

BLOOMING WHERE THEY'RE PLANTED: CLOSING COGNITIVE ACHIEVEMENT
GAPS WITH NON-COGNITIVE SKILLS

Elaina Michele Sabatine

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in
partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of
Social Work

Chapel Hill
2019

Approved by:

Melissa Lippold

Amy Blank Wilson

Natasha Bowen

Din-Geng Chen

Jill Hamm

©2019
Elaina Sabatine
ALL RIGHTS RESERVED

ABSTRACT

ELAINA SABATINE: Blooming Where They're Planted: Closing Cognitive Achievement Gaps With Non-Cognitive Skills
(Under the direction of Dr. Melissa Lippold)

For the last several decades, education reform has focused on closing achievement gaps between affluent, white students and their less privileged peers. One promising area for addressing achievement gaps is through promoting students' non-cognitive skills (e.g., self-discipline, persistence). In the area of non-cognitive skills, two interventions – growth mindset and stereotype threat – have been identified as promising strategies for increasing students' academic achievement and closing achievement gaps. This dissertation explores the use of growth mindset and stereotype threat strategies in the classroom, as a form of academic intervention.

Paper 1 examines the extant evidence for growth mindset and stereotype threat interventions. Paper 1 used a systematic review and meta-analysis to identify and analyze 24 randomized controlled trials that tested growth mindset and stereotype threat interventions with middle and high school students over the course of one school year. Results from meta-analysis indicated small, positive effects for each intervention on students' GPAs. Findings highlight the influence of variation among studies and the need for additional research that more formally evaluates differences in study characteristics and the impact of study characteristics on intervention effects.

Paper 2 explored middle school students' perceptions of teacher behaviors related to growth mindset and stereotype threat theories. Paper 2 used qualitative analysis of data from

9 focus groups with 44 middle school students in 3 rural, low income middle schools.

Emergent themes included participants' beliefs that all students are smart, that some teachers feel that certain students are smarter than others, and that students feel smartest when their teachers provide instrumental support and show emotional care. Findings highlight the capacity for middle school students to observe concepts like ability, identity and stereotypes in their teachers' behavior and the resulting impact on their motivation and beliefs about the nature of intelligence.

Paper 3 explored teacher strategies to implement growth mindset and stereotype threat theories in the classroom. Paper 3 used qualitative analysis of classroom observations and individual interviews with 9 middle school teachers in a rural, low income school. Findings include a theme related to teachers' support of students' *productive struggle*, in part by allowing students to resubmit assignments and retake assessments. Findings also indicate the potential for unintended consequences of implementing growth mindset and stereotype threat theories, as teachers reported that "retakes" also had the potential to demotivate students.

**Per Frances Sabatino e Abramo Manzi, chi hanno capito il valore delle donne che
hanno ricevuto un'istruzione,**

E per Filomena Damico e Gloria Pennesi, che non hanno avuto la possibilita' -

Ci avete dato quello che i soldi non possono comprare.

ACKNOWLEDGMENTS

I'd like to begin by thanking my chair, Melissa Lippold, for her guidance and support throughout my doctoral program and, especially, during the dissertation process. I would also like to extend my gratitude to my committee members: to Amy Blank Wilson, for going above and beyond in providing her expertise and knowledge of qualitative methods and analysis; to Din Chen for his encouragement and willingness to discuss in great detail each analytic decision of my meta-analysis; to Jill Hamm, for sharing her wealth of knowledge and experience with education research and school-based intervention; and finally, to Natasha Bowen, who encouraged me to pursue a doctorate in social work and who was instrumental in helping me to view education from a social work lens – without her, I would never have ended up in this program nor would I hold this degree. I'd also like to thank Kirsten Kainz for sharing her support, humor and perspective throughout my time in the program.

My biggest debt of gratitude is to my parents, Jan and Rich Sabatine. You have shown me the value of working hard, asking meaningful questions, holding different perspectives and operating with intentionality and purpose. I'm equally thankful to my big brother Vince – my original growth mindset mentor – who in helping me learn to ride a bike once told me, “The trick is not to try to go far. The trick is that each time, you just have to go a little farther than the last time.” At eight, you were able to communicate so succinctly, beautifully and effectively the theme of this dissertation and the attitude you have fostered in me since the day I was born. It is a great privilege to grow up surrounded by three brilliant

scientists who refused to let me succumb to the stereotypes imposed on me by my own educational experiences: that a tender heart and creative mind is necessarily less empirical, methodical or capable. My great hope is that all students could be so lucky to grow up surrounded by people who believe so fully in their abilities.

To the Dream Team, I cannot express how thankful I am for the incredible fortune of having arrived in Chapel Hill at the same time as y'all. I can say with 100% certainty that I would not have completed the process without y'all, nor would I want to. Thanks for your humor, brilliance and strength. To the Less Fam, much of this dissertation is the direct result of our friendship – from the chats in our classrooms at the end of long days or in the Belly of the Whale on long road trips. Thanks for always being willing to deconstruct worldviews, to view things from different angles and – especially – to bring Thomas to the long hours and tough days. Y'all are a gift.

To the students of ENC STEM, I hope y'all never doubt your capacity for growth and the power embedded in the core of who you are. Y'all never cease to surprise me with your brilliance, energy and love. Thanks for continually investing in our family. You are changing the narrative of what's possible for students in Eastern North Carolina.

Most importantly, I want to thank my partner Seth for, among so many things, your optimism and wholeheartedness and the sense of ease with which you move through the world. You bring balance and humor to my most stressful moments, and I aspire to your capacity for calm. And, finally, I can't forget Pretty Rey, who laid patiently and supportively at my feet while every single page of this dissertation was written. You are the sweetest jewel.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES.....	xi
INTRODUCTION.....	1
References: Introduction.....	5
PAPER I: STEREOTYPE THREAT AND GROWTH MINDSET: A META-ANALYSIS OF INTELLIGENCE BELIEFS AS INTERVENTION TARGETS FOR IMPROVING ACADEMIC OUTCOMES.....	7
Introduction.....	7
Methods: Systematic Review	14
Results: Systematic Review	16
Methods: Meta-Analysis.....	21
Results: Meta-Analysis	34
Discussion	36
Conclusion.....	47
References: Paper I.....	48
Tables: Paper I.....	57
Figures: Paper I	73
PAPER II: MINDSETS AND STEREOTYPES OF INTELLIGENCE: TEACHER BEHAVIORS THAT MOTIVATE STUDENTS TO ACHIEVE.....	80
Introduction.....	80
Methods.....	90

Results	101
Discussion	113
Conclusion.....	126
References: Paper II.....	127
Tables: Paper II.....	136
Appendices: Paper II.....	138
PAPER III: TRANSLATING GROWTH MINDSET AND STEREOTYPE THREAT THEORIES TO THE CLASSROOM: A QUALITATIVE ANALYSIS.....	139
Introduction.....	139
Methods.....	148
Results	157
Discussion	165
Conclusion.....	173
References: Paper III	174
Tables: Paper III	182
Appendices: Paper III	183
SUMMARY	185
References: Summary.....	192

LIST OF TABLES

Table 1.1: List of excluded studies	57
Table 1.2: Description of stereotype threat studies	58
Table 1.3: Description of growth mindset studies	62
Table 1.4: Effects of stereotype intervention	64
Table 1.5: Effects of growth mindset intervention	66
Table 1.6: Meta-analysis of stereotype threat intervention on GPA of threatened students ...	68
Table 1.7: Meta-analysis of growth mindset intervention on GPA of all students	69
Table 1.8: One study removed sensitivity test for stereotype threat meta-analysis	70
Table 1.9: One study removed sensitivity test for growth mindset meta-analysis	71
Table 1.9: One study removed sensitivity test for growth mindset meta-analysis, with Brougham (2016) removed	72
Table 2.1: Demographics of focus groups.....	136
Table 2.2: Summary of themes	137
Table 3.1: Final summary of themes.....	182

LIST OF FIGURES

Figure 1.1: Identification of eligible studies.	73
Figure 1.2: Risk of bias assessment across stereotype threat studies.....	74
Figure 1.3: Risk of bias assessment across growth mindset studies... ..	75
Figure 1.4: Summary of outcomes, populations and samples for stereotype threat studies.....	76
Figure 1.5: Summary of outcomes, populations and samples for growth mindset studies... ..	77
Figure 1.6: Forest plot for meta-analysis of stereotype threat intervention on GPA... ..	78
Figure 1.7: Forest plot for meta-analysis of growth mindset intervention on GPA.....	79

INTRODUCTION

BLOOMING WHERE THEY'RE PLANTED: CLOSING COGNITIVE ACHIEVEMENT GAPS WITH NON-COGNITIVE SKILLS

For decades, education reform has focused on identifying and addressing factors that contribute to gaps in academic performance between affluent, white students and their less privileged peers – often, students of color and students from low income families (Kena et al., 2016). Many reform efforts have focused on meeting students' social and emotional needs and the resulting impact on their school performance (Brown, 2013). To describe the importance of supporting students' social and emotional development, Heckman and Rubenstein (2001) popularized the term *non-cognitive skills* to refer to factors like self-discipline and perseverance, which they observed as key characteristics separating students who dropped out of high school and those who persisted through graduation. Subsequent research has examined the schools' role in supporting students' acquisition of non-cognitive skills as a strategy for improving students' academic performance.

Growth Mindset and Stereotype Threat as Interventions

Specifically, several studies have looked at students' beliefs and experiences with school, examining students' perceptions of ability and sense of belonging as critical factors in promoting students' development of non-cognitive skills and in predicting students' academic success. Dweck's mindset research (1999; 2006) has investigated the impact of students' beliefs about their ability on their effort, engagement and performance in school. Students who have a growth mindset believe that intelligence grows with effort, and

encouraging students to hold a growth mindset of ability has been associated with gains in academic performance, particularly among low income students and students of color (e.g., Blackwell, Trzesniewski, & Dweck, 2007; Good, Aronson, & Inzlicht, 2002; Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015; Yeager et al. 2014; Yeager et al., 2016). In addition, Cohen and colleagues (2006) have extended Steele and Aronson's (1995) work on the phenomenon of stereotype threat to explore the interplay between students' beliefs about their identities, their sense of belonging in school and how well they perform in class.

Negative stereotypes that link specific identities (e.g., gender, race/ethnicity) to low ability can create threat that suppresses academic achievement, but interventions to reduce threat via identity affirmation have been linked with long-term improvements in students' achievement (e.g., Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Sherman et al., 2013).

Continued replication of growth mindset and stereotype threat studies has contributed to a growing evidence base for boosting students' achievement by shaping their intelligence beliefs, but little work has explored how best to adapt growth mindset and stereotype threat strategies for daily school practice, particularly for classroom teachers (Edwards, Esmonde, Wagner, & Beattie, 2017). This dissertation aims to fill these gaps by aggregating the evidence on growth mindset and stereotype threat interventions; by investigating students' experiences and how they attend to their teachers' beliefs and behaviors related to ability; and by exploring strategies for supporting growth mindset and reducing stereotype threat that have been developed and implemented by classroom teachers. Because the research on growth mindset and stereotype threat interventions as evidence-based practice is relatively new, this dissertation focuses on understanding more about how teachers and students respond to intervention strategies organically, within the school environment, to support future, formal

evaluation efforts where teachers implement growth mindset and stereotype threat intervention strategies.

Dissertation Papers

Paper 1, “Stereotype Threat and Growth Mindset: A Meta-Analysis of Intelligence Beliefs as Intervention Targets for Improving Academic Achievement,” consolidates the research that has been conducted on growth mindset and stereotype threat interventions and their effects on student achievement (e.g., grades, GPA, test scores). Paper 1 assesses intervention effects first by conducting a systematic literature review to consolidate the randomized controlled trials that have tested each intervention and second by conducting meta-analyses that separately summarize each intervention’s effects. Paper 1 also includes a discussion of factors that may affect intervention success, contributing to the knowledge base around how and for whom these interventions should be implemented.

