to be even stronger for students from marginalized groups—which not surprisingly—has led to the common practice of incorporating academic advising into the framework of the types of support programs that often target these students (Smith & Allen, 2014; Tinto, 1999). That said, the findings from the current research suggest that by communicating to marginalized students that they are perceived as being 'at-risk', 'underprepared', or 'disadvantaged', some advisors may be inadvertently squandering the opportunity to develop meaningful relationships with them. Based on findings from prior work, as well as the current findings, there is reason to believe that these 'interactions' might be perceived as threatening by students and promote resentment and distrust. As a result, students may become considerably less motivated to continue meeting with their advisors and/or receptive to any subsequent feedback they receive from these individuals (Inzlitch & Good, 2006).

Consequences of Students' Personal Deficit-Oriented Labeling Experiences

Across Study 1 and 2, there were no differences in the extent to which LIFG and NLIFG students reported negative affective and motivational consequences as a result of being characterized with deficit-oriented labels. Across both groups of students, a greater number of such labeling experiences was associated with a greater degree of negative affect in Study 1 (for three out of four labels) and a greater degree of negative effects on students' sense of belonging, academic self-perceptions, and affect in Study 2. Findings from Study 2 also yielded evidence to

suggest that greater stereotype vulnerability was associated with higher levels of negative academic self-perceptions and negative affect in response to labeling experience. This finding support Castro's (2014) argument that characterizing marginalized students by deficit-oriented descriptors pathologizes them and undermine efforts to help them succeed in college. The findings in relation to racial identity beliefs were somewhat inconsistent, such that stronger racial identification was associated with a greater degree of consequences for LIFG and NLIFG students' academic engagement, but seemed to protect LIFG students only from negative affective consequences of their labeling experiences.

With respect to stereotype vulnerability, there are several ways in which deficit-oriented labeling experiences might increase the saliency of negative academic stereotypes. For instance, even if students are not aware of the specific stereotypes associated with these descriptors, the inherently negative nature of labels like *at-risk* or *underprepared* might be sufficient to make other race- or income-based academic stereotypes salient (Aronson & Steele, 2005). Considering that the LIFG students in this study were affiliated with multiple academically stigmatized groups and showed greater susceptibility to negative academic stereotypes, being characterized by a label that communicates negative information about their academic competence may ultimately compound the motivational effects of their academically stigmatizing experiences (Aronson & Steele, 1995), though group differences in the self-reported motivational consequences of labeling were not found in the current studies.

Additionally, endorsing stronger growth mindsets were negatively associated with the effects of deficit-oriented labeling experiences on the academic self-perceptions of students in both groups. These results are consistent with findings from other work showing that stronger growth mindsets are generally beneficial to all students and can help mitigate the effects of

stereotype threat on motivation and achievement outcomes (e.g., Aronson, Fried, & Good, 2002; Good, Rattan, & Dweck, 2012). This finding, though tentative, is encouraging because its suggests that promoting the endorsement of stronger growth mindsets may help students develop a more resilient disposition that can buffer them from the negative effects of these labeling experiences.

Deficit-Oriented Labels & Stereotype Threat Effects

The current dissertation also explored the possibility that recalling a deficit-oriented labeling experience might trigger effects of stereotype threat on students' subsequent verbal test scores. Although the findings did not yield evidence to support this idea, it is also possible that the effects of briefly recalling a stigmatizing experience—and in some cases, a hypothetical experience—may not have been strong enough to generate the cognitive imbalance required to trigger the physiological, cognitive, and self-regulatory processes that are said to potentially drive the effects of stereotype threat on subsequent performance (e.g., Schmader, Johns, & Forbes, 2008). However, I address this point further in my discussion of the limitations of this research and suggest some recommendations for addressing these issues in future work.

Students' Interpretations of Advisor's Use of Deficit-Oriented & Neutral Labels

Although the findings were slightly inconsistent across Studies 1 and 3 with respect to the perceived valence of the advisor's feedback, participants in both studies expected the affective and motivational consequences of the feedback to be more negative when it included a deficit-oriented label compared to when it included a neutral label. Interestingly, a more encouraging finding was that LIFG students were more likely than NLIFG students to pick up on subtle differences in the way the advisor *intentionally* characterized the student when providing them with feedback. When the feedback characterized the student as a 'first-generation student' (i.e.,

the neutral label condition), LIFG students interpreted the advisor as being more intentional in trying to convey their positive beliefs about the student. However, the adding context to the deficit-oriented label (i.e., deficit-oriented label + context condition) did lead to the same positive interpretation.

In contrast to LIFG students, it seems that NLIFG students were generally less likely to either perceive nuanced differences in the advisor's intentionality or to take them into account when interpreting the advisor's feedback. In addition, NLIFG participants in Study 3 (but not Study 1) seemed to interpret the advisor's feedback more negatively than LIFG students across conditions. These effects could potentially be explained by the fact that NLIFG students had less of exposure to these types of labels on average, as shown in Studies 1 and 2. Another possibility is that NLIFG and LIFG students may have interpreted the hypothetical scenarios in the context of the university or college they currently attend, such that differing characteristics between these contexts could influence the manner in which students interpret these situations. Given that the majority of NLIFG students for Study 3 were recruited from Boston College—a private elite college where their low-income, first-generation peers are heavily underrepresented—this context may lend itself to a more negative interpretation of any situation in which an authority figure characterizes a marginalized student in a deficit-oriented manner. More specifically, these students might interpret the advisor's feedback as demeaning a student who is already marginalized within that context. However, it is worth noting these sample group patterns of interpretations (i.e., NLIFG students' pattern of more negative interpretations and LIFG students' more positive interpretations) were present within the samples within each recruitment method as well, which suggests that differences in the types of institutions NLIFG and LIFG students are

attending are likely not the driving force for these findings. Moreover, the effects of sample group in Study 3 emerged after accounting for recruitment method in the analysis.

With respect to the difference in how LIFG students interpreted the advisor's intentions in using the label 'at-risk' versus 'first-generation,' prior research exploring *self-affirming* strategies to reduce stereotype threat offer some potential insight (e.g., Cohen et al., 2006; Martens, Johns, Greenberg, & Schimel, 2006). In broad terms, self-affirming strategies—which draw on Steele's (1997) Self-Affirmation Theory—operate under the assumptions that by encouraging students to reflect on their strengths (e.g., values; characteristics; skills), they can reinforce their self-worth under conditions that might otherwise present challenges that could potentially diminish it. For instance, in a study by Cohen and colleagues (2006), the authors tested a short 'self-affirming' intervention for reducing stereotype threat in African American high school students. To do this, they assigned White and African American students in their sample to one of two conditions. Students in the experimental intervention condition (i.e., self-affirming task) were presented with a list of values (e.g., "My family is important to me") and then asked to choose the value that was most important to them and write a brief explanation as to why it was important to them, whereas students in the control condition were presented with the same list of values but asked to select the value that was least important to them but might be important to others, and then wrote about why that value might be important to other people. The authors found that for African American students, those who wrote about their own values showed significant increases in GPA, relative to African American students in the control condition.

In regard to the current research, it is possible that in contrast to the 'at-risk' label (when used ambiguously), the 'first-generation' label may have a self-affirming element that the former

label lacks. For example, although the first-generation label is undoubtedly associated with negative stereotypes (e.g., Gray, 2013), it is also a label that distinguishes these students as having accomplished something that no one else in their family had up until that point. Thus, it is possible that LIFG who read the scenario in which the advisor characterizes the students as a 'first-generation student' may have been interpreted the advisor's feedback as their way of trying to highlight what the student had already accomplished by being the first in their family to attend college and as conveying their confidence in the student's ability to overcome any struggles they were currently experiencing. In contrast, LIFG students failed to ascribe this same level of intention to the advisor's ambiguous characterization of the student as 'at-risk/underprepared'. Moreover, unlike NLIFG students, LIFG students could personally relate to identifying with the 'first-generation' label, which may explain why they were increasingly likely to make these distinctions in their interpretations of the advisor's feedback. Lastly, similar to the findings from Study 2, the findings from Study 3 also suggest that students' motivational beliefs influence the manner in which students respond to academically stigmatizing experiences, or in this case, how they expect other students would react to these experiences.