Paper 2, “Mindsets and Stereotypes of Intelligence: Teacher Behaviors that Motivate Students to Achieve” focuses on students’ perceptions of how their teachers message the nature of intelligence (whether it’s fixed or malleable) and intelligence-based stereotypes (whether certain students are smarter than others). Paper 2 uses qualitative data from focus group interviews with adolescent students in low income, rural middle schools to investigate the experiences of high need students. Through a greater understanding of students’ opinions, beliefs and experiences, Paper 2 serves as a basis for future classroom research on growth mindset and stereotype threat strategies. For example, the findings of Paper 2 could be used to generate new strategies or refine existing strategies that could be evaluated as pathways to integrate growth mindset and stereotype threat theories into the classroom. Findings from Paper 2 could also inform the development of measures that quantify student ratings of

teachers' behavior. Such measures would be a helpful resource in future intervention evaluations, including structural models that can provide more detail on the mechanisms among teacher behaviors, students behaviors and student achievement.

Paper 3, "Translating Growth Mindset and Stereotype Threat Theories to the Classroom: A Qualitative Analysis," provides a basis for understanding how teachers organically implement growth mindset and stereotype threat strategies in their classrooms when they have been exposed to information and resources about the two theories. Paper 3 uses qualitative data from classroom observations and interviews with middle school teachers in a high need school to explore ways that teachers integrate these theories into their daily practice. In addition, teachers discussed the impact of using these strategies – including both benefits and drawbacks – which provides helpful context for future implementation of these strategies in schools.

Together, all three papers contribute to the evidence base for reducing gaps in student achievement. This dissertation evaluates strategies that have been identified as evidence-based practices for implementing growth mindset and stereotype threat theories in schools and draws upon the perspectives of both teachers and students in identifying how these strategies may be used to promote student achievement.

REFERENCES: INTRODUCTION

- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Brown, E. (2013). No child left behind and the teaching of character education. *ABNF Journal, 24*(3), 77-82.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*, 1307–1310. doi:10.1126/science.1128317
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*(5925), 400-403. doi:10.1126/science.1170769
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Edwards, A. R., Esmonde, I., Wagner, J. F., & Beattie, R. L. (2017). Learning mathematics. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 57–80). New York, NY: Routledge.
- 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*(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149. doi:10.1257/aer.91.2.145
- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., ... & Barmer, A. (2016). The Condition of Education 2016. NCES 2016-144. *National Center for Education Statistics*.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017

- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., . . . Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology, 104*(4), 591-618. doi:10.1037/a0031495
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797–811. doi:10.1037/0022-3514.69.5.797
- 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*(3), 374. doi:10.1037/edu0000098
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General, 143*(2), 804. doi:10.1037/a0033906

PAPER I

STEREOTYPE THREAT AND GROWTH MINDSET: A META-ANALYSIS OF INTELLIGENCE BELIEFS AS INTERVENTION TARGETS FOR IMPROVING ACADEMIC OUTCOMES

Despite several decades of curricular reform, increased standardized testing and heightened accountability, academic disparities persist between low-income and racial/ethnic minority youth and their more affluent, white peers (Dee & Jacob, 2011; Kena et al., 2016; Vanneman, Hamilton, Anderson, & Rahman, 2009). In light of evidence that cognitive intervention (e.g., curricular reform, heightened accountability and high-stakes testing) has had little impact on closing performance gaps, school-based intervention has shifted toward strategies that promote achievement by addressing social and emotional determinants of learning (Brown, 2013). One important social and emotional determinant of learning is youths' attitudes and perceptions towards school (Durlak, Weissberg, Dymnicki, Taylor, Schellinger, 2011). In particular, strategies that promote students' positive beliefs about their intelligence – identified here as growth mindset and stereotype threat interventions – have been linked to students' academic achievement (Blackwell, Trzesniewski, & Dweck, 2007; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Prior research has demonstrated that, as students internalize beliefs that their performance can increase with effort (growth mindset) and is not predetermined by their group identity (stereotype threat), they increase their engagement and persistence in school, subsequently boosting academic achievement (Blackwell, et al., 2007; Cohen, Garcia, Apfel, & Master, 2006; Cohen et al. 2009; Dinger Dickhäuser, Spinath, & Steinmayr, 2013).

Growth Mindset

Several decades of mindset research reveal that students' beliefs about their intelligence are powerful predictors of behavior and achievement (Gonida, Kiosseoglu, & Leondari, 2006; Kloosterman, 1988; Stipek & Gralinski, 1996). Based on seminal work in social psychology by Dweck (1999; 2006) and colleagues, the concept of a growth mindset was borne out of literature on learned helplessness and implicit intelligence theories.

Whereas an incremental view of intelligence (growth mindset) states that ability can be increased with effortful practice, an entity view (fixed mindset) states that ability is a stable and predetermined quality that cannot be changed. Research and classroom data consistently show that students' mindsets drive their level of classroom engagement (Dinger et al., 2013; Elliot & Church, 1997), which is necessary for academic success (Johnson, Crosnoe, & Elder, 2001; Park, Holloway, Arendtsz, & Bempechat, 2012; Sciarra & Seirup, 2008).

Students with a growth mindset show greater effort, higher school enjoyment and more positive affect than their peers - and they maintain these behaviors even after they are unsuccessful on academic tasks (Cimpian, Arce, Markman, Dweck, 2007; Dweck & Leggett, 1988; Elliot & Dweck, 1988). On practically every measure of adaptive school behaviors, students with a growth mindset outperform their fixed mindset peers: They are more resilient, cooperative, self-regulated and positive, and they are more likely to succeed academically (Senko, Hulleman, & Harackiewicz, 2011).

Fortunately, mindsets are highly malleable. Children learn from parents and teachers whether to perceive ability as stable or fixed (Gunderson, Gripshover, Romero, Dweck, Goldin-Meadow, & Levine, 2013; Henderlong & Lepper, 2002), and differences in parent and teacher praise and messaging about the nature of intelligence influence students' sense of

helplessness, goal choice and strategy use in overcoming failures (Cimpian, et al., 2007; Mueller & Dweck, 1996). Based on this work, various strategies to promote students' growth mindset beliefs have been developed and tested as academic interventions. Aronson, Fried and Good (2002) designed a pen pal experiment where undergraduate students corresponded with fake middle school students who were purportedly struggling in school. End of semester GPAs were higher among undergraduate students who received growth mindset training to assist their pen pal compared to those students who did not receive growth mindset training. Similarly, middle school students who were taught about the growth mindset in two 90-minute sessions by a real college-aged mentor (with whom they corresponded via email throughout the school year) scored higher on end of grade standardized assessments in math than those in a control group (Good, Aronson, & Inzlicht, 2003). Finally, Blackwell et al. (2007) found improvements in course grades among students who participated in a six-week growth mindset course. Subsequent studies have focused on the efficacy of scaling up the intervention developed by Blackwell and colleagues (2007), using web-based modules that have replicated achievement effects in samples as large as 3,000 students (Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015; Yeager et al. 2016).

Stereotype Threat

Stereotype threat was first defined by Steele and Aronson (1995) as the "risk of confirming, as self-characteristic, a negative stereotype about one's group" (p. 797). As it relates to academic performance, stereotype threat can occur in two ways: (1) students who belong to minority groups internalize negative stereotypes about the academic abilities of their group or (2) students' efforts to disconfirm a stereotype about their intelligence create a cognitive burden that reduces learning and task performance (Aronson, 1999; Aronson,

Quinn, & Spencer, 1998; Blascovich, Spencer, Quinn, & Steele, 2001; Mangels, Good, Whiteman, Maniscalco, & Dweck, 2012; Rydell, Shiffrin, Boucher, Van Loo, & Rydell, 2010; Steele & Aronson, 1995). The phenomenon of stereotype threat is well-documented among black and Hispanic students, as well as women in male-dominated subjects like math and science (Steele and Aronson, 1995; Steele & Aronson, 2000). Threat can be reduced by providing messages about the equality of students' abilities, regardless of racial, ethnic or gender identity. For example, in some studies, a simple statement denoting that men and women perform equally well on a difficult math exam before its administration has been sufficient to eliminate gender gaps in performance (Spencer, Steele, & Quinn, 1999). The same is true for racial differences in performance between white and non-white students (Steele & Aronson, 1995).

In school, stereotype threat can manifest in response to a particular task (e.g., standardized tests), but a more chronic form of threat also exists in the day-to-day classroom environment. A powerful source of stereotype threat occurs when students' identities are not positively affirmed by their school experiences, leading them to internalize negative stereotypes about their intelligence. Extending Steele and Aronson's work to the classroom setting, Cohen et al. (2006) identified that negative school experiences can compound students' beliefs about intelligence-based stereotypes in reducing academic achievement. In response, they conceptualized an identity-affirming intervention to combat stereotype threat in schools using brief, written reflections on personal values. Unlike growth mindset interventions, the value affirmation does not explicitly address the nature of intelligence, intelligence-based stereotypes or commonly stereotyped identities (i.e., gender and race). Instead, to reduce stereotype threat, Cohen and colleagues (2006; 2009) implemented the

value reflections prior to potentially stressful school events (e.g., before the end of the grading period, before standardized exams), when threat is likely to be high. They hypothesized that having students engage in a positive reflection about themselves and their values just prior to a potentially threatening school event may encourage students to make positive attributions between their identities and their school performance, thereby reducing threat posed by stereotypes. Cohen and colleagues (2006; 2009) suggest that this process of identity affirmation may provide the same initial threat reducing benefits conceptualized by Steele and Aronson (1995), but may also trigger a “cascade” of positive school experiences that produce long term effects on students’ academic achievement. Indeed, in some studies, the values reflection intervention has been associated with a reduction in achievement gaps and improvements in students’ grade point averages (GPAs) for up to three years (Cohen et al., 2009; Sherman et al., 2013).

Distinction between Stereotype Threat and Growth Mindset Approaches

Although intelligence-based stereotypes may represent a specific type of fixed mindset beliefs (e.g., I can’t get better at math *because* I am a girl), the two interventions are conceptually distinct (Good et al., 2003), and the interventions often target different populations. Teaching students about the malleability of intelligence — growth mindset interventions — is intended to change their attitudes about intelligence, while the values reflection is intended to inoculate students against stereotypes about their abilities. Thus, the interventions target different messages and social determinants of learning. There is some evidence to suggest that growth mindset beliefs may be inherently stereotype threat-reducing (Good et al., 2003; Paunesku et al., 2015; Yeager et al., 2014 – Study 2), suggesting the interventions may be related. However, growth mindset and stereotype threat interventions

have often been targeted toward and tested on different populations, making it difficult to compare them.

In general, the values affirmation task has primarily been tested on and shown effects mostly among stereotyped groups (e.g., racial minority groups). In contrast, growth mindset interventions have been tested on general student populations and have shown effects among a broad range of students. Thus, it would be inappropriate to aggregate and analyze combined data from both interventions, as there is little reason to believe that the interventions are directly interchangeable (Blackwell et al., 2007; Cohen et al., 2006; Cohen et al., 2009). Accordingly, in this meta-analysis, growth mindset and stereotype threat interventions will be evaluated separately for their effects.

Adolescence and the School Context

Growth mindset and stereotype threat interventions have shown promise with undergraduate students (e.g., Aronson, Fried, & Good, 2002), but much of the recent research has been conducted with adolescents (i.e., middle and high school students; Blackwell et al., 2007; Cohen et al., 2009; Paunesku et al., 2015; Yeager et al., 2014; Yeager et al. 2016). Broadly, adolescence is an ideal time for academic intervention for several reasons: The middle school transition marks a time when school becomes more evaluative and competitive, and students are more likely to adopt fixed mindset beliefs that match the increased emphasis on grades and test scores (Eccles & Midgley, 1989; Midgley, Feldlaufer, & Eccles, 1995). As adolescents matriculate through middle school and into high school, they become increasingly self-aware, navigating a critical process of identity development while making frequent judgments about their academic abilities (Eccles, Midgley, Wigfield, Buchanan, Reuman, Glanafan, & Mac Iver, 1993). In the absence of intervention, students

can form intelligence and identity beliefs that negatively affect their school performance – a process that may help explain a characteristic pattern of achievement that drops as students enter middle school (Blackwell et al., 2007). Though on average, achievement begins to rise as students progress through high school, for some, disengagement in early adolescence cements a trajectory of poor performance that affects the rest of their academic careers (Eccles, Lord, & Midgley, 1991; Gutman & Midgley, 2000; Midgley, Feldlaufer, & Eccles, 1989; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991). Given this possibility, early intervention is critical, and evaluating how growth mindset and stereotype threat interventions operate with middle and high school students is important for identifying effective strategies for promoting academic achievement at a pivotal time in students' lives.