Limitations

The research conducted for this dissertation has several limitations that are worth noting. One, administering the studies fully online, rather than having participants complete them from a computer in the lab, had its benefits and disadvantages. I conducted the studies online in hopes that it would extend my reach in terms of the number of participants that could be recruited to complete them (particularly with respect to LIFG students). That said, conducting the studies in this manner limited my control over who completed the studies and the extent to which they were carefully reading and processing survey instructions, prompts, and response items—all of

which could have been controlled for to at least some extent in a laboratory setting. Conducting the study as a survey also meant I was limited to self-reported data, which are associated with participant biases, such as social desirability effects (see Joinson, 1999). Another limitation was with respect to the sources used to recruit participants for all three studies. Given that the participants for Studies 1 and 2 were all recruited from MTurk, in Study 3, I attempted to recruit students from various colleges and universities through emails to course listservs and emails associated with on-campus organizations for undergraduate students (e.g., student government associations). However, the only emails that received responses were those to course listservs at Boston College, which introduced another limitation, because most of the participants recruited through listservs at Boston College were NLIFG students, which made the proportion of NLIFG and LIFG students imbalanced across both recruitment methods.

Yet another limitation associated with the samples of students used for this research was that it excluded students who did not meet the specific criteria for the LIFG and NLIFG student samples. Although this was done purposely due to a limited amount of funds to pay participants and to contrast the experiences of students who might be the most and least likely to be labeled by deficit-oriented descriptors, operating under these constraints came at a cost. By excluding all other students, the insights yielded through this research can only be generalized to other LIFG and NLIFG college students, which obviously limit the reach of this research. However, given that this research was the first to explore this issue, there are several ways in which future research could use the findings from this dissertation as a basis from which to build on.

For instance, a recent report by the Pew Research Center (PRC) indicated that Black men reported being disproportionately stigmatized in certain contexts, relative to Black women (PRC, 2019). Given this data, it would be important to explore the possibility that students' gender may

influence the likelihood that they are characterized by a deficit-oriented descriptor. Although the current research did include gender in most analyses, given the limitations in sample size, it was for the purpose of controlling for gender, rather than examining the effects of gender. Moreover, although there is often a considerable amount of overlap between students who come from low-income backgrounds and those that are first-generation students, we would benefit from understanding if one of these factors disproportionately influence the likelihood that students will be characterized by these labels—or even the extent to which they influence the manner in which students interpret these experiences. Given that the 'first-generation' label is unique, in that it is associated with negative stereotypes but also represents a meaningful accomplishment on the part of the student, students who are from low-income backgrounds but not first-generation college students may interpret these labels differently, compared to those who are both.

Another set of limitations is associated with some of the methods used to conduct this research. One, although relying on hypothetical scenarios to specifically assess students' interpretations of deficit-oriented labels versus alternative options was practical, it was certainly not ideal because it did not draw on students' own experiences and therefore limited the interpretation of the findings. However, given that I was interested in examining NLIFG students' perceptions of LIFG students' deficit-oriented labeling experiences, the nature of this task had to be hypothetical to some extent. That said, future research could address this limitation in one of two ways. The first—and most ideal—approach would be to draw on students' actual experiences of being labeled by a deficit-oriented descriptor. This strategy would require comprehensive recruitment efforts, but would likely yield the most valuable insights. Conversely, future studies could continue to explore students' interpretations through hypothetical scenarios, but employ creative methods to make these scenarios more realistic and

relevant to students' themselves. For instance, one potential way to do this would be to show participants a video where an individual portraying the part of the advisor provides the participant with the same type of feedback the advisor provided the hypothetical student within the current research. Receiving the feedback themselves from an actual person would likely elicit more genuine interpretations of that feedback.

Lastly, another limitation of this research, which was noted earlier in this chapter, relates to the methodology used to activate stereotype threat. Given that recalling a brief (and potentially hypothetical) experience may not have been powerful enough to elicit stereotype threat effects on students' subsequent performance, future research could address this limitation a couple of different ways. One possibility would be to induce a deficit-oriented labeling experience in-person—which is a strategy often used in stereotype threat research. For instance, participants could be invited to an information session to hear about a program they 'qualify' to participate in, and during this brief information session, the experimenter could characterize the program as one that helps promote academic success for students who might be considered as 'at-risk'. Alternatively, for online studies, this could also be done using videos, similar to the hypothetical advising session previously described.

It is also possible that slight refinements in the methodology used in this research might successfully yield evidence of stereotype threat effects. For instance, in lieu of limiting the amount of time students spend describing their experience of being labeled as 'at-risk' to a few minutes (as was the case with this study), future research could ask participants to write about their experience for 5 or more minutes, to ensure that they have spent a sufficient amount of time thinking about how they felt. Moreover, asking participants to rate their current affective state following the manipulation might be more effective in gauging its effectiveness than basing that

assessment on whether participants indicated they experienced negative effects as a result of their labeling experience. And finally, considering the possibility that recalling an experience may not yield effects easily measured on brief subsequent analytical tasks, future studies could utilize either longer achievement tasks or employ more subtle measures of stereotype threat effects, such as measuring participants' avoidance of challenge.

Implications for Practice

In conjunction with the anecdotal evidence, the findings from this dissertation highlight the necessity for promoting the success of all students in higher education by *empowering*—rather than pathologizing them. To achieve this, we must start by eliminating the discourse that frames low-income, first-generation students of color as burdens that their institutions must contend with, as well as the ambiguous use of deficit-oriented labels that sustains this interpretation. An easy start would be to change the way we frame academic support programs within higher education. For instance, rather than referring to these programs as 'interventions, we can simply refrain from qualifying these initiatives; at the very least, students should be made explicitly aware of the concrete criteria used to target participants for these programs.

Further, empowering all college students will require nurturing their resilience to adversity, reaffirming the value they add to their institution's community, and encouraging them to believe in their capacity to control their academic outcomes. These are bold objectives that require dynamic solutions, but there is an abundance of novel research being conducted on various fronts of the educational literature which we can draw on for guidance (e.g., Broda et al., 2018; Cerezo & McWhirter, 2012; Davidson, Feldman, & Margalit, 2012; Yeager & Dweck, 2012). Given that the current research was the first to examine the potential consequences of labeling students as 'at-risk' and 'underprepared,' further research is needed that explores

potential strategies to help students overcome academically stigmatizing experiences. However, the findings from both the prior work and the current research suggest that self-affirming strategies and methods designed to promote students' endorsement of a growth mindset may be particularly promising.

Given what we know about the influence that academic advisors and instructors have over students' outcomes, the findings from this dissertation underscore the importance of addressing this issue from the practitioners' perspective as well. The encouraging news is that the there is some evidence to suggest that more often than not, deficit-oriented characterizations are driven by genuine concern and interest in helping students succeed (Castro, 2014). Assuming this is the case, the findings from this research suggest that either practitioners use these descriptors so often with colleagues that they may not be explicitly aware that they are also using them during their interactions with students, or they are purposely using these descriptors during their interactions with students because they have not fully considered the negative implications associated with doing so. Either way, this demonstrates a pressing need for professional development that is specifically designed to tackle this issue head on, which could be accomplished in several different ways.