Purpose of Present Study

Based on results from seminal studies (e.g., Aronson, Fried, & Good, 2002; Bowen, Wegmann, & Webber, 2013; Blackwell, et al., 2007; Cohen et al., 2009; Good, et al., 2003; Sherman et al., 2013), the research on growth mindset and stereotype threat interventions has been disseminated within school systems, where there is an increasing emphasis on evidence-based practice (e.g., included in the What Works Clearinghouse through the Institute of Education Sciences; Aronson, Cohen, McColskey, Montrosse, Lewis, & Mooney, 2009). Yet, little research has synthesized the overall effectiveness of these interventions. Without a comprehensive analysis that assesses the consistency and magnitude of intervention effects, it is unclear how successfully strategies to increase growth mindset and reduce stereotype threat buffer against academic underachievement. To fill this gap, our study synthesizes the available data from randomized controlled trials of growth mindset and stereotype threat interventions in a systematic review and meta-analysis, in pursuit of the following research

questions:

1. In middle and high school students, how does growth mindset intervention compared to control activities affect students' GPA within one academic year?
2. In middle and high school students of threatened identities (e.g., black and Hispanic students), how does stereotype threat intervention compared to control activities affect students' GPA within one academic year?

Methods: Systematic Review

Study Selection and Data Abstraction

Eligibility criteria. This study adhered to the PRISMA guidelines for systematic review and meta-analysis (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). The review includes only randomized interventions designed to increase academic achievement either by reducing stereotype threat or by increasing students' growth mindset beliefs.

Studies must have used at least one experimental group and one control group, with random assignment to each condition. Interventions were considered stereotype threat-reducing if they met one or more of the following criteria: (1) affirming students' identities within the school context; (2) promoting students' sense of belonging in school; or (3) promoting the idea that anxiety about school is common among all racial, ethnic and gender identities.

Growth mindset interventions were categorized as (1) teaching students that intelligence is malleable and/or that they can "grow their brains"; (2) promoting effortful practice as a strategy for academic improvement; or (3) encouraging students to perceive challenge as a sign of learning.

Studies were included only if the interventions were conducted with middle or high school students (Grades 5-12). Interventions must have taken place in school, with a primary

aim to improve long-term academic performance. The outcomes of interest for both interventions are indicators of academic achievement, specifically: course grades, standardized exam scores and/or grade point average (GPA). Studies were included if they measured at least one of these outcomes. Randomized control trials (RCTs) that manipulate stereotype threat or growth mindset beliefs to assess immediate outcomes (e.g., test performance directly following cognitive primes) were excluded.

To avoid differences in the structure of education systems and the cultural norms surrounding stereotypes and intelligence, studies were excluded if they were conducted outside of the United States or were not published in English. Studies that were published before 1980 were excluded from this review because the conceptual frameworks for growth mindset and stereotype threat, as they are defined here, were not identified prior to this date. Conference proceedings, dissertations and other unpublished work were eligible for inclusion to reduce the risk of publication bias.

Information sources. Seven electronic databases were used to identify articles for review: PubMed, Web of Science, PsychINFO, ERIC, Education Full Text, Social Work Abstracts and the publications and products search tool of the Institute of Education Sciences. The following keyword search was used in each database: “*growth mindset*” OR “*fixed mindset*” OR “*incremental intelligence*” OR “*incremental theory*” OR “*incremental beliefs*” OR “*entity beliefs*” OR “*entity theory*” OR “*implicit intelligence theor**” OR “*stereotype threat*” AND *school* OR *learn** OR *academ** OR *grade* OR *educ**. In all EBSCO databases (PsychINFO, ERIC, Education Full Text, and Social Work Abstracts), the advanced search option of “empirical studies” was selected to reduce the high volume of non-relevant work identified by this search string. For the same reason, in Web of Science,

the advanced search options of “psychology social,” “psychology educational,” “education,” “education research,” and “psychology multidisciplinary” were used.

Study selection and data abstraction. In each search, studies were initially reviewed by title and abstract simultaneously. Documents that potentially met inclusion criteria were retrieved from each search, and duplicates were removed before full-text review. See Table 1.1 and Figure 1.1 for more details on study selection. Data abstractions of included studies were conducted with coding spreadsheets that captured both study characteristics and empirical findings, in addition to data regarding effect sizes (e.g., test statistics, standard deviations, standard errors, variance, sample sizes). If studies reported standardized effect sizes, they were noted in the abstraction forms.

Results: Systematic Review

A total of 3,062 articles were identified using this search protocol, including 687 duplicates that were removed before further review. After title and abstract review, an additional 2,332 articles were removed because they failed to meet inclusion criteria. Full text review was conducted on 38 articles, where an additional 17 articles did not meet inclusion criteria. While our search protocol did return four eligible documents from the unpublished literature (e.g., conference proceedings, reports from government agencies), all of these papers summarized data from one or multiple of the peer-reviewed studies that were returned by the search string. These unpublished articles were excluded, in addition to two other studies (Bancroft, Bratter, & Rowley, 2017; Hanselman, Bruch, Gamoran, & Borman, 2014) that presented secondary analyses of a primary intervention study that was already included in the sample. Of the remaining 11 excluded articles, five were not academic interventions or did not measure a long term academic outcome, two did not use random

assignment, one was on observational study, and three used non-American or undergraduate samples. See Figure 1.1 and Table 1.1 for more detail on the selection process. In total, this search resulted in 13 stereotype threat studies published in 12 articles and 11 growth mindset studies published in 12 articles. Sherman et al. (2013) and Eskreis-Winkler et al. (2016) contain evaluations of two separate implementations of stereotype threat and growth mindset interventions, respectively.

Assessing risk of bias. The risk of bias assessment tool provided by the Cochrane Collaboration (Higgins & Green, 2011) identifies five areas of potential bias for systematic reviews/meta-analyses: selection bias, performance bias, detection bias, attrition bias and reporting bias. The issue of selection bias, or systematic variation among groups at the outset of an intervention, is low in this review because it includes only studies with random assignment. Similarly, the issue of performance bias – in this case, students receiving differential treatment apart from their participation in the experimental and control groups – was mitigated in stereotype threat interventions because 12 of 13 studies used a double-blind design where both teachers and students were unaware of students’ treatment condition. One study (Good et al., 2003) did not report on blinding of participants, but did report blinding of implementers. The investigators of growth mindset interventions rarely provided information on treatment concealment, but of the three that did (Good et al., 2003, Yeager et al. 2014; Yeager et al., 2016), it was the implementers who were blinded to students’ condition. No information was available regarding treatment concealment with students, which could increase the risk for bias in these studies. In some studies (e.g., Pauneksu et al., 2015; Yeager et al. 2014; Yeager et al., 2016), assignment was conducted randomly when students logged into a web-based program, but no further detail was provided on treatment concealment. In

growth mindset interventions with more lengthy procedures (e.g., Blackwell et al., 2007; Good et al., 2003), where students were assigned to different in-school courses, students in the control group participated in programs structurally similar to the intervention but with different messaging. For example, control group students experienced the same number and length of sessions as experimental group students, but learned about brain anatomy or study skills and not the malleable nature of intelligence. Again, because these studies did not explicitly mention efforts to keep students unaware of their treatment condition, they were not considered to be blinded, a potential source of bias.

Both interventions have been conducted throughout the school year, ranging from a semester to one or several academic years, suggesting that attrition may be a potential source of bias in these studies. Across the 24 studies from both types of intervention, three studies failed to mention retention of participants (Good et al., 2003, Protzko & Aronson, 2016; Shnabel, Purdie-Vaughns, Cook, Garcia, & Cohen, 2013). Of the remaining studies, all reported that attrition occurred throughout the study. Eight studies reported attrition rates of around 5% or less and/or reported findings to support that attrition was not systematic to treatment group, race/ethnicity, gender or other demographic variables (Blackwell et al., 2007; Borman et al., 2016; Bowen et al., 2013; Burnette, Russell, Hoyt, Orvidas, & Widman, 2017; Cohen et al., 2009; Eskreis-Winkler et al., 2016, Study 4; Hanselman, Rozek, Grigg, Borman, 2017; Sherman et al., 2013, Study 1). Five studies reported high attrition rates, between 20-25% (Bratter, Rowley, Chukhray, 2016; Brougham, 2016; Sherman et al., 2013, Study 2; Simmons, 2013; Wilkins, 2014). Of these, two reported that attrition varied systematically, leaving the analytic sample to be more advantaged than the original sample (Bratter et al., 2016; Sherman et al., 2013, Study 2).

Included studies were also analyzed for reporting bias, which revealed that five studies failed to report findings from all planned analyses (Borman et al., 2016; Cohen et al., 2009; Good et al., 2003; Oyserman, Bybee, & Terry, 2003; Paunesku et al., 2015). However, for two of these studies, supplemental data was available to include results from analyses that were not published (Borman et al., 2016; Cohen et al., 2009), leaving only three studies with incomplete reporting around outcomes.

In general, though both intervention categories are subject to bias due to attrition, the overall risk of bias in this review is likely low, as bias is low across most categories (Higgins et al., 2011). See Figures 1.2 and 1.3 for more detail on the risk of bias in this review.

Characteristics of Included Studies

Intervention design. Among the 13 studies in the stereotype threat category, there was strong consistency in study procedures and outcomes. All but one study (Good et al., 2003) closely replicated the protocol identified by Cohen et al. (2006), where students write about a list of values (e.g. relationships with friends and family, being good at sports). The process of reflecting on personal values, rather than values in general, is believed to strengthen students' sense of identity in the school context, reducing stereotype threat that may negatively affect their academic performance (Cohen et al., 2006). Students in the experimental group were instructed to write about the value most important to them, whereas students in the control group were instructed to write about the value least important to them. All students were given approximately 15 minutes to write a brief paragraph about why their value is important to them (experimental) or why their value may be important to someone else (control). This procedure is typically repeated several times throughout the year, with variations made to the list of values. When possible, to increase internalization of the

affirmation, the reflections were messaged as a regular school activity (rather than a research task) and were conducted prior to a stressful or evaluative event (e.g., the end of the grading period or before a standardized test) to affirm students' identities when threat is likely high. Nine studies (70%) reported effects on student GPA, four (31%) reported effects on course grades and three (23%) reported effects on standardized test scores (some studies reported on more than one outcome). Two studies were dissertations (Simmons, 2012; Wynne, 2011), and the remaining ten studies were published in peer-reviewed journals. See Table 1.2 and Figure 1.4 for a full description of included stereotype threat studies.

Compared to the stereotype threat interventions, among the 11 studies in the growth mindset category, there was more variability in the methods and duration of the intervention, ranging from one 45-minute lesson to 11 weekly hour-long sessions. Some studies exclusively taught students about the malleable nature of intelligence, while others added mentorship, emphasized effortful practice or focused on retraining students' attributions about teachers' critical feedback on assignments. Despite this variability in procedures, all interventions were designed to promote students' effortful engagement in school by modifying their perceptions of intelligence and ability. Eight studies (73%) reported effects on student GPA, two (18%) reported effects on course grades and two (18%) reported effects on standardized test scores (some studies reported effects on more than one outcome). See Table 1.3 and Figure 1.5 for a full description of included growth mindset studies.

Participants. The stereotype threat studies included in this review have a combined sample size of 7,801, which varied widely among studies, with an average of 558 students per study ($SD = 685$). Ten studies (77%) were conducted with middle school students and three studies (23%) with high school students (see Figure 1.4). The focus of these studies was

largely on the academic performance of black, Hispanic and white students, but in some cases, students were categorized simply as having potentially threatened identities (i.e., black or Hispanic) or non-threatened identities (i.e., white or Asian), but were not described further. In eight stereotype threat studies, results were presented on intervention effects for both threatened and nonthreatened identities, but in the remaining five studies, results were reported for students of threatened identities only (see Figure 1.4). Six studies were conducted among multiple schools with varying populations, combining data from students in predominantly black schools, predominantly white schools, low/high performing and high/low income schools, etc. Though multilevel models were used to consider between school differences in these studies, analyses were typically not broken down further by school type. The remaining seven studies that were conducted in one school were largely low income and low performing schools, with at least half of the student population identifying as black or Hispanic. See Table 1.2 for more detail.

The growth mindset studies included in this review have a combined sample size of 7,781, which also varied widely among studies, with an average of 708 students per study ($SD = 1,093$). Six studies (55%) were conducted with middle school students and five (45%) with high school students. Schools characteristics varied among studies, but of the four studies that were conducted in one school, three were low income schools with student populations that were at least half black or Hispanic. See Table 1.3 and Figure 1.5 for more detail.

Methods: Meta-Analysis

Analytic Strategy

To evaluate an intervention, meta-analysis can be used to combine findings across a collection of studies to more closely estimate its true effect size (Lipsey & Wilson, 2001). Several important decisions must be made before attempting to conduct a methodologically sound and trustworthy meta-analysis, including: (1) choosing and calculating the effect size, (2) choosing a fixed effect or random effects model and assessing heterogeneity, (3) defining outcome variables (4) specifying the populations of interest and (5) considering issues of power. Each of these decisions will be discussed further in the following sections.

Calculating effect sizes. This study uses Cohen's d to calculate the average effect of growth mindset and stereotype threat interventions on students' GPAs. The included studies used linear and multilevel regression models, independent sample t-tests and one-way ANOVA – and they reported regression coefficients and t and F test statistics, respectively. When Cohen's d was not reported, test statistics were converted to a Cohen's d , a standardized mean difference (Littell, Corcoran, & Pillai, 2008). Per notation from Borenstein, Hedges, Higgins and Rothstein (2011), Cohen's d is calculated by dividing the difference in means between a treatment and control group by their pooled standard deviation:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s_{pooled}}$$

If a study reported a Cohen's d , to be entered into the meta-analysis software used for this study (Comprehensive Meta-Analysis, 3.0; Borenstein, Hedges, Higgins, & Rothstein, 2005), data must also be available on the variance of the effect size or the group sample sizes (i.e., individual samples for treatment and control groups). Many of the studies in our sample reported a Cohen's d (six of 13 stereotype threat studies reported d , and four of 11 growth mindset studies reported a Cohen's d). For two stereotype threat studies, Cohen's d was

available, but the variance or group sample size was unavailable (Good et al., 2003; Shnabel et al., 2014). For three growth mindset studies, Cohen's d was available, but the variance was unavailable (Good et al., 2003; Yeager et al., 2014 – Study 3; Yeager et al., 2016 – Study 2). The effect sizes for these studies were recalculated using available test statistics (i.e., t and total sample size). Of the five studies for both stereotype threat and growth mindset where published effect sizes needed to be recalculated, in three studies, the calculated effect size was the same as the published effect size. In two studies (Good et al., 2003, Yeager et al., 2014 – Study 3), the calculated and published effect sizes did not match. For the Good et al., (2003) stereotype threat study, the reported effect size was $d = 1.51$ and the calculated effect size was $d = 1.71$. For the Good et al. (2003) growth mindset study, only one of three calculated effect sizes did not match the three published effect sizes. The reported Cohen's d for female's math standardized test scores was $d = 1.13$ and the calculated effect size was $d = 1.31$. For the Yeager et al. (2014) study, $d = .62$ was the calculated effect size vs. $d = .54$ that was published). See Tables 1.6 and 1.7 for more detail on available statistics used to enter or calculate effect sizes for each study.

An additional sensitivity test was run using a Hedge's g rather than Cohen's d effect size on a reduced sample of studies. For studies with small sample sizes, the best practice for estimating standardized mean differences is to convert effect sizes to a Hedge's g , which reduces the potential that effect sizes are over-inflated by small sample size (Hedges, 1981). However, converting Cohen's d to Hedge's g requires sample sizes for both the treatment and control group, which was not available in one study, where the effect size for meta-analysis was calculated using only the reported Cohen's d and variance (Hanselman et al., 2017). To retain this study in the meta-analysis and to maintain as large an analytic sample as

possible, Cohen's d was used to estimate average treatment effects. A comparison of estimates using Cohen's d on the full sample versus using Hedge's g on a reduced sample showed that substantive findings did not change. In sum, it is unlikely that the use of Cohen's d inflated the results presented here, and Cohen's d was used in all subsequent analyses.

Separate meta-analyses were conducted by intervention type (i.e., growth mindset and stereotype threat) using Comprehensive Meta-Analysis (3.0; Borenstein, Hedges, Higgins, & Rothstein, 2005), according to the following formula:

$$M^* = \frac{\sum_{i=1}^k W_i^* Y_i}{\sum_{i=1}^k W_i^*}$$

where the average effect size is the weighted mean effect size (M^*), a function of the sum of the products of the individual effect sizes (Y_i) and their weights (W_i^*), divided by the sum of their weights (Borenstein et al., 2011). All effect sizes were weighted based on sample size and standard error of the effect size (Borenstein et al., 2011). Studies were weighted using an inverse variance approach, according to the following formula:

$$W_i = \frac{1}{\left(\frac{\sigma^2}{n}\right) + \tau^2}$$

where σ^2 is the variance of the effect size, n is the sample size of the study, and τ^2 is the variance in effect sizes between studies (Borenstein et al. 2011). Studies were weighted to avoid overrepresentation of potentially inaccurate or exaggerated effects from individual studies. Through this process, studies with large sample sizes and/or small standard errors were upweighted and studies with small sample sizes and/or large standard errors were downweighted. A visual representation of study weights is included in two forest plots (see Figures 1.6 and 1.7), where the diamond represents the overall point estimate, and each square represents the weight of an individual study's effect size in the overall point estimate.

The squares becomes larger with higher weight values. Studies with large confidence intervals around their effect sizes have smaller boxes because these estimates are considered to be less precise. Studies with small confidence intervals around their effect sizes have larger boxes because these estimates are considered to be more precise. To produce the most precise point estimate possible, studies represented by small boxes factor into the overall estimate less than studies whose effect sizes are represented by larger boxes. This weighting process increases confidence in the overall point estimate, as it emphasizes studies with effect sizes that are more likely to be accurate estimates of the true effects of the interventions.

Fixed effects vs. random effects model. This study used a random effects model, rather than a fixed effect model. Fixed effect models assume that studies included in the meta-analysis are highly similar to each other, akin to calculating an average by sampling the same population multiple times (Borenstein et al., 2011). In a random effects model, an additional error term is added to account for between-study variance – in other words, the potential that studies differ from each other. The random effects model assumes a distribution of effects, where each study's individual effect size has its own variance and confidence interval and where the estimated average effect size across studies has its own variance and confidence interval (Borenstein, Hedges, Higgins, & Rothstein, 2010). Because the studies in these analyses were conducted in different schools by different researchers, it was expected that there would be between-study variation. The random effects model accounts for this heterogeneity across studies and, because of the added error term, is a more conservative point estimate than a fixed effect model (Borenstein, et al., 2010).

Random effects models are especially important when the heterogeneity in studies is high; therefore, as a preliminary step, two heterogeneity statistics were calculated: The Q and I^2 statistic (Borenstein et al., 2010, Borenstein et al., 2011). Heterogeneity of the sample of studies in each intervention category was first assessed using the Q statistic, according to the following formula:

$$Q = \sum_{i=1}^k W_i (Y_i - M)^2$$

where W_i is the study weight, Y_i is the study's effect size, M is the average effect size of all studies, and k is the number of studies in the sample. In essence, the Q statistic is a weighted sum of squares that captures how much each study's individual effect size deviates from the average effect size (Borenstein et al., 2011). The Q statistic is evaluated through a chi-square test to determine whether studies in the sample are significantly different from one another (i.e., $p < .05$), which was the case for both the meta-analysis of stereotype threat interventions ($Q = 26.61, p < .001$) and growth mindset interventions ($Q = 16.245, p < .01$). These Q statistics suggest there is high heterogeneity among studies for both types of interventions.

Heterogeneity was also assessed using the I^2 statistic, which indicates the proportion of the variance in the results of the meta-analysis that may be attributable to differences among studies (Higgins et al., 2003; Higgins & Thompson, 2002). In a random effects model, the distribution of effect sizes has its own variance, expressed as τ^2 . Heterogeneity between studies can be assessed using the I^2 statistic, which is a function of the variation in effect sizes between studies in the analytic sample (τ^2) and sampling error, according to the following equation:

$$I^2 = \tau^2 / (\tau^2 + V_{\text{error}})$$

I^2 values are expressed as a percentage, with larger values indicating greater heterogeneity in effect sizes across studies. In the final samples, stereotype threat studies had an I^2 value of 69.94 and growth mindset studies had an I^2 value of 63.06, meaning that in each category, roughly 60-70% of the variation in the average effect size may be attributable to heterogeneity among studies (Higgins & Thompson, 2002). These I^2 values fall at or just outside the high end of the Cochrane Collaboration's range for moderate heterogeneity (30-60%; Higgins & Green, 2011), which indicates that a portion of the variance in the point estimates may be attributable to differences between studies.

In sum, both the Q and I^2 statistics suggest high heterogeneity in our sample. Though heterogeneity is not uncommon in meta-analyses and is addressed through the use of a random effects model, it can also indicate that there may be moderating effects of study characteristics (e.g., type of academic outcome, student population; Borenstein et al., 2010; 2011, Higgins et al., 2003). To reduce the likelihood that these between-study differences confound our findings, we specified outcome variables and student populations to be as homogeneous as possible, explained in further detail in the next two sections.

Defining academic outcomes. In this study, we focus on overall GPA as the primary indicator of academic achievement in our meta-analyses. Seven of the 11 growth mindset interventions and nine of the 12 stereotype threat interventions provided data on GPA and were used for meta-analysis (See Tables 1.6 and 1.7 for included studies). In other published meta-analyses of academic outcomes, the operationalization of achievement varies widely. Some meta-analyses combine outcomes like standardized test scores and GPA (e.g., Durlak et al., 2011), but in this analysis, we focus solely on GPA for several reasons. First, GPA was the most frequently used outcome across studies in both interventions. Second, though many

studies reported on more than one type of outcome, there is evidence that effects of academic interventions vary substantially between effects on standardized outcomes (i.e., standardized test scores) and effects on non-standardized outcomes (i.e., grades; Castro et al., 2015). In particular, changes to students' beliefs about themselves are often more strongly linked with changes in their grades, which would be reflected in their GPA, compared to standardized test scores (Hansford & Hattie, 1983; Valentine, DuBois, & Cooper, 2004). Third, even among non-standardized outcomes, students' full GPA can often be a more reliable measure of achievement than single course grades (Bacon & Bean, 2006), and the difference between course grades and full GPA provides a level of distinction between whether the intervention effects generalized to students' academic performance broadly or whether effects were confined to a specific course (Cohen et al., 2009). Finally, GPA is most commonly used for future academic decisions and outcomes (e.g., college acceptance), and identifying interventions that boost students' overall school performance would be helpful in supporting future academic success (Belfield & Crosta, in press; Noble & Sawyer, 2004; Noble & Sawyer, 2002).

Our analysis included GPA effects during the school year in which the intervention was conducted. In one study (Sherman et al., 2013 – Study 2), effect sizes were reported from two time points during the academic year (i.e., end of first semester, end of year). Because the intervention was not concluded at the end of the first semester, and the implementation of the values reflections continued into the second semester, only the end of year effect size was used for analysis. In all other studies, effect sizes were calculated after all intervention sessions were completed; thus, the end of year estimate from Sherman et al. (2013 – Study 2) maintains consistency in the effect sizes used across studies. In addition,

one study (Wynne, 2011) reported effects across students' courses individually. To address this, per Borenstein (2011), when a study reports multiple outcomes from the same participants at the same time, they can be averaged together to create a composite effect size. Because effect sizes from all courses were presented in the Wynne (2011) study, they were combined to create a composite effect size for GPA (Bacon & Bean, 2006; Borenstein, et al., 2011). See Tables 1.6 and 1.7 for more detail. One growth mindset study (Oyserman et al., 2015) and four stereotype threat studies (Borman et al., 2016; Cohen et al., 2009; Hanselman et al., 2014; Sherman et al., 2013 – Study 1) provided one-year follow up data, but because there are only a small number of studies with follow-up data, they were not included in the meta-analyses but will be discussed further descriptively. In addition to the quantitative summary of GPA effects, we also discuss effects on single course grades and standardized test scores. Because few studies provided data on these outcomes, they will be discussed descriptively, rather than via meta-analysis, as descriptive analysis can still contribute to the understanding of intervention effects (Loeb, Dynarski, McFarland, Morris, Reardon, & Reber, 2017).