One potential approach could involve presenting data to practitioners that provide them with some type of evidence that shows them (a) that these labels are often communicated to students by individuals in similar roles and (b) that these experiences are associated with real consequences for students. A bolder strategy would be to have students from within the institution speak to practitioners—perhaps as part of a panel—in relation to their experiences of having these labels communicated to them, including the impact these experiences had on them. Given that the majority of individuals who work closely with students in higher education do so

because they genuinely care about helping students succeed, presenting them with some evidence that they may be inadvertently undermining their efforts to help these students may motivate them to take action. Although findings from the current research did identify one potential alternative for characterizing students in a more constructive manner, this strategy would need to be investigated further prior to being recommended for use by practitioners in a variety of contexts.

Another possible approach would be to bring educational researchers whose work focuses on educational deficit-thinking to speak with practitioners. The goal of this strategy would be twofold. One, the ‘experts’ would be there for the purpose of facilitating practitioners’ thinking in regard to the ways in which labels like ‘at-risk’ and ‘underprepared’ characterize students in a manner that reflects deficit-thinking. Two, given the perceived disconnect between the current form of deficit-thinking and its links to cultural deprivation and genetic inferiority theories, providing practitioners with information related to the framework’s history and evolution might prove to be a much needed, eye-opening experience for educators and practitioners. Finally, given that it is likely that many practitioners do use these labels or communicate them to students, institutions can also choose to adopt broader strategies. For instance, providing professional development that centers around the influence of language and framing on student motivation would presumably benefit any practitioner that works closely with a large and/or diverse population of students.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA, US: Sage Publications, Inc.

Alonso, J., Buron, A., Rojas-Farreras, S., De Graaf, R., Haro, J. M., De Girolamo, G., ... & ESEMeD/MHEDEA 2000 Investigators. (2009). Perceived stigma among individuals with common mental disorders. *Journal of Affective Disorders, 118*, 180-186.

Angermeyer, M. C., & Matschinger, H. (2005). Labeling—stereotype—discrimination. *Social Psychiatry and Psychiatric Epidemiology, 40*, 391-395.

Apple, M. W. (2007). Ideological success, educational failure? On the politics of No Child Left Behind. *Journal of Teacher Education, 58*, 108-116.

Aronson, J. (2004). The threat of stereotype. *Educational Leadership, 62*, 14-20.

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on african american college students by shaping theories of intelligence.*Journal of Experimental Social Psychology, 38*, 113-125.

Aronson, J., & Inzlicht, M. (2004). The ups and downs of attributional ambiguity: Stereotype vulnerability and the academic self-knowledge of african american college students. *Psychological Science, 15*, 829-836.

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*, 29-46.

Aronson, J., & Steele, C.M. (2005). Stereotypes and the fragility of human competence, motivation, and self-concept. In C. Dweck & E. Elliot (Eds.), *Handbook of Competence & Motivation,* pp. 436-456. New York: Guilford.

Banks, J. A. (1993). Multicultural education: Historical development, dimensions, and practice. *Review of Research in Education, 19*, 3-49.

Banks, M., & Woolfson, L. (2008). Why do students think they fail? The relationship between attributions and academic self-perceptions. *British Journal of Special Education, 35*, 49-56.

Barnard, L., Burley, H., Olivarez, A., & Crooks, S. (2008). Measuring vulnerability to stereotype threat. *Electronic Journal of Research in Educational Psychology, 6*, 51-64.

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*, 246-263.

Butler-Barnes, S. T., Leath, S., Williams, A., Byrd, C., Carter, R., & Chavous, T. M. (2018). Promoting resilience among African American Girls: Racial identity as a protective factor. *Child Development, 89*, 552-571.

Bauer, L., & Schanzenbach, D. W. (2016). The long-term impact of the Head Start program. *The Hamilton Project*.

Beasley, M. A., & Fischer, M. J. (2012). Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education, 15*, 427-448.

Bedini, L. A. (2000). Just sit down so we can talk: Perceived stigma and the pursuit of community recreation for people with disabilities. *Therapeutic Recreation Journal, 34,* 55-68.

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*, 317-338.

Cadinu, M., Maass, A., Lombardo, M., & Frigerio, S. (2006). Stereotype threat: The moderating role of locus of control beliefs. *European Journal of Social Psychology, 36*, 183-197.

Cahalan, M., Perna, L. W., Yamashita, M., Ruiz, R., & Franklin, K. (2017). Indicators of Higher Education Equity in the United States: 2017 Trend Report, Washington, DC: Pell Institute for the Study of Higher Education, Council for Education Opportunity (COE) and Alliance for Higher Education and Democracy (AHEAD) of the University of Pennsylvania.

Camp, D. L., Finlay, W. M. L., & Lyons, E. (2002). Is low self-esteem an inevitable consequence of stigma? An example from women with chronic mental health problems. *Social Science & Medicine, 55*, 823-834.

Castro, E. L. (2014). "Underprepared" and "at-risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *Journal of Student Affairs Research and Practice, 51*, 407–419.

Cerezo, A., & McWhirter, B. (2012). A brief intervention designed to improve social awareness and skills to improve Latino college student retention. *College Student Journal, 46*, 867-879.

Chavous, T. M., Hilkene Bernat, D., Schmeelk-Cone, K., Caldwell, C. H., Kohn-Wood, L., & Zimmerman, M. A. (2003). Racial identity and academic attainment among African American adolescents. *Child Development, 74*, 1076–1090.

Chavous, T. M., Rivas-Drake, D., Smalls, C., Griffin, T., & Cogburn, C. (2008). Gender matters, too: The influences of school racial discrimination and racial identity on academic

engagement outcomes among African American adolescents. *Developmental Psychology, 44*, 637–654.

Connors, M. H., & Halligan, P. W. (2015). A cognitive account of belief: A tentative road map. *Frontiers in Psychology, 5,* 1588.

Colbeck, C. L., Cabrera, A. F., & Terenzini, P. T. (2001). Learning professional confidence: Linking teaching practices, students' self-perceptions, and gender. *The Review of Higher Education, 24*, 173-191.

Corrigan, P. W., Larson, J. E., & Rusch, N. (2009). Self-stigma and the "why try" effect: impact on life goals and evidence-based practices. *World Psychiatry*, *8*, 75–81.

Corrigan, P. W., & Rao, D. (2012). On the self-stigma of mental illness: Stages, disclosure, and strategies for change. *The Canadian Journal of Psychiatry, 57*, 464-469.

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*, 588-594.

Croninger, R. G., & Lee, V. E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record, 103*, 548-581.

Davidson, O. B., Feldman, D. B., & Margalit, M. (2012). A focused intervention for 1st-year college students: Promoting hope, sense of coherence, and self-efficacy. *The Journal of Psychology, 146*, 333-352.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19*, 109-134.

Donovan, J., & Johnson, L. (2005). Exploring the experiences of first-generation, multiethnic undergraduate college students. *Journal of Student Affairs*, 14, 32-38.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256.

Elbaum, B., Vaughn, S., Tejero-Hughes, M., & Watson-Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92,* 605.

Elbaum, B. (2002). The self–concept of students with learning disabilities: A meta–analysis of comparisons across different placements. *Learning Disabilities Research & Practice, 17,* 216-226.

Erickson, F. (1987). Transformation and school success: The politics and culture of educational achievement. *Anthropology & Education Quarterly, 18*, 335-356.

Field, S., Hoffman, A., Peter, S. S., & Sawilowsky, S. (1992). Effects of disability labels on teachers' perceptions of students' self-determination. *Perceptual and Motor Skills, 75*, 931-934.

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education, 75*, 203-220.

Fusarelli, L. D. (2004). The potential impact of the No Child Left Behind Act on equity and diversity in American education. *Educational Policy, 18*, 71-94.

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28,* 659-670.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*, 645-662.

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,* 700-717.

Graham, S. (1984). Communicating sympathy and anger to Black and White children: The cognitive (attributional) antecedents of affective cues. *Journal of Personality and Social Psychology, 47,* 40-54.

Gray, S., S. (2013). Framing "at risk" students: Struggles at the boundaries of access to higher education. *Children and Youth Services Review*, *8,* 1245-1251.