Populations of interest. We assessed the impact of each intervention on its target population. Growth mindset interventions target the entire population of students, and data from full study samples were included in this analysis, which compares achievement differences between students who received the growth mindset intervention and those who did not. One study (Paunesku et al., 2015) only reported effects for a subgroup of students who were historically lower performers, and data from this study was therefore not included in the point estimate for growth mindset interventions.

Stereotype threat interventions typically target a threatened group (e.g., black and Hispanic students), and some studies report effects on the threatened group only. Thus, the analysis in this study included data from black/Hispanic students only. In only one study (Good et al., 2003), gender was used to determine threatened vs. nonthreatened groups (i.e., boys and girls). Because this study assessed effects on standardized test scores, it is not included in the meta-analysis but will be discussed further descriptively. As a result, this meta-analysis reflects the difference in achievement between black and Hispanic students who received the intervention and black and Hispanic students who did not.

Data from students of non-threatened identities (i.e., white and Asian students) were excluded for two reasons. First, stereotype threat interventions target those students whose performance may be suppressed by negative stereotypes about their intelligence or academic ability – which many studies identify as students of color. Second, data from threatened groups was presented in all stereotype threat studies included in this review, while data from the whole sample or from non-threatened students was not available in most studies. One study (Wynne, 2011) was conducted in a school where all students identified as either black or Hispanic, and four other studies (Borman et al., 2016; Good et al., 2003; Protzko & Aronson, 2016; Simmons, 2013) reported analyses for threatened students only. In sum, focusing on intervention effects among students of threatened identities not only retains the largest analytic sample possible but also estimates effects for the population the intervention was designed to support.

In our analysis, black and Hispanic students were collapsed into a single threatened group for two reasons. First, prior research on the reduction of threat for black and Hispanic students has indicated that effects are similar between races (Nadler & Clark, 2011). Second,

several studies included in this analysis do not report results by racial subgroup and instead collapse student data to report comparisons between threatened and non-threatened students (Borman et al., 2016; Hanselman et al., 2017).

Issues of power. For many quantitative analyses, power calculations are used to assess whether the sample size is sufficiently large to detect effects, reducing the chance of a Type II error – the conclusion that there are no intervention effects when, in actuality, there are (Bachetti, 2013; Hoenig & Heisey, 2001). For a random effects model in a meta-analysis, power is calculated according to the following equation (Borenstein et al., 2011; Valentine, Pigott, & Rothstein, 2010):

$$p = 1 - \Phi(c_\alpha - \lambda^*)$$

where Φ represents the standard normal distribution of c_α , the critical value for the standard normal distribution ($c_\alpha = 1.64$ for a one-tailed test at $\alpha = .05$) minus λ^* . λ^* is the estimated average effect size over the square root of the estimated variance of the random effects model (v^*):

$$\lambda^* = \frac{ES - 0}{\sqrt{v^*}}$$

v^* is a function of the variance of the effect size over the number of effect sizes in the analysis (i.e., the number of studies), k :

$$v^* = \frac{v^*}{k}$$

The variance v^* is a function of the average group sample sizes (i.e., treatment, n_T , and control, n_C) across studies, the estimated effect size and the estimated between-study variance, τ^2 :

$$v^* = \frac{n_T + n_C}{n_T * n_C} + \frac{ES^2}{2(n_T + n_C)} + \tau^2$$

In this study, power was calculated separately for both the meta-analysis of stereotype threat and the meta-analysis of growth mindset studies. Nine stereotype threat studies assessed effects on GPA. For the meta-analysis of stereotype threat studies, $k = 9$, which had average group sample sizes of approximately $n = 200$ students, power was estimated conservatively, assuming a high level of heterogeneity among studies (per Valentine et al., 2010, $\tau^2 = 3$) and a small effect size ($d = .20$; Cohen, 1988). Estimated power for this meta-analysis is .99. Three of the nine stereotype threat studies did not provide group sample sizes and were not able to be included in the power analysis (Cohen et al., 2009; Protzko & Aronson, 2016; Shnabel et al., 2013). However, these studies had an average overall sample size of $M = 350$ ($SD = 105$), suggesting that their group sample sizes were likely similar to the studies included in the power analysis.

Seven growth mindset studies assessed effects on GPA. For the meta-analysis of growth mindset studies, $k = 7$, with average group sample sizes of approximately $n = 100$ students, assuming a high level of heterogeneity among studies, $\tau^2 = 3$, and a small effect size, $d = .20$, power was estimated at .56. Though estimated power is low for the meta-analysis of growth mindset studies, it should be noted that group sample sizes were available for only three of the seven included studies; therefore, full data were not available for this power analysis. It is likely that if the studies without published group sample sizes were included, the estimates of power would have been larger. In the four studies that were not able to be included in the estimate of average group sample size (Eskreis-Winkler et al., 2016 – Study 4, Eskreis-Winkler et al., Study 5, Yeager et al., 2014 – Study 3, Yeager et al., 2016 – Study 2), overall sample sizes were quite large ($M = 1,393$, $SD = 1,985$). Likely, group sample sizes of these studies were also large, which suggests that the average group sample

sizes for growth mindset studies is larger than the sample sizes used in our power analysis, which would increase power considerably.

Our power calculations suggest that the meta-analyses in this study are likely adequately powered for stereotype threat studies but may be underpowered for growth mindset studies; however, there is debate in the literature regarding the utility of calculating power for meta-analyses. Some scholars argue that power calculations are most useful in primary studies, when conducted a priori, in order to estimate sufficient sample sizes before participants are recruited (Cumming, 2014; Hoenig & Heisey, 2001; Smith & Bates, 1992). Similarly, others argue that for meta-analysis, calculating power is most helpful for planning reviews in advance (i.e., planning several primary studies that are intended to be used for meta-analysis; Borenstein et al. 2011).

In contrast, for meta-analysis, sample size is determined by available research and inclusion/exclusion criteria. When a high number of studies is available, sample size can be increased by relaxing inclusion criteria. Considering the aims and research questions of this study, the implementations of stereotype threat and growth mindset interventions in schools are relatively new areas of research, where a limited number of studies is available (Borenstein et al., 2011). Because these interventions have emerging evidence bases, maintaining rigorous study design as part of the inclusion criteria (i.e., RCTs, school-based rather than research laboratory implementation) is important for a quality assessment of these interventions, despite the small sample sizes of studies that test them (Cumming, 2014).

In addition, power calculations are not included as best practice for meta-analysis in the Cochrane manual (Higgins & Green, 2011). In fact, in their database of meta-analyses, which must meet rigorous methodological and reporting standards, the median sample size is

six studies, suggesting that meta-analyses conducted on a small number of studies can still provide important statistical and practical information about interventions (Borenstein et al., 2011).

Results: Meta-Analysis

Intervention Effects: Stereotype Threat

Stereotype threat and GPA. Of the nine studies that examined the effect of identity affirmation on students' GPAs, four indicated statistically significant improvement among students with threatened identities. Across all studies, effect sizes ranged from $d = .00$ - $.56$. Meta-analysis, which averaged the effects across studies, revealed an overall point estimate of $d = .19$ ($SE = .06$, $p = .01$; see Tables 1.4 and 1.6). This estimate is considered a small effect (Cohen, 1988).

Subsequent sensitivity testing was conducted through one study removed analysis (Borenstein et al., 2005; Borenstein et al., 2011), which compares the effect of removing an individual study on the overall point estimate. One study removed analysis provides an additional check on outlier studies that may be influencing overall findings (e.g., Kedzior & Laeber, 2014). For the meta-analysis presented here, the one study removed technique indicated that findings were robust across studies, showing that if any one study was removed from the sample, findings did not change substantially (See Table 1.8).

Not included in this meta-analysis are the effect sizes from one-year follow-up data: Of the four studies that provided long term follow-up, two demonstrated that effects not only persisted but were magnified over time. Cohen et al. (2009) and Sherman et al., (2013, Study 1) reported increases from $d = .28$ to $.38$ and from $d = .29$ to $.43$, respectively, at one year follow-up. Borman et al., (2016) and Hanselman et al. (2017), who found no initial

significant effects of the intervention, did not report significant effects at one-year follow-up (See Table 1.4).

Stereotype threat and exam scores. Although there were not sufficient studies to conduct meta-analysis, in this section, we descriptively review findings for other outcomes. Only one of the four studies that examined effects on standardized test scores found significant improvements. Wynne (2011) demonstrated that the values affirmation task boosted students' math standardized test scores ($d = .45$). No other studies found significant effects on students' test scores, with effect sizes ranging from $d = -.001$ to $.28$. See Table 1.4 for a summary of the effect sizes for these studies.

Stereotype threat and course grades. Only one of the three studies that assessed effects in a single course found significant improvements. Bowen et al., (2013) found significant improvements in students' social studies grades ($d = .57$), though this sample included both white and black students. Dee (2015) and Bratter et al. (2016) did not find improvements in students' social studies and English grades, respectively, with effect sizes ranging from $d = -.001$ to $.16$. See Table 1.4 for a summary of the effect sizes for these studies.

Intervention Effects: Growth Mindset

Growth mindset and GPA. Among the seven studies that tested the effect of growth mindset interventions on students' GPAs, three found statistically significant improvement, with effect sizes ranging from $d = .09$ to $.62$. Together, meta-analysis revealed these studies had a statistically nonsignificant point estimate of $d = .11$ ($SE = .08$, $p = .18$). Of note, one study (Brougham, 2016) found a statistically significant *decrease* in GPA ($d = -.68$, $p = .01$). Sensitivity testing through one study removed analysis indicated that the inclusion of

Brougham (2016) in the sample altered the point estimate substantially, changing substantive findings. Without the study, the point estimate is statistically significant ($d = .14$, $SE = .049$, $p = .004$; see Table 1.9), and these results are robust across all other studies (see Table 1.10), suggesting that the Brougham (2016) study may be an influential outlier in these data (Borenstein et al., 2011). The estimate without Brougham (2016) in the sample, $d = .14$, would also be considered a small effect (Cohen, 1988).

Not included in this estimate is one year follow-up data from Oyserman et al. (2006). Similar to the results from the stereotype threat intervention, the effect size from this study was larger at one year follow-up, increasing from $d = .25$ to $.30$.

Growth mindset exam scores. Two studies examined the effects of growth mindset intervention on students' standardized test scores. Oyserman et al. (2006) found significant improvements in the proportion of exams that students passed ($d = .36$). Good et al. (2003) saw increases in all students' reading test scores ($d = .51$), with significant effects for females' math scores ($d = 1.31$) and marginal effects for males' math scores ($d = .62$, $p = .05$).

Growth mindset and course grades. Only two studies assessed effects in specific course grades. Blackwell et al. (2007) found that participation in the intervention boosted students' math grades ($d = .62$), while Wilkins (2014) found effects for students' science grades, but not math grades ($d = .26$ and $d = .02$, respectively). See Tables 1.5 and 1.7 for a summary of effect sizes from all growth mindset studies.

Discussion

This systematic review and meta-analysis tested effects from randomized controlled trials of interventions for growth mindset and stereotype threat on students' academic

performance. Results indicated that both interventions may produce positive, albeit small, effects that may be highly contingent upon school context.

Reducing Stereotype Threat

Meta-analysis of stereotype threat interventions demonstrated a small overall effect of the intervention on academic performance ($d = .19$). Although some studies did not replicate the magnitude of findings from the seminal Cohen et al. (2006) values affirmation intervention, most studies did demonstrate some degree of academic improvement for threatened students: Seven of twelve studies found effects of the intervention on at least one academic outcome.

There is some evidence that the effects of stereotype threat intervention may grow over time, although there was not an adequate number of studies to test longitudinal effects through meta-analysis. Studies that used growth curve models of students' GPAs over the course of two to three years (e.g., Sherman et al. 2013, Study 1) reveal that on average, the GPAs of all middle school students decline from sixth to eighth grade, regardless of race. This trend aligns with a wealth of literature describing a "middle school slump" that plagues motivation and achievement in early adolescence, particularly for at-risk students (Akos, Rose & Orthner, 2015; Wang and Eccles, 2013). Thus, to be clear, the effects of the values affirmation activity depict a reduction *in the decline* of GPAs for threatened students who receive the intervention. In light of the available statistics on the achievement gaps between racial/ethnic groups, reducing the decline in GPA of non-white students to match their white peers is still noteworthy (Kena et al., 2016). Though the increased magnitude of effects in one- to two-year follow-up assessments is promising, these findings should be interpreted with caution, as very few studies provided additional data. There is ample evidence that

intervention effects often decay over time (DiClemente, Santelli, & Crosby, 2009), and it is possible that other studies did not publish follow up data due to lack of significant findings.