Gray, S. W., & Klaus, R. A. (1970). The early training project: A seventh-year report. *Child Development, 41*, 909-924.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology, 95*, 124-136.

Gummadam, P., Pittman, L. D., & Ioffe, M. (2016). School belonging, ethnic identity, and psychological adjustment among ethnic minority college students. *The Journal of Experimental Education, 84*, 289-306.

Guyll, M., Madon, S., Prieto, L., & Scherr, K. C. (2010). The potential roles of self-fulfilling prophecies, stigma consciousness, and stereotype threat in linking Latino/a ethnicity and educational outcomes. *Journal of Social Issues, 66*, 113-130.

Haring, K. A., Lovett, D. L., Haney, K. F., Algozzine, B., Smith, D. D., & Clarke, J. (1992). Labeling preschoolers as learning disabled: A cautionary position. *Topics in Early Childhood Special Education, 12*, 151-173.

Harper, B. E., & Tuckman, B. W. (2006). Racial identity beliefs and academic achievement: does being black hold students back? *Social Psychology of Education, 9*, 381-403.

Hausmann, L. R., Ye, F., Schofield, J. W., & Woods, R. L. (2009). Sense of belonging and persistence in White and African American first-year students. *Research in Higher Education, 50*, 649-669.

Hatzenbuehler, M. L. (2009). How does sexual minority stigma "get under the skin"? A psychological mediation framework. *Psychological Bulletin, 135*, 707-730.

Inzlicht, M., & Good, C. (2006). How environments can threaten academic performance, self-knowledge, and sense of belonging. In *Stigma and Group Inequality*, pp. 143-164. Psychology Press.

Jameson, M. M., & Fusco, B. R. (2014). Math anxiety, math self-concept, and math self-efficacy in adult learners compared to traditional undergraduate students. *Adult Education Quarterly, 64,* 306-322.

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. Behavior *Research Methods, Instruments, & Computers, 31*, 433-438.

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter?. *Learning and Individual Differences, 25*, 67-72.

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' math performance. *British journal of Educational Psychology, 77*, 323-338.

Kim, J. S., & Sunderman, G. L. (2005). Measuring academic proficiency under the No Child Left Behind Act: Implications for educational equity. *Educational Researcher, 34*, 3-13.

Lai, E. R. (2011). Motivation: A literature review research report. Pearson's Publication, 4-36.

Latalova, K., Kamaradova, D., & Prasko, J. (2014). Perspectives on perceived stigma and self-stigma in adult male patients with depression. *Neuropsychiatric Disease and Treatment, 10,* 1399-1405.

Leyens, J. P., Désert, M., Croizet, J. C., & Darcis, C. (2000). Stereotype threat: Are lower status and history of stigmatization preconditions of stereotype threat? *Personality and Social Psychology Bulletin, 26*, 1189-1199.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology, 27*, 363-385.

Lundell, D. B., & Collins, T. (1999). Toward a theory of developmental education: The centrality of discourse. *The Expanding Role of Developmental Education*, 3-20.

Macaruso, P., Hook, P. E., & McCabe, R. (2006). The efficacy of computer-based supplementary phonics programs for advancing reading skills in at-risk elementary students. *Journal of Research in Reading, 29*, 162-172.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods, 44*, 1-23.

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80,* 366.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397-416.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*, 236-243.

Maxwell, M. (1997). *Improving student-learning skills. From Access to Success: A Book of Readings on College Developmental Education and Learning Assistance Programs*. Clearwater, FL: H&H Publishers.

McKown, C., & Weinstein, R. S. (2002). Modeling the Role of Child Ethnicity and Gender in Children's Differential Response to Teacher Expectations 1. *Journal of Applied Social Psychology, 32*, 159-184.

Meece, J. L., & Courtney, D. P. (1992). Gender differences in students' perceptions: Consequences for achievement-related choices. In D. H. Schunk & J. L. Meece (Eds.), *Student Perceptions in the Classroom* (pp. 209-228). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Meltzer, L., Reddy, R., Pollica, L. S., Roditi, B., Sayer, J., & Theokas, C. (2004). Positive and negative self-perceptions: Is there a cyclical relationship between teachers' and students'

perceptions of effort, strategy use, and academic performance? *Learning Disabilities Research & Practice, 19*, 33-44.

Moore, K. E., Stuewig, J. B., & Tangney, J. P. (2016). The effect of stigma on criminal offenders' functioning: A longitudinal mediational model. *Deviant Behavior, 37*, 196-218.

Morrow, J., & Ackermann, M. (2012). Intention to persist and retention of first-year students: The importance of motivation and sense of belonging. *College Student Journal, 46,* 483-491.

Miu, A. S., & Yeager, D. S. (2015). Preventing symptoms of depression by teaching adolescents that people can change: Effects of a brief incremental theory of personality intervention at 9-month follow-up. *Clinical Psychological Science, 3*, 726-743.

Ogbu, J. U. (1979). Minority education and caste: The American system in cross-cultural perspective. *Crisis, 86*, 17-21.

O'Keeffe, P. (2013). A sense of belonging: Improving student retention. *College Student Journal, 47*, 605-613.

Osterholm, K., Nash, W. R., & Kritsonis, W. A. (2011). Effects of Labeling Students" Learning Disabled": Emergent Themes in the Research Literature 1970 Through 2000. *FOCUS on Colleges, Universities & Schools, 6*, 1-11.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly, 19*, 139-158.

Pattyn, E., Verhaeghe, M., Sercu, C., & Bracke, P. (2014). Public Stigma and Self-Stigma: differential association with attitudes toward formal and informal help seeking. *Psychiatric Services*, *65*, 232–238.

Pearl, A. (1991). 'Systemic and institutional factors in Chicano school failure'. In *Chicano School Failure and Success: Research and Policy Agendas for the 1990s*, pp. 273-320, London, England: Falmer Press.

Pellegrini, A. D. (1991). A critique of the concept of at risk as applied to emergent literacy. *Language Arts, 68*, 380-385.

Pettigrew, T. F. (1964). Race, mental illness, and intelligence: A social psychological view. *Eugenics Quarterly, 11*, 189-215.

Pew Research Center (2019). *Race in America 2019*. Retrieved from: https://www.pewsocialtrends.org/2019/04/09/race-in-america-2019-methodology/.

Phelan, J. E., & Basow, S. A. (2007). College students' attitudes toward mental illness: An examination of the stigma process. *Journal of Applied Social Psychology, 37*, 2877–2902.

Phelan, S. M., Burgess, D. J., Puhl, R., Dyrbye, L. N., Dovidio, J. F., Yeazel, M., ... & Burke, S. E. (2015). The adverse effect of weight stigma on the well-being of medical students with overweight or obesity: Findings from a national survey. *Journal of General Internal Medicine, 30*, 1251-1258.

Placier, M. (1996). The cycle of student labels in education: The cases of culturally deprived/disadvantaged and at risk. *Educational Administration Quarterly, 32*, 236.

Pollard, D. S. (1993). Gender, achievement, and African-American students' perceptions of their school experience. *Educational Psychologist, 28,* 341-356.

Putwain, D. W., Woods, K. A., & Symes, W. (2010). Personal and situational predictors of test anxiety of students in post-compulsory education. *British Journal of Educational Psychology, 80,* 137-160.

Reisig, M. D., Bales, W. D., Hay, C., & Wang, X. (2007). The effect of racial inequality on black male recidivism. *Justice Quarterly, 24*, 408-434.

Rhee, B. S., & Hurtado, S. (2009). *Practical and statistical issues in assessing college student outcomes: Evaluating the use of survey measures and standard cognitive measures.* Paper presented at the Association for Institutional Research Annual Forum, Atlanta, GA.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130,* 261–288.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68-78.

Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles, 66*, 175-183.

Shifrer, D. (2013). Stigma of a Label: Educational Expectations for High School Students Labeled with Learning Disabilities *Journal of Health and Social Behavior, 54,* 462 – 480.

Soria, K. M., & Stebleton, M. J. (2012). First-generation students' academic engagement and retention. *Teaching in Higher Education, 17*, 673-685.

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

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

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education, 17*, 213-226.

Tepper, K. (1994). The role of labeling processes in elderly consumers' responses to age segmentation cues. *Journal of Consumer Research, 20*, 503-519.

Valencia, R. (2010). *Dismantling contemporary deficit thinking: Educational thought and practice*. New York, NY: Routledge.

Walters, P. B. (2001). Educational access and the state: Historical continuities and discontinuities in racial inequality in American education. *Sociology of Education, 74*, 35-49.

Walton, G. M., & Carr, P. B. (2012). Social belonging and the motivation and intellectual achievement of negatively stereotyped students. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, Process, and Application* (pp. 89-106). New York, NY, US: Oxford University Press.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *Journal of Personality and Social Psychology, 92*, 82-96.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331*, 1447-1451.

Winne, P. H., Woodlands, M. J., & Wong, B. Y. (1982). Comparability of self-concept among learning disabled, normal, and gifted students. *Journal of Learning Disabilities, 15*, 470-475.

Winograd, G., & Rust, J. P. (2014). Stigma, Awareness of Support Services, and Academic Help-Seeking Among Historically Underrepresented First-Year College Students. *The Learning Assistance Review, 2*, 19-43.

Woodcock, A., Hernandez, P. R., Estrada, M., & Schultz, P. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology, 103*, 635.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302-314.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., ... & Trott, J. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*, 374.

Yosso, T. J. (2005). "Whose Culture Has Capital? A Critical Race Theory Discussion of Community Cultural Wealth.". *Race, Ethnicity and Education*, *8*, 69–91.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3-17.

Appendix A

Glossary of Terms

Term	Definition
Attainment Gaps	Disparities in 6-year college graduation rates between students from different backgrounds, where the differences in this outcome are statistically significant (NCES, 2016).
Four-Year Institution	Universities and colleagues that offer at least one 4-year program of college-level studies (NCES, 2016).
Full-Time Enrollment	Being enrolled in a total credit load equal to at least 75 percent of the normal full-time course load. At the undergraduate level, full-time enrollment typically includes students who have a credit load of 12 or more semester or quarter credits (NCES, 2016).
Black or African American Undergraduate Students	Undergraduate students having origins in any of the black racial groups of Africa and is used interchangeably with the shortened term Black (NCES, 2016).
Hispanic or Latino Undergraduate Students	Undergraduate students of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race (NCES, 2016).
Retention Rates	A measure of the rate at which students persist in their educational program at an institution and are expressed as the percentage of first-time bachelor's (or equivalent) degree-seeking undergraduates from the previous fall who are again enrolled in the current fall (NCES, 2016).
White Undergraduate Students	Undergraduate students having origins in any of the original peoples of Europe, the Middle East, or North Africa (NCES, 2016).
First-Generation College Students	Undergraduate students for whom neither parent has a college degree from a 4-year institution (Cahalan et al., 2017).
Low-Income College Students	Students who are eligible to receive the Pell Grant or other Federal Grants that are not required to be paid back, or as those whose families have a combined income that fell under $37, 679.00 (Cahalan et al., 2017).

Appendix B

Study 1: Pre-Screening Questionnaire

a) Are you currently a full-time student at a 4-year university or college?
- ☐ Yes
- ☐ No

b) Are you currently eligible for the Federal Pell Grant?
- ☐ Yes
- ☐ No

c) Are you currently eligible for Federal Work Study?
- ☐ Yes
- ☐ No

d) What is the highest level of education completed by your mother or female guardian?
- ☐ Less than high school completed
- ☐ High school diploma or equivalent
- ☐ Some college, vocational, or trade school (including 2-year degrees)
- ☐ Bachelor's degree (e.g., BS, BA, AB)
- ☐ Master's degree (e.g., MS, MA, MBA)
- ☐ Professional degree (e.g., JD, LLB, MD, DDS, DVM)
- ☐ Doctorate (e.g., PhD, DSc, EdD)

e) What is the highest level of education completed by your father or male guardian?
- ☐ Less than high school completed
- ☐ High school diploma or equivalent
- ☐ Some college, vocational, or trade school (including 2-year degrees)
- ☐ Bachelor's degree (e.g., BS, BA, AB)
- ☐ Master's degree (e.g., MS, MA, MBA)
- ☐ Professional degree (e.g., JD, LLB, MD, DDS, DVM)
- ☐ Doctorate (e.g., PhD, DSc, EdD)

f) Are you Hispanic or Latino (i.e., a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race)?

g) Please indicate your racial background (select all that apply):
- ☐ American Indian or Alaska Native: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

- ☐ Asian: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
- ☐ Black or African American: A person having origins in any of the black racial groups of Africa.
- ☐ Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
- ☐ White: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

h) Please indicate your age: (dropdown list; *"under 18" – "over 80"*):

i) Please indicate your sex:
- ☐ Male
- ☐ Female

j) Is your college or university a "public" or "private" institution?
- ☐ Yes
- ☐ No

k) Is your college or university religiously affiliated?
- ☐ Yes
- ☐ No

Appendix C

Study 1 Survey

Section 1

Prompt 1

"We are interested in specific labels, such as "at risk" or "underprepared," that may have been used to negatively characterize your academic potential or the academic potential of people like you. These labels may have been used by other people at your college (such as instructors, advisors, administrators, peers, etc.) or included in communications disseminated by the college (e.g., webpages, emails, brochures, etc.)"

Take a couple of minutes to think about this and list a maximum of (5) negative labels, that if possible, are different than the ones provided above. You can provide labels you have heard used by others or seen used in brochures or online websites—used to describe the academic potential that you and other students like you possess.

Please note: we are NOT interested in general adjectives commonly used to describe students based on personal attributes—such as "motivated", "lazy", or "dumb".

Prompt 2

Here are the labels you provided:

[Label 1]
[Label 2]
[Label 3]
[Label 4]
[Label 5]

Using the text box below, please provide some examples of:

1) The people who have used these labels to describe you (e.g., peers, faculty, school staff, advisor)
2) The contexts in which the labels were communicated to you (e.g., one-on-one meeting with an instructor or advisor, a group orientation, online on your school's website, program brochure, etc.)

Prompt 3

Here are the labels you provided:

[Label 1]
[Label 2]
[Label 3]
[Label 4]
[Label 5]

Using the text box below, please provide some examples of:

3) How these experiences of being categorized by these labels made you feel (e.g., happy, sad, frustrated, encouraged)

Section 2

Items

1. How often (if ever) have you heard yourself or other students like you described as [LABEL]? [1= "Never" to 5= "Often"]

1a. If you have heard yourself or other students like you described as an "at-risk student": In what context(s) did this occur? (Check all that apply)

- ☐ On-campus Event (e.g., orientation, first-year experience)
- ☐ On-campus Academic Advising Session
- ☐ On-campus Academic Support Services (e.g., tutoring)
- ☐ In class (college level)
- ☐ At home
- ☐ High School
- ☐ University or College Website
- ☐ Other or Not Applicable (please specify)

1b. Who referred to you or students like you with the label [LABEL]? (Check all that apply)

- ☐ Peers/Classmates
- ☐ Academic Advisor
- ☐ Instructor
- ☐ University or College Staff or Administrators
- ☐ Roommates
- ☐ High School Teacher or Counselor
- ☐ Parents
- ☐ Other or Not Applicable (please specify)

1c. How negatively did it make you feel to hear yourself or other students like you described as an [LABEL]? [1= "Not at all negative" to 6= "Extremely negative"; or "Not Applicable"]

Section 3

Scenario

"Aaron [April] is a college freshman. He [she] graduated at the top of his [her] class from a public high school in a working-class county, and is the first person in his [her] family to attend college. Today, Aaron [April] met with his [her] advisor to discuss which classes he [she] should register for in the spring. As they were wrapping up the advising session, Aaron's [April's] advisor gave him [her] the feedback below.