The fact that not all replications of the intervention produced similar results suggests that individual study characteristics may affect the likelihood of the intervention's success, a possibility that is supported by the heterogeneity we observed among studies. In terms of intervention implementation, in essentially all replications of the Cohen et al. (2006) study, investigators were careful to implement factors that had been previously identified as central to the intervention mechanism: (a) conducting the reflections just prior to an event where threat of evaluation would be high (i.e., before the end of the grading period or administration of a standardized test) and (b) messaging the intervention as a normal classroom activity to promote authenticity of the assignment. In only one study (Protzko & Aronson, 2016) were the investigators required to inform students that the activity was for a research study, and in light of null findings, they called to question whether messaging the activity as a research task interfered with students' internalization of the affirmation. However, in general, investigators were able to replicate the intervention procedures with fidelity, suggesting other contextual factors (e.g., student population) may explain the differences in effects.

One source of variation among studies that may affect intervention effects is the racial composition of the study body, as there is evidence that students of potentially threatened identities experience threat differently by school (Tyson, 2011). Threat may manifest most intensely when students of threatened identities are considerably outnumbered by students of non-threatened identities. If representation of a student's identity in a group is low, identity can factor more strongly into their self-assessments of ability, increasing threat

(Deaux & Major, 1987; Devine, 1989; McGuire & Padawer-Singer, 1976). For example, in studies that examine gender, stereotype threat and test performance, the reduction in scores for female students is directly proportional to the number of males in the room. Inzlicht and Ben Zeev (2000) found that performance was highest among females who tested with only other females and lowest among those who tested as the only female in a room with all males. Females who tested in a room with both males and females still experienced a reduction in scores, even if female students were in the numerical majority. Replications of the Inzlicht and Ben Zeev (2000) study have shown that the effects in mixed group settings (i.e., males and females) are not mitigated by the promise that scores will not be made public to others, suggesting that the composition of students itself has a strong impact on the level of threat in a given context (Inzlicht & Ben Zeev, 2003). Though these studies have primarily focused on gender effects in math, the same may be true for race or other stereotyped identities.

In our sample, many of the studies where the stereotype threat intervention boosted student achievement had sizeable numbers of both non-threatened and threatened students, where it's possible that the intervention may have reduced threat created by a more diverse student body (Cohen et al., 2009; Sherman et al., 2013 - Study 1; Sherman et al., 2013 – Study 2; Shnabel et al., 2013). In the studies where data were pooled across several different schools, it is difficult to assess the degree to which a school's racial composition influenced results. However, the potential for moderation by student composition was explored further by Hanselman et al. (2014), who re-analyzed data from the Borman et al. (2016) study included here, by creating an additional variable for threat that accounted for the racial composition of the school. Their results indicated a suppression effect in the original

findings, showing an interaction between the intervention and school context. When data were re-analyzed, intervention effects were present only for students of threatened identities in high threat schools. Evidence from this follow-up study indicates that the racial makeup of the school may play a role in a school's threat context and whether the values affirmation task would benefit students of threatened identities.

In addition, as demonstrated by Bowen et al. (2013), who found significant improvements in a low-income school among white students who received the intervention, there may be aspects of students' identities beyond race or gender (e.g., socioeconomic status) that are under threat in schools. None of the studies identified in this systemic review and meta-analysis addressed other aspects of identity, such as socioeconomic status, ability status or sexual identity. Yet these identities may be salient to students, and students may use them to make attributions about their academic performance and intelligence. Spencer and Castano (2007) found that presenting an assessment as a diagnostic of natural ability reduced confidence for low-income students and negatively affected their performance on the test. Similarly, Quinn, Kahang, & Cocker (2004) tested whether stigmatized identities that are potentially concealable (unlike gender and race) can still affect performance when activated, showing that, among students who receive mental health services, disclosing treatment before the assessment reduced students' test performance. Future studies should explore other types of stereotype threat beyond race to examine possible intervention effects.

Several other contextual factors may help explain a lack of findings in certain school contexts. In the Wynne (2011) study, for example, all students were either black or Hispanic. In this school context, threat may be low because each student is surrounded by other students of color, making identity less salient in their attributions about ability. Threat may

also be low in the Wynne (2011) study because this school has a strong academic record as a high performing charter. Threat is intensified by relative underperformance of a stereotyped group in comparison to a non-stereotyped group – meaning, if students of color have historically underperformed in comparison to white or Asian students within the school, the effects of stereotype threat may be stronger (Tyson, 2011). A uniformly high performing school that exclusively serves students of color, such as the school in the Wynne (2011) study, may not be an ideal context for stereotype threat intervention. It may be the case that performance gaps between students are quite narrow or that students do not attribute any performance gaps to race, reducing the impact of stereotype threat on their academic achievement. Future replications or moderator analyses in existing studies should further explore the characteristics of a high threat context and how context influences intervention effects.

Promoting Growth Mindset

All but two of the eleven studies testing growth mindset interventions found significant effects in at least one outcome, though effect sizes varied in magnitude. Meta-analysis revealed a small and nonsignificant effect size on GPA, though sensitivity testing revealed that this estimate may have been unduly affected by the Brougham (2016) study. When this study was excluded, meta-analysis resulted in a small positive effect ($d = .14$). The Brougham (2016) study found highly discrepant effects, showing that students who received the growth mindset intervention experienced a decrease in academic achievement. It should be noted that the group sample sizes for this study were small (just over 30 students in the treatment and control groups), and both treatment and control groups were pooled from students from two different high schools. The treatment group in one school, for example,

contained as few as 10 students. The average drop in GPA for these ten students ($M = -.35$, $SD = .84$) and null effects among students in the treatment group in the other school ($M = .09$, $SD = .47$), paired with unexpected gains in GPA among control group students in both schools ($M = .32$, $SD = .38$; $M = .34$, $SD = .59$), produced the overall negative effect size.

It is important to note that, though the effect size is large in magnitude ($d = -.68$), it is driven primarily by large gains in the control group rather than substantial iatrogenic effects in the experimental group. Nonetheless, the results of the Brougham study (2016) are inconsistent with other studies in this sample and raise questions about methodological rigor and study quality. In particular, the author does not provide data from any preliminary testing to check for successful randomization between treatment and control groups to assess whether they differed by GPA at the outset of the study. It's possible that in such a small sample, randomization was not sufficient to produce balance across groups. However, the author does provide a personal assessment of group dynamics that indicates that all control groups sessions were rated with a score of 3 (seemingly, the highest score). In comparison, treatment group sessions were rated with 2s and 1s, suggesting that the dynamic among students in the treatment groups was worse than among those students in the control groups. Her qualitative notes indicate that treatment groups were more unruly and required more behavior management and encouragement for participation. In addition, the author noted that one student in the treatment group of ten students who showed a decline in achievement experienced a significant trauma throughout the course of the intervention – her father was killed in between intervention sessions. In the absence of original study data, it is not possible to assess whether the average GPA of the treatment group was unduly influenced by

one or more outlier students whose achievement may have been affected by factors outside the study context.

In light of the results of sensitivity testing, and in comparison to the consistently neutral or positive findings among other studies, which had larger samples and greater fidelity of implementation, it appears as though the Brougham (2016) study is an outlier. In comparison to other studies in this sample whose data depict a trend of declining GPA among students in the control group and an even sharper decline for high risk students (e.g., Blackwell et al., 2007; Burnette et al., 2017; Paunesku et al., 2015), the significant increase in achievement demonstrated by the control group students in the Brougham (2016) study appears especially unusual and is potentially attributable to group composition or systematic attrition more than to any influence of intervention sessions for either group. However, the possibility that growth mindset intervention could negatively affect students' achievement should not be ignored and should be considered in future research.

The implementation of growth mindset interventions lacked the consistency and specificity of the values affirmation task, creating several potential explanations for the variation in effects. Most notably, studies varied widely in the length of time that students were exposed to the intervention. However, effects were found in studies that conducted mindset lessons over the course of an entire year (e.g., Blackwell et al., 2007, Good et al., 2003) and studies where students received one or two lessons over the course of a few days (e.g., Paunesku et al., 2015; Yeager et al., 2014 – Study 3; Yeager et al., 2016 – Study 2). Similarly, effects were found across studies that taught students about the brain and studies that retrained students' attributions about academic challenge and critical feedback (Eskreis-Winkler et al., 2016 – Study 1; Eskreis-Winkler et al., 2016 – Study 2; Oyserman et al.,

2006). Such variation in findings makes it difficult to assess the critical components of a growth mindset intervention. In the only large-scale study to compare various intervention elements, Yeager and colleagues (2016) found differences in proximal outcomes (e.g., challenge-seeking behavior, academic engagement) based on small tweaks in the intervention design, including the addition of a celebrity endorsement ($p = .07$), scientific evidence ($p = .35$), greater elaboration on the benefits of a growth mindset ($p = .006$), increased focus on the direct impact to participating students ($p = .03$), and more evidence to refute fixed mindset beliefs ($p = .002$). The Yeager et al. (2016) study did not monitor changes in academic outcomes, and it studied iterations of the intervention that are relatively narrow in scope. However, in light of their findings that small tweaks to growth mindset messaging can have effects on student outcomes, we echo their call for more research that directly examines iterations on growth mindset intervention elements to identify those that most effectively drive differences in student behavior and learning.

In addition, despite initial findings from seminal studies like Blackwell et al., (2007) and Good et al. (2003), which indicated that growth mindset interventions were universally beneficial, it is possible that growth mindset intervention works differently for different students. In subsequent moderator analyses, Yeager et al. (2014 – Study 3) demonstrated that significant overall effects were driven by black students, and interaction analyses in three other studies (Eskreis-Winkler, 2016, Study 4; Paunesku et al., 2015; Yeager et al, 2016 – Study 2) showed that effects were driven by low performing students. It is possible that in other studies with significant findings, effects of growth mindset interventions were driven by the most vulnerable groups. Similarly, it is possible that in studies with non-significant findings, effects in low performers may be masked by null effects among high performers.

As more studies are conducted on the effects of growth mindset interventions, it will be necessary to test for differential impacts among students.

Strengths and Limitations

Though this study provides a comprehensive synthesis of the extant literature on growth mindset and stereotype threat interventions, it is not without limitations. First, a small number of studies have been conducted on these interventions, and they have not all looked at similar outcomes. In addition, there was considerable heterogeneity among studies for both interventions, but with a small number of studies, it was not possible to directly test potential moderators, reducing our ability to draw strong conclusions about the differences among studies and their impact on intervention effects. We addressed the potential for moderating effects by limiting meta-analysis to specific populations and specific outcomes, but it should be noted that this does not eliminate all sources of variation that may influence study effects. Also, because the focus of this review was limited to achievement outcomes, rather than other aspects of students' school experience (e.g., sense of belonging, stress, motivation, enjoyment of school), our study does not address whether these interventions positively impact students in other aspects of their well-being.

As is the case with much research on academic achievement, it is possible that ceiling effects stunt the measurement of intervention effects for students who appear to be unaffected by the interventions in these studies (e.g., white students, historically high performing students). High performing students may be learning more or may be acquiring knowledge at a greater depth, but such outcomes may not be captured due to the scope or scales for classrooms grades or standardized tests. It is also possible that high performing students may experience no difference in grades or test scores but may shift their engagement

in school to growth mindset rather than fixed mindset goals for achievement (Yeager et al., 2016). A review of growth mindset studies (see Senko, Hulleman, & Harackiewicz, 2011) showed that growth mindset beliefs are linked with cooperativeness, help-seeking behavior, better self-regulation, more positive affect, and deeper learning strategies, whereas fixed mindset beliefs are associated with increased anxiety, low self-efficacy and disorganized work habits. It is possible that, as a result of intervention, high performing students may perform equally to their prior performance but may be choosing more adaptive behaviors and/or feeling less anxious and more competent. Future research and meta-analyses could summarize effects on outcomes other than academic achievement to assess the full impact of growth mindset intervention.

Nonetheless, this study has several strengths that contribute to the knowledge base around growth mindset and stereotype threat interventions. Great length was taken to reduce bias by conducting a thorough systematic review of the literature and by choosing only studies with rigorous designs and methodologies, where each intervention has been tested in large scale RCTs with thousands of students. To promote accuracy of findings, when possible, we conducted sensitivity testing around analytic decisions (e.g., choice of effect size, potential for outlier studies) to ensure that results were not biased by methodological choices.

This review supports the potential of these interventions to promote school success and reduce achievement disparities, and it provides trailheads for future research. Additional studies could build upon the analysis presented here by exploring the roles of student identity and school's threat context and by consolidating evidence for outcomes beyond grades and exams so that the full effect of these interventions may be captured. We chose to avoid

introducing bias in our analysis by not excluding any studies that met pre-established inclusion criteria, but when more intervention studies are available in each area, it could be beneficial to conduct systematic ratings of study quality to determine whether potential outliers like the Brougham (2016) study are clouding conclusions drawn from average effects. In addition, as more research becomes available, future meta-analysis could be also be improved by the use of cumulative analysis, which compiles evidence over time to show the accumulation of evidence for an intervention and how results may have shifted over periods of time (Borenstein et al., 2011).

Conclusion

Overall, this study indicates that, while both growth mindset and stereotype threat interventions have demonstrated effects on academic outcomes like grades and standardized test scores, the magnitude of effects varies considerably among studies, and average effects may be small. These findings should be interpreted with caution, given the small number of available studies and the heterogeneity among studies in each intervention category. More research is needed to examine the specific conditions under which and for whom these interventions are most effective. As it stands, however, this review provides a summary of the current knowledge on two popular interventions, providing practitioners and researchers with direction in improving the academic performance of all students, and especially for those at greatest risk.

REFERENCES: PAPER I

- Akos, P., Rose, R. A., & Orthner, D. (2015). Sociodemographic moderators of middle school transition effects on academic achievement. *The Journal of Early Adolescence*, 35(2), 170-198. doi:10.1177/0272431614529367
- Aronson, J. (1999). The effects of conceiving ability as fixed or improvable on responses to stereotype threat. Unpublished manuscript, New York University.
- Aronson, J., Cohen, G., McColsky, W., Montrosse, B., Lewis, K., & Mooney, K. (2009). Reducing stereotype threat in classrooms: A review of social-psychological intervention studies on improving the achievement of Black students (Issues & Answers Report, REL 2009–No. 076). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- 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(2), 113-125. doi:10.1006/jesp.2001.1491
- Aronson, J., Quinn, D., & Spencer, S. (1998). Stereotype threat and the academic underperformance of women and minorities. In J. Swim & C. Stangor (Eds.), *Stigma: The target's perspective*. New York: Academic Press. doi: 10.1016/B978-012679130-3/50039-9
- Bacchetti, P. (2013). Small sample size is not the real problem. *Nature Reviews Neuroscience*, 14(8), 585. doi:10.1038/nrn3475-c3
- Bacon, D. R., & Bean, B. (2006). GPA in research studies: An invaluable but neglected opportunity. *Journal of Marketing Education*, 28(1), 35-42. doi:10.1177/0273475305284638
- Belfield, C. R., & Crosta, P. M. (2012). Predicting Success in College: The Importance of Placement Tests and High School Transcripts. CCRC Working Paper No. 42. *Community College Research Center, Columbia University*.
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, 12(3), 225-229. doi:10.1111/1467-9280.00340

- Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2013). Enhancing a brief writing intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology, 105*(2), 427-435. doi:10.1037/a0031177
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2005). *Comprehensive Meta-Analysis 3.0*. Englewood, NJ: Biostat.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis*. John Wiley & Sons.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97-111. doi:10.1002/jrsm.12
- Borman, G. D., Grigg, J., & Hanselman, P. (2016). An effort to close achievement gaps at scale through self-affirmation. *Educational Evaluation and Policy Analysis, 38*(1), 21-42. doi:10.3102/0162373715581709
- Bratter, J. L., Rowley, K. J., & Chukhray, I. (2016). Does a self-affirmation intervention reduce stereotype threat in black and Hispanic high schools?. *Race and Social Problems, 8*(4), 340-356. doi:10.1007/s12552-016-9187-4
- Brougham, L., & Kashubeck-West, S. (2017). Impact of a Growth Mindset Intervention on Academic Performance of Students at Two Urban High Schools. *Professional School Counseling, 21*(1), doi:10.1177/2156759X18764934
- Brougham, L. (2016). *Impact of a growth mindset intervention on academic performance of students at two urban high schools*. Retrieved from: <https://irl.umsl.edu/dissertation/48>
- Brown, E. (2013). No child left behind and the teaching of character education. *ABNF Journal, 24*(3), 77-82.
- Burnette, J. L., Russell, M. V., Hoyt, C. L., Orvidas, K., & Widman, L. (2017). An online growth mindset intervention in a sample of rural adolescent girls. *British Journal of Educational Psychology, 10.1111/bjep.12192*
- Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational Research Review, 14*, 33-46. doi:10.1016/j.edurev.2015.01.002
- Cimpian, A., Arce, H. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science, 18*(4), 314-316. doi:10.1111/j.1467-9280.2007.01896.x

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* 2nd edn.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310. doi:10.1126/science.1128317
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, *324*(5925), 400-403. doi:10.1126/science.1170769
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7-29.
- Deaux, K., & Major, B. (1987). Putting gender into context: An interaction model of gender-related behavior. *Psychological Review*, *94*, 369–389. doi:10.1037/0033-295X.94.3.369
- Dee, T. S., & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, *30*(3), 418-446. doi: 10.1002/pam.20586
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1) 5-18. doi:10.1037/0022-3514.56.1.5
- DiClemente, R. J., Santelli, J. S., & Crosby, R. A. (Eds.). (2009). *Adolescent health: Understanding and preventing risk behaviors*. John Wiley & Sons.
- Dinger, F. C., Dickhäuser, O., Spinath, B., & Steinmayr, R. (2013). Antecedents and consequences of students' achievement goals: A mediation analysis. *Learning and Individual Differences*, *28*, 90-101. doi:10.1016/j.lindif.2013.09.005
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256-273. doi:10.1037/0033-295X.95.2.256

- Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, 3, 139-186.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48(2), 90. doi:10.1037/0003-066X.48.2.90
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents? The impact of educational contexts on early adolescents. *American Journal of Education*, 99(4), 521-542. doi:10.1086/443996
- Eskreis-Winkler, L., Shulman, E. P., Young, V., Tsukayama, E., Brunwasser, S. M., & Duckworth, A. L. (2016). Using wise interventions to motivate deliberate practice. *Journal of Personality and Social Psychology*, 111(5), 728. doi:10.1037/pspp0000074
- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 72(1), 218-232.
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5-12. doi:10.1037/0022-3514.72.1.218
- Gonida, E., Kiosseoglou, G., & Leondari, A. (2006). Implicit theories of intelligence, perceived academic competence, and school achievement: Testing alternative models. *The American Journal of Psychology*, 119(2), 223-238. doi:10.2307/20445336
- 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(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development*, 84(5), 1526-1541. doi:10.1111/cdev.12064
- Gutman, L. M., & Midgley, C. (2000). The role of protective factors in supporting the academic achievement of poor African American students during the middle school transition. *Journal of Youth and Adolescence*, 29(2), 223-249. doi:10.1023/A:1005108700243

- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education*, 87(2), 106-124. doi:10.1177/0038040714525970
- Hanselman, P., Rozek, C. S., Grigg, J., & Borman, G. D. (2017). New evidence on self-affirmation effects and theorized sources of heterogeneity from large-scale replications. *Journal of Educational Psychology*, 109(3), 405. doi:10.1037/edu0000141
- Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, 52(1), 123-142.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107-128. doi:10.3102/10769986006002107
- Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, 128(5), 774-795. doi:10.1037/0033-2909.128.5.774
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, 327(7414), 557. doi:10.1136/bmj.327.7414.557
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., ... & Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, 343, doi:10.1136/bmj.d5928.
- Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558. doi:10.1002/sim.1186
- Higgins, J. & Green, S. (2011). Cochrane handbook for systematic reviews of interventions. *The Cochrane Collaboration*.
- Hoening, J. M., & Heisey, D. M. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. *The American Statistician*, 55(1), 19-24. doi:10.1198/000313001300339897
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11(5), 365-371. doi:10.1111/1467-9280.00272
- Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The Implications of threatening environments on intellectual processing. *Journal of Educational Psychology*, 95(4), 796. doi:10.1037/0022-0663.95.4.796

- Johnson, M. K., Crosnoe, R., & Elder, G. H. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education*, *74*, 318–340. doi:10.2307/2673138
- Kedzior, K. K., & Laeber, L. T. (2014). A positive association between anxiety disorders and cannabis use or cannabis use disorders in the general population-a meta-analysis of 31 studies. *BMC Psychiatry*, *14*(1), 136. doi:doi.org/10.1186/1471-244X-14-136
- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., ... & Barmer, A. (2016). The Condition of Education 2016. NCES 2016-144. *National Center for Education Statistics*.
- Kloosterman, P. (1988). Self-confidence and motivation in mathematics. *Journal of Educational Psychology*, *80*(3), 345-351. doi:10.1037/0022-0663.80.3.345
- Lipsey, M. & Wilson, D. B. (2001). *Practical Meta-analysis* (Vol. 49). Thousand Oaks, CA: Sage publications.
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic Reviews and Meta-analysis*. Oxford University Press. doi:10.1093/acprof:oso/9780195326543.001.0001
- Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). Descriptive Analysis in Education: A Guide for Researchers. NCEE 2017-4023. *National Center for Education Evaluation and Regional Assistance*.
- Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience*, *7*, 230–241. doi:10.1093/scan/nsq100
- McGuire, W. J., & Padawer-Singer, A. (1976). Trait salience in the spontaneous self-concept. *Journal of Personality and Social Psychology*, *33*(6), 743-754. doi:10.1037/0022-3514.33.6.743
- Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self-and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, *81*(2), 247. doi:10.1037/0022-0663.81.2.247
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264-269. doi:10.7326/0003-4819-151-4-200908180-00135
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33-52. doi:10.1037/0022-3514.75.1.33

- Nadler, J. T., & Clark, M. H. (2011). Stereotype threat: A meta-analysis comparing African Americans to Hispanic Americans. *Journal of Applied Social Psychology, 41*(4), 872-890. doi:10.1111/j.1559-1816.2011.00739.x
- Noble, J., & Sawyer, R. (2002). Predicting Different Levels of Academic Success in College Using High School GPA and ACT Composite Score. ACT Research Report Series.
- Noble, J. P., & Sawyer, R. L. (2004). Is high school GPA better than admission test scores for predicting academic success in college?. *College and University, 79*(4), 17.
- Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology, 91*(1), 188. doi:10.1037/0022-3514.91.1.188
- Park, S., Holloway, S. D., Arendtsz, A., Bempechat, J., & Li, J. (2012). What makes students engaged in learning? A time-use study of within- and between-individual predictors of emotional engagement in low-performing high schools. *Journal of Youth and Adolescence, 41*(3), 390-401. doi:10.1007/s10964-011-9738-3
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017
- Protzko, J., & Aronson, J. (2016). Context moderates affirmation effects on the ethnic achievement gap. *Social Psychological and Personality Science, 7*(6), 500-507. doi:10.1177/1948550616646426
- Quinn, D. M., Kahng, S. K., & Crocker, J. (2004). Discreditable: Stigma effects of revealing a mental illness history on test performance. *Personality and Social Psychology Bulletin, 30*(7), 803-815. doi:10.1177/0146167204264088
- Rydell, R. J., Shiffrin, R. M., Boucher, K. L., Van Loo, K., & Rydell, M. T. (2010). Stereotype threat prevents perceptual learning. *Proceedings of the National Academy of Sciences, 107*, 14042-14047. doi:10.1073/pnas.1002815107
- Sciarra, D. T., and Seirup, H. J. (2008). The multidimensionality of school engagement and math achievement among racial groups. *Professional School Counseling, 11*, 218-228. doi:10.5330/PSC.n.2010-11.218
- Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist, 46*(1), 26-47. doi:10.1080/00461520.2011.538646
- Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a