Advisor: I've received some feedback from a few of your professors, who said that you've scored below average on some of your exams and assignments. You know, a lot of my at-risk students [first-generation students] get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."

Items

Instructions: Now, please imagine that you are Aaron and respond to the following questions about this scenario from his perspective.

1. Please indicate the extent to which you believe that Aaron's [April's] advisor is communicating positive or negative beliefs about Aaron [April] with the statement below: [1 = "Very Negative Beliefs" – 6 = "Very Positive Beliefs"]

   "You know, a lot of my at-risk students [first-generation students] get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."

2. Next, please indicate the extent to which you believe that Aaron's advisor is intentionally communicating positive or negative beliefs about Aaron with the statement below: [1 = "Very Negative Beliefs" – 6 = "Very Positive Beliefs"]

   "You know, a lot of my at-risk students [first-generation students] get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."

3. Please indicate your level of agreement or disagreement with the following

statements. [1 = “Strongly Disagree” – 6 = “Strongly Agree”]

a. The advisor’s feedback probably made Aaron feel like doing well in school isn’t really that important.
b. The advisor’s comments probably made Aaron feel less confident about his ability to do well in college-level courses.
c. The advisor’s comments probably made Aaron feel frustrated.

d. The advisor’s comments probably made Aaron feel like he doesn’t belong at that university.
e. The advisor’s comments probably made Aaron feel like he can overcome his academic challenges if he works hard.
f. The advisor’s comments probably made Aaron feel like he is not as smart as most of the students at that university.
g. The advisor’s comments probably made Aaron feel like he needs more help than most students at that university in order to do well in his classes.

Appendix D

Study 1 Demographics Questionnaire

1. How financially "well off" was your family when you were growing up?

- ☐ Extremely well off / money was never a concern for my family
- ☐ Well off / money was usually not a concern for my family
- ☐ Fairly well off / money was occasionally a concern for my family
- ☐ Not very well off / money was often a concern for my family
- ☐ Poor / money was constantly a concern for my family
- ☐ Not sure / Not applicable (not included in analyses)

2. Would you say that most of the students at your college/university come from:

- ☐ Low-income families
- ☐ Middle-class families
- ☐ Upper-class families
- ☐ Different income backgrounds (i.e., more or less an equal mix of students who come from low-income, middle-class, and upper-class families)

3. What is your GPA? [text-box]

4. What category does your primary major fall under?

- ☐ Sciences (e.g., Biology, Environmental Sciences, Chemistry, Physics)
- ☐ Humanities (e.g., Art History, Literature)
- ☐ Social Sciences (e.g., Psychology, Sociology, Anthropology)
- ☐ Education
- ☐ Creative Arts (e.g., Art, Music)
- ☐ Other (please specify):

5. Please indicate your gender:

- ☐ Female
- ☐ Male
- ☐ Non-binary

6. Are you a native English speaker?

- ☐ Yes
- ☐ No

7. How many years have you spoken English?

- ☐ Less than 1 year

- ☐ 1 to 3 years
- ☐ 3 to 5 years
- ☐ 5 to 10 years
- ☐ More than 10 years

8. What is your current class rank?

- ☐ Freshman
- ☐ Sophomore
- ☐ Junior
- ☐ Senior

Appendix E

Study 1-3 Debriefing Form

Secret Completion Code: PSRHQP98457

Thank you for your participation in this study!

As we told you initially, the purpose of this study is to better understand the types of labels that are used to categorize some college students. Now that you have finished participating, we would like to tell you more about what we are investigating. The primary goal of the study was to identify the types of deficit-oriented labels (e.g., "at risk") that are used to categorize low-income, first-generation Black and Hispanic students, and examine the potential effects of these labels on students' affect and motivation. In addition to answering questions about various labels, participants were asked to read a scenario about a college student and the student's advisor. There were multiple versions of this scenario. Some participants read a version that used a deficit-oriented label to describe students, while other participants read a version that used a different label.

Additionally, in order to identify the types of labels that are disproportionately used to categorize low-income, first-generation Black and Hispanic students, we are sampling students who come from this background, as well as White, non-first-generation, middle class students. This will enable us to compare the types of labels reported by students from these different backgrounds. Your responses to the pre-screening questionnaire were used to determine which sample you were included in, but all participants complete the same study.

Your participation in this project will help our efforts in understanding the impacts of labels like "at-risk" and "disadvantaged" may impact students' motivation in college. So, that’s a basic description of what the experiment is about. It is very important for other participants to complete the survey without knowing what we are studying. For this reason, please do not talk about this study with other individuals who may participate. Prior expectations may influence the findings unintentionally and thus make our efforts (and yours) potentially less useful and informative.

If you have any additional questions, comments, or concerns, please feel free to email Shenira Perez, the principal investigator, at perezsj@bc.edu. And thank you again for your participation, we truly appreciate it!

Appendix F

Study 2 & 3 Pre-Screening Questionnaire

e) Are you currently a full-time student at a 4-year university or college?
- ☐ Yes
- ☐ No

f) Are you currently eligible for the Federal Pell Grant?
- ☐ Yes
- ☐ No

g) Are you currently eligible for Federal Work Study?
- ☐ Yes
- ☐ No

h) What is the highest level of education completed by your mother or female guardian?
- ☐ Less than high school completed
- ☐ High school diploma or equivalent
- ☐ Some college, vocational, or trade school (including 2-year degrees)
- ☐ Bachelor's degree (e.g., BS, BA, AB)
- ☐ Master's degree (e.g., MS, MA, MBA)
- ☐ Professional degree (e.g., JD, LLB, MD, DDS, DVM)
- ☐ Doctorate (e.g., PhD, DSc, EdD)

f) What is the highest level of education completed by your father or male guardian?
- ☐ Less than high school completed
- ☐ High school diploma or equivalent
- ☐ Some college, vocational, or trade school (including 2-year degrees)
- ☐ Bachelor's degree (e.g., BS, BA, AB)
- ☐ Master's degree (e.g., MS, MA, MBA)
- ☐ Professional degree (e.g., JD, LLB, MD, DDS, DVM)
- ☐ Doctorate (e.g., PhD, DSc, EdD)

g) Are you Hispanic or Latino (i.e., a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race)?

h) Please indicate your racial background (select all that apply):
- ☐ American Indian or Alaska Native: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.

- ☐ Asian: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
- ☐ Black or African American: A person having origins in any of the black racial groups of Africa.
- ☐ Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
- ☐ White: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

l) Please indicate your age: (dropdown list; *"under 18" – "over 80"*):

m) Please indicate your sex:
- ☐ Male
- ☐ Female

n) Is your college or university a "public" or "private" institution?
- ☐ Yes
- ☐ No

o) Is your college or university religiously affiliated?
- ☐ Yes
- ☐ No

Appendix G

Study 2 Survey

Section 1

Frequency Item

1. How often (*if ever*) have you experienced being categorized as an *at-risk [underprepared; disadvantaged]* student?

- ☐ 1= "Not at all this past academic year"
- ☐ 2= "At least once this past academic year"
- ☐ 3= "A couple of times this past academic year"
- ☐ 4= "About 3-4 times this past academic year"
- ☐ 5= "More than 5 times this past academic year"

Affective/Motivational Consequence Items

Instructions: Please take a moment to think about the way you felt after your experience(s) of being labeled as *at-risk,* and/or *underprepared,* and/or *disadvantaged*, and indicate your level of agreement with the following statements, from 1= "Strongly Disagree" to 6= "Strongly Agree".

Being labeled as *at-risk*, and/or *underprepared*, and/or *disadvantaged* made me feel…

Sense of Belonging Items

1. …like I'm not a valued member of my university's community.
2. …like I don't belong at my university.
3. …like my university supports me and wants me to succeed.