- key to buffering against identity threat. *Personality and Social Psychology Bulletin*, 39(5), 663-676. doi:10.1177/0146167213480816
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., . . . Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, 104(4), 591-618. doi:10.1037/a0031495
- Simmons, C. M. (2013). *Reducing stereotype threat in academically at-risk Africa American students: A self-affirmation intervention*. (2013-99010-209).
- Smith, A. H., & Bates, M. N. (1992). Confidence limit analyses should replace power calculations in the interpretation of epidemiologic studies. *Epidemiology*, 449-452. doi:10.1097/00001648-199209000-00011
- Spencer, B., & Castano, E. (2007). Social class is dead. Long live social class! Stereotype threat among low socioeconomic status individuals. *Social Justice Research*, 20(4), 418-432. doi:10.1007/s11211-007-0047-7
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4-28. doi:10.1006/jesp.1998.1373
- Steele, C. M., & Aronson, J. (2000). Stereotype threat and the intellectual test performance of African Americans. In C. Stangor, & C. (Stangor (Eds.), (pp. 369-389). New York, NY, US: Psychology Press.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811. doi:10.1037/0022-3514.69.5.797
- Stipek, D., & Gralinski, J. H. (1996). Children's beliefs about intelligence and school performance. *Journal of Educational Psychology*, 88(3), 397-407. doi:10.1037/0022-0663.88.3.397
- Tyson, K. 2011. *Integration interrupted: Tracking, Black students, and acting White after Brown*. New York: Oxford University Press.
- Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215-247. doi:10.3102/1076998609346961
- Valentine, J., DuBois, D., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2), 111-133. doi:10.1207/s15326985ep3902_3

- Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2009) Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12-23.
doi:10.1016/j.learninstruc.2013.04.002
- Wilkins, P. (2014). Efficacy of a growth mindset intervention to increase student achievement. *Education Theses, Dissertations and Projects*. Paper 24.
- Wigfield A., Eccles J.S., MacIver D., Reuman, D., & Midgley C. (1991). Transitions at adolescence: changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, 27, 552–565. doi:10.1037/0012-1649.27.4.552
- Wynne, C. B. (2012). *When hindsight isn't twenty-twenty: Unsuccessful and unresolved after the implementation of a promising self-affirmation intervention*. (2012-99150-217).
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143(2), 804. doi:10.1037/a0033906
- 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(3), 374. doi:10.1037/edu0000098

Table 1.1 List of excluded studies.

Author(s)	Type of Publication	Reason for Exclusion
1. Aronson et al. (2009)	White paper	Summary of several published studies already included
2. Bancroft et al. (2017)	Peer-reviewed article	Data from a published study already included (substantive findings were not different)
3. Barron et al. (2015)	Conference paper	Not an academic intervention
4. Borman et al. (2015)	Conference paper	Data from a published study already included
5. Casad et al. (2017)	Peer-reviewed article	Not a long term achievement outcome
6. Chao et al. (2017)	Peer-reviewed article	Non-American population
7. Cohen et al. (2006)	Peer-reviewed article	Data included in follow-up analyses from Cohen et al. (2009)
8. Cutts et al. (2010)	Peer-reviewed article	Non-American population
9. Ganley et al. (2013)	Peer-reviewed article	Not a long term achievement outcome
10. Hanselman et al. (2014)	Peer-reviewed article	Data from a published study already included (substantive findings were different)
11. Mousseau (2013)	Dissertation	Not a long term achievement outcome
12. Murphy & Zirkel (2015)	Peer-reviewed article	Observational study
13. Rozek et al. (2015)	Conference paper	Data from a published study already included
14. Saunders (2013)	Dissertation	Did not use random assignment
15. Schmidt et al. (2015)	Peer-reviewed article	Did not use random assignment
16. Schmidt et al. (2016)	Peer-reviewed article	Not a long term achievement outcome
17. Walton & Cohen (2007)	Peer-reviewed article	Undergraduate sample

Table 1.2 Description of stereotype threat studies.

Citation	Participants	Description of Participants	Description of School(s)	Intervention Description	Study Length
1. Borman et al., 2016	1,048 middle school students	Varied across schools, on average, potentially threatened students: 37%	11 schools Varied across schools, on average 43% FRL	Self-affirming paragraph Four brief reflections in language arts or homeroom, once before an assessment, presented as a regular class assignment	1 school year
2. Bowen et al., 2013	132 middle school students	Black: 81% White/Other: 13.9%	1 school Low income Low performing	Self-affirming paragraph One brief reflection, two days before end of grading period in fall semester Presented in homeroom, students told it would not be graded	1 school year
3. Bratter et al., 2016	886 high school students	Black: 25% Hispanic: >60% White: 11%	3 schools Two low performing One high performing	Self-affirming paragraph Four brief reflections, all before standardized assessments Completed in English class, presented as a regular class assignment	1 school year

4. Cohen et al., 2009 ²	385 middle school students	Black: 51% White: 49%	1 school Suburban Middle to lower class families	Self-affirming paragraph One brief reflection, beginning of the school year, presented as a regular class assignment in target course 2 school years
5. Dee, 2015	2,564 middle school students	Black: 5-65% Hispanic: 16-36% White: 17-59%	6 schools Three low performing One high performing Two average	Self-affirming paragraph Two brief reflections, once at the beginning of the year and once 6-8 weeks later 1 school year Completed before an assessment, when possible
6. Good et al., 2003	138 middle school students	Black: 13% Hispanic: 67% White: 20%	1 school Rural Low income 70% free or reduced lunch program participation	Mentoring and Lessons Two 90-minute in person sessions and weekly e-mail correspondence and web site development with undergraduate mentors related to growth mindset, identity affirmation or both 1 school year
7. Hanselman et al., 2017	1,170 middle school students	Varied across schools, on average,	11 schools	Self-affirming paragraph Three to four sessions throughout the first school year, at the beginning of the year 2 school years

<p>potentially threatened students: 38%</p>	<p>Varied across schools, on average 46% FRL</p>	<p>and before assessments, presented as an exercise the school does a few times a year</p>
<p>8. Protzko & Aronson, 2016</p>	<p>454 High school students</p>	<p>School 1: 96% Black/Hispanic School 2: 26% Black/Hispanic</p> <p>2 schools</p> <p>One low income, urban school One wealthy, suburban school</p> <p>Self-affirming paragraph</p> <p>One session in the first three weeks of school, students told it was an outside research team</p> <p>1 semester</p>
<p>9. Sherman et al., 2013 – Study 1</p>	<p>199 middle school students</p>	<p>Hispanic: 41% White: 52%</p> <p>1 school</p> <p>Self-affirming paragraph</p> <p>Four to five sessions throughout the first school year, presented as an exercise the school does a few times a year, at the beginning of the year and before assessments</p> <p>3 school years</p>
<p>10. Sherman et al., 2013 – Study 2</p>	<p>185 middle school students</p>	<p>Hispanic: 37% White: 62%</p> <p>1 school</p> <p>Low income 50% free or reduced lunch program participation</p> <p>Self-affirming paragraph</p> <p>Two sessions near beginning and end of school year, presented as a regular class activity</p> <p>1 school year</p>

11. Shnabel et al., 2013	355 middle school students	Black: 48% White: 52%	1 school Middle to lower middle class students	Self-affirming paragraph Two to three sessions at the beginning of the year	1 grading period
12. Simmons, 2013¹	47 high school students	Black: 100%	3 schools Mostly middle or working class families	Self-affirming paragraph One reflection, middle of the fall semester, students were moved to cafeteria or separate classroom, they were also compensated	1 semester
13. Wynne, 2011¹	100 middle school students	Black: 71% Latino: 29%	1 school High performing Low income 74% FRL	Self-affirming paragraph Three to five sessions throughout school year	1 school year

Note: ¹ = dissertation, ² = data from Cohen et al. (2009) includes data from Cohen et al. (2006) that is compiled across cohorts for 2 years

Table 1.3. Description of growth mindset studies.

Citation	Participants	Description of Participants	Description of School(s)	Intervention Description	Study Length
1. Blackwell et al., 2007	91 middle school students	Black: 52% Hispanic: 45% White/Asian: 3%	1 school Average math achievement in 35 th percentile 79% free or reduced lunch program participation	Growth mindset lessons Eight sessions on incremental nature of intelligence and the ability to “grow your brain”	1 school year
2. Brougham, 2016 ¹	84 high school students	Black: 65% Hispanic: 1% White: 26%	2 schools Urban Low income High performing charter school Low performing traditional school	Growth mindset lessons Three 45-minute sessions where students wrote written summaries of informational materials and participated in discussions about the brain and learning	1 semester
3. Burnette et al., 2017	222 high school students	Black: 25% Hispanic: 29% White: 38% Female: 100%	4 schools Rural Low income	Growth mindset lessons One session, with three modules on the incremental nature of intelligence and the ability to “grow your brain”	1 school year
4. Eskreis-Winkler et al., 2016 – Study 1	427 middle school students	Black: 5% Hispanic: 8% White: 65%	3 schools 7% FRL	Deliberate practice One session, emphasizing deliberate practice is more important than talent for improvement and success	1 grading quarter
5. Eskreis-Winkler et al., 2016 – Study 2	248 middle school students	White: 65% Black: 14% Hispanic: 10%	1 school 12% FRL	Deliberate practice Two sessions, emphasizing deliberate practice is more important than talent for improvement and success	2 grading quarters

6. Good et al., 2003	138 middle school students	Black: 13% Hispanic: 67% White: 20%	1 school Rural Low income 70% free or reduced lunch program participation	Mentoring and Lessons Two 90-minute in person sessions and weekly e-mail correspondence and web site development with undergraduate mentors related to growth mindset, identity affirmation or both	1 school year
7. Oyserman et al., 2006	228 middle school students	Black: 72% Hispanic: 17% White: 11%	3 schools Urban Low income 67% FRL	Academic possible selves Eleven sessions that taught students how to make meaning of difficulty and set academic goals and strategies. Emphasized self-regulatory behavior	2 schools years
8. Paunesku et al., 2015	1,594 high school students	Black: 11% Hispanic: 33% White: 23%	11 schools On average, ~43% FRL	Growth mindset lesson One 45-minute session	1 semester
9. Wilkins, 2014¹	1,012 middle school students	Black: 25% Hispanic: 23% White: 46%	7 schools 50% FRL	Growth mindset lessons Brainology curriculum – twelve online modules about the growth mindset	1 school year
10. Yeager et al., 2014 – Study 3	76 high school students	Black: 66% White: 24%	1 school Urban Low income 21% of participants lived in home below poverty line	Wise feedback One online session in attribution retraining to help students attribute teachers’ critical feedback to teachers’ high expectations and belief in their ability	1 semester
11. Yeager et al., 2016 – Study 2	3,676 high school students	Black: 17% Hispanic: 29% White: 30%	10 schools Within medium range for poverty indicators (i.e. FRL)	Growth mindset intervention Two one-period online sessions about the growth mindset	1 semester

Note: ¹ = dissertation

Table 1.4. Effects of stereotype intervention.

Citation	Group(s)	Outcome(s)	Effect size (within Year 1)	Effect size (end of Year 2)	Method
1. Borman et al., 2016	Black and Hispanic students	GPA	.09*		Reported
		Standardized test scores English Math	-.03 ³ .06 ³	.00 .08	
2. Bowen et al., 2013	Black and White students Black students	Course grades Social studies Social studies	.57*³ .33 ³		Reported and Calculated
		Standardized test scores English English	-.28 ³ -.10 ³		
3. Bratter et al., 2016	Black students Hispanic students	Math Math	-.31 ³ -.03 ³		Calculated
		Course grades English English	.157 ³ -.032 ³		
4. Cohen et al., 2009²	Black students	GPA	.28**	.38***	Calculated
5. Dee, 2015	Black students Hispanic students	Course grades Social studies Social studies	-.001 ³ .025 ³		Calculated

6.	Hanselman et al., 2017	Black and Hispanic students	GPA	-.002	-.07	Reported
			Standardized test scores			
			Math	-.09 ³	-.08	
			English	-.005 ³	-.005	
7.	Protzko & Aronson, 2016	Black and Hispanic students	GPA	.14		Calculated
8.	Sherman et al., 2013 – Study 1	Hispanic students	GPA	.29**	.43**	Reported
9.	Sherman et al., 2013 – Study 2	Hispanic students	GPA	.45* ³ ; .56**		Reported
10.	Shnabel et al., 2013	Black students	GPA	.38***		Reported
11.	Simmons, 2013 ¹	Black students	GPA	-.255		Calculated
12.	Wynne, 2011 ¹	Black and Hispanic students	GPA	.13		Calculated
			Standardized test scores			
			English	.28 ³		
			Math	.45* ³		

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation, ² = data from Cohen