Academic Self-Perception Items

4. …less confident in my ability to do well in college
5. …like I need more help than other students at my university to pass my classes.
6. …like I'm not as smart as most of the other students at my university.
7. …like I'm not 'college material'.

Engagement Items

8. …motivated to work harder in my classes.
9. …hesitant to take any challenging courses moving forward.

Affective Items

10. …discouraged about my future in college.
11. …embarrassed and/or ashamed.

Section 2

Prompt

"In this section of the study, we would like you to think about the *most recent* experience in which someone at your university/college (e.g., an instructor, advisor, faculty) used a *negative label* to characterize your academic potential as a college student, or the general academic potential of students who come from similar *racial, ethnic,* or *income* backgrounds as you.

Some examples of the types of labels we are referring to are: *at-risk*; *underprepared*; and *disadvantaged*. Please note: we are **NOT** interested in adjectives commonly used to describe students based on personal attributes—such as "motivated", "lazy", or "dumb".

Once you recall your most recent experience of being labeled, use the text box below to describe this experience.

Please try to recall as many details about this experience as possible, including:
a) When this experience occurred (e.g., last week, last semester, your freshman year)
b) The specific label that was used
c) The context in which this experience occurred (e.g., in class, an orientation)
d) The individual that used the label to categorize you (e.g., your advisor, an instructor)
The way you felt afterwards (e.g., encouraged, sad, motivated, frustrated)

Please note that after 5 minutes, you will automatically be advanced to the next screen

Appendix H

Study 3 Survey

Section 1

***Deficit-oriented label* condition: Scenario 1**

"Aaron [April] is a college freshman, and is the first person in his [her] family to attend college. Today, Aaron [April] met with his [her] advisor to discuss which classes he [she] should register for in the spring.

As they were wrapping up the advising session, Aaron's [April's] advisor gave him [her] the feedback below.

"I've received some feedback from a few of your professors, who said that you've scored below average on some of your exams and assignments. You know, a lot of my *at-risk students* get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."

***Deficit-oriented label* condition: Scenario 2**

"Ryan [Casey] is in his [her] first semester in college, and is currently finishing a month-long academic support program for first-generation college freshmen. During the last week of the program, Ryan [Casey] met with his [her] advisor to discuss his [her] progress.

As they were wrapping up the session, Ryan's [Casey's] advisor gave him [her] the feedback below.

"I know the past few weeks have been very challenging for you and that you've been struggling in some of your courses. But, I want you to know that many of my *underprepared students* experience similar obstacles in adapting to the expectations of college-level coursework, so we have a lot of resources available for students like you."

***Neutral label* condition: Scenario 1**

"Aaron [April] is a college freshman, and is the first person in his [her] family to attend college. Today, Aaron [April] met with his [her] advisor to discuss which classes he [she] should register for in the spring.

As they were wrapping up the advising session, Aaron's [April's] advisor gave him [her] the feedback below.

"I've received some feedback from a few of your professors, who said that you've scored below average on some of your exams and assignments. You know, a lot of my *first-generation students* get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."

***Neutral label* condition: Scenario 2**

"Ryan [Casey] is in his [her] first semester in college, and is currently finishing a month-long academic support program for first-generation college freshmen. During the last week of the program, Ryan [Casey] met with his [her] advisor to discuss his [her] progress.

As they were wrapping up the session, Ryan's [Casey's] advisor gave him [her] the feedback below.

"I know the past few weeks have been very challenging for you and that you've been struggling in some of your courses. But, I want you to know that many of my *first-generation students* experience similar obstacles in adapting to the expectations of college-level coursework, so we have a lot of resources available for students like you."

***Deficit-oriented label + context* condition: Sample Scenario 1**

"Aaron [April] is a college freshman, and is the first person in his [her] family to attend college. Today, Aaron [April] met with his [her] advisor to discuss which classes he [she] should register for in the spring.

As they were wrapping up the advising session, Aaron's [April's] advisor gave him [her] the feedback below.

"I've received some feedback from a few of your professors, who said that you've scored below average on some of your exams and assignments. You know, a lot of my *at-risk students* get overwhelmed with coursework their first semester, because they haven't had the same level of access to the types of opportunities that have helped their peers prepare for college. Luckily, we have a lot of resources and academic support services available for students who need help."

***Deficit-oriented label + context* condition: Sample Scenario 2**

"Ryan [Casey] is in his [her] first semester in college, and is currently finishing a month-long academic support program for first-generation college freshmen. During the last week of the program, Ryan [Casey] met with his [her] advisor to discuss his [her] progress.

As they were wrapping up the session, Ryan's [Casey's] advisor gave him [her] the feedback below.

"I know the past few weeks have been very challenging for you and that you've been struggling in some of your courses. But, I want you to know that many of my *underprepared students* experience similar obstacles in adapting to the expectations of college-level coursework because they haven't had access to the types of opportunities that have helped their peers prepare for college. So, we have a lot of resources available for students like you."

Valence & Intentionality Items[30]

Instructions: Now, please imagine that you are Aaron [April] and respond to the following questions about this scenario from his [her] perspective.

> *"I've received some feedback from a few of your professors, who said that you've scored below average on some of your exams and assignments. You know, a lot of my at-risk students get overwhelmed with coursework their first semester. Luckily, we have a lot of resources and academic support services available for students who need help."*

1. To what extent to do you believe that the feedback from Aaron's [April's] advisor (*see statement above*) is communicating positive or negative beliefs about Aaron [April]? [1 = "Very Negative Beliefs" – 5 = "Very Positive Beliefs"]

2. To what extent do you believe that Aaron's [April's] advisor is (*see statement above*) intentionally communicating their positive or negative beliefs about Aaron [April]? [1 = "Not at All Intentionally" – 5 = "Very Intentionally"]

3. Please indicate your level of agreement or disagreement with the following statements. [1 = "Strongly Disagree" – 6 = "Strongly Agree"]

Aaron's [April's] advisor…
a) …didn't think much about how he [she] would word his [her] feedback to Aaron [April].
b) …wouldn't deliberately say something to make Aaron [April] feel bad.
c) …is intentionally trying to make Aaron [April] feel better.

4. Next, please indicate your level of agreement or disagreement with the following statements. [1 = "Strongly Disagree" – 6 = "Strongly Agree"]

The advisor's feedback to Aaron [April] probably made Aaron [April] feel:

Sense of Belonging Items
j) …like he [she] is not a valued member of my university's community.
k) …like he [she] doesn't belong at his [her] university.
l) …like his [her] university supports him [her] and wants him [her] to succeed.

Academic Self-Perception Items

[30] Participants will complete these items for each of the (3) scenarios they are presented with. Items will be tailored to the context of each specific scenario (these have been tailored for sample scenario 1).

m) …less confident in his [her] ability to do well in college
n) …like he [she] need more help than other students at his [her] university to pass his [her] classes.
o) …like he [she] is not as smart as most of the other students at his [her] university.
p) …like he [she] is not 'college material'.

Engagement Items
1. …motivated to work harder in his [her] classes.
2. …hesitant to take any challenging courses moving forward.

Affective Items
3. …discouraged about his [her] future in college.
4. …embarrassed and/or ashamed.

Section 2

Labeling Experience & Verbal Tasks

Deficit-Oriented Prompt: In this section of the study, we would like you to think about an experience in which someone at your university/college (e.g., an instructor, advisor, faculty) labeled you as an *at-risk student*.

Once you recall your most recent experience of being labeled, use the text box below to briefly describe this experience.

If you cannot think of a specific instance in which you were categorized as an *at-risk student*, then take a moment to imagine you were meeting with your academic advisor and they used this label to categorize you, and describe this hypothetical scenario in the box below.
Some example of details you may want to include in your description are:

a) When this experience occurred (e.g., last week, last semester, your freshman year)
b) The specific label that was used
c) The context in which this experience occurred (e.g., in class, an orientation)
d) The individual that used the label to categorize you (e.g., your advisor, an instructor)
e) The way you felt afterwards (e.g., encouraged, sad, motivated, frustrated)

Please note that after 3 minutes, you will automatically be advanced to the next screen

Neutral Prompt: In this section of the study, we would like you to think about an experience in which someone at your university/college (e.g., an instructor, advisor, faculty) labeled you as a *first-year student*.

Once you recall your most recent experience of being labeled, use the text box below to briefly describe this experience.

If you cannot think of a specific instance in which you were categorized as a *first-year student*, then take a moment to imagine you were meeting with your academic advisor and they used this label to categorize you, and describe this hypothetical scenario in the box below.

Some example of details you may want to include in your description are:

f) When this experience occurred (e.g., last week, last semester, your freshman year)
g) The specific label that was used
h) The context in which this experience occurred (e.g., in class, an orientation)
i) The individual that used the label to categorize you (e.g., your advisor, an instructor)
j) The way you felt afterwards (e.g., encouraged, sad, motivated, frustrated)

Please note that after 3 minutes, you will automatically be advanced to the next screen

Section 2

Sample Verbal Test

Directions: Each question below contains a pair of words in capital letters and five answer choices. Each answer choice contains a pair of words. Please select the pair that **best** expresses the relationship expressed by the pair in all capital letters.

1. **PRIZE : CONTESTANT:**
   A. trophy : presenter
   B. diploma : principal
   C. medal : runner
   D. book : author
   E. mortgage : lender

2. **CLASSROOM : STUDENTS:**
   A. podium : lecturers
   B. stadium : athletes
   C. cafeteria : trays
   D. garage : vehicles
   E. auditorium : ushers

3. **ENDORSE : CANDIDATE:**
   A. sign : affidavit
   B. endure : trial
   C. idolize : celebrity
   D. espouse : idea
   E. devise : plan

4. **STUDY : TEST:**
   A. script : composition
   B. rehearse : performance
   C. interpret : decision
   D. operate : cure
   E. record : parody

5. **CHRONICLE : JOURNEY:**
   A. assume : debt
   B. enumerate : demands
   C. banish : doubts
   D. juxtapose : positions
   E. clarify : intentions

6. **ANNOTATE : ESSAY:**
   A. elevate : level
   B. research : theory
   C. abridge : chapter
   D. elaborate : plan
   E. mitigate : damage

7. **CAPRICIOUS : IMPULSIVE:**
   A. magnanimous : generous
   B. articulate : critical
   C. petty : deceptive
   D. diligent : precise
   E. provocative : appealing

8. **NOTES : SONG:**
   A. conductors : orchestra
   B. pictures : frame
   C. keys : door
   D. lawyers : courtroom
   E. ingredients : recipe

9. **MARATHON : RACE:**
   A. victory : competition
   B. sprint : finish
   C. filibuster : speech
   D. novel : author
   E. deposition : question

10. **CASTLE : MOAT:**
    A. island : ocean
    B. king : soldier

C. school : playground
D. embryo : placenta
E. bacteria : germ

11. BLIZZARD : SNOW:
A. harvest : garden
B. flood : lake
C. water : ice
D. exhibits : zoo
E. deluge : rain

12. APATHETIC : EMOTION:
A. eloquent : precision
B. lenient : permanence
C. perceptive : awareness
D. zealous : passion
E. glib : sincerity

13. EXULTANT : KUDOS:
A. focused : support
B. joyful : praise
C. honorable : criticism
D. enigmatic : puzzles
E. exceptional : qualities

14. NOXIOUS : POISON:
A. egregious : crime
B. benign : leader
C. dubious : concoction
D. judicious : statement
E. pragmatic : decision

15. UTILITARIAN : QUIXOTIC:
A. disconcerting : unsettling
B. ephemeral : fleeting
C. malevolent : kind
D. loquacious : talkative
E. obdurate : stubborn

Appendix I

Study 2 & 3 Motivational Measures

Stereotype Vulnerability Scale (SVS)

1. Professors/Instructors expect me to do poorly in class because of my race [ethnic background].

2. My academic success may have been easier because of my race [ethnic background]. *(reverse scored)*

3. I doubt that others would think I have less academic success because of my race [ethnic background]. *(reverse scored)*

4. Some people feel I have less academic success because of my race [ethnic background]. *(reverse scored)*

5. People of my race [ethnic background] rarely face unfair evaluations in academic classes. *(reverse scored)*

6. In the academic setting, people of my race [ethnic background] often face biased evaluations from others.
My race [ethnic background] does not affect people's perception of my academic achievement. *(reverse scored)*

7. When I am in academic settings, I often feel that others look down on me because of my race [ethnic background].

Multi-Dimensional Black Identity Scale *(revised)*

1. Overall, being Black [White; Hispanic/Latino; Black Hispanic] has very little to do with how I feel about myself as a college student. *(reverse scored)*

2. In general, being Black [White; Hispanic/Latino; Black Hispanic] is an important part of my self-image as a college student.

3. Being Black [White; Hispanic/Latino; Black Hispanic] is unimportant to my sense of what kind of college student I am. *(reverse scored)*

4. I have a strong sense of belonging when I am around Black [White; Hispanic/Latino; Black Hispanic] college students.

5. I have a strong attachment to other Black [White; Hispanic/Latino; Black Hispanic] college students.

6. Being Black [White; Hispanic/Latino; Black Hispanic] is an important reflection of who I am as a college student.

7. Being Black [White; Hispanic/Latino; Black Hispanic] is not a major factor in my social relationships in college. *(reverse scored)*

Academic Mindsets

1. You have a certain amount of intelligence, and you really can't do much to change it.
2. Your intelligence is something about you that you can't change very much.
3. No matter who you are, you can significantly change your intelligence level.
4. To be honest, you can't really change how intelligent you are.
5. You can always substantially change how intelligent you are.
6. You can learn new things, but you can't really change your basic intelligence.
7. No matter how much intelligence you have, you can always change it quite a bit.
8. You can change even your basic intelligence level considerably.

Appendix J

Study 2 & 3 Demographics Questionnaire

1. How financially "well off" was your family when you were growing up?

- ☐ Extremely well off / money was never a concern for my family
- ☐ Well off / money was usually not a concern for my family
- ☐ Fairly well off / money was occasionally a concern for my family
- ☐ Not very well off / money was often a concern for my family
- ☐ Poor / money was constantly a concern for my family
- ☐ Not sure / Not applicable (not included in analyses)

2. Would you say that most of the students at your college/university come from:

- ☐ Low-income families
- ☐ Middle-class families
- ☐ Upper-class families
- ☐ Different income backgrounds (i.e., more or less an equal mix of students who come from low-income, middle-class, and upper-class families)

3. What is your GPA? [text-box]

4. What category does your primary major fall under?

- ☐ Sciences (e.g., Biology, Environmental Sciences, Chemistry, Physics)
- ☐ Humanities (e.g., Art History, Literature)
- ☐ Social Sciences (e.g., Psychology, Sociology, Anthropology)
- ☐ Education
- ☐ Creative Arts (e.g., Art, Music)
- ☐ Other (please specify):

5. Please indicate your gender:

- ☐ Female
- ☐ Male
- ☐ Non-binary

6. Are you a native English speaker?

- ☐ Yes
- ☐ No

7. How many years have you spoken English?

- ☐ Less than 1 year

- ☐ 1 to 3 years
- ☐ 3 to 5 years
- ☐ 5 to 10 years
- ☐ More than 10 years

8. What is your current class rank?

- ☐ Freshman
- ☐ Sophomore
- ☐ Junior
- ☐ Senior

9. Please indicate your level of agreement or disagreement with the following statement on scale from 1 = "Strongly Disagree" to 6 = "Strongly Agree".

"Doing well in college is very important to me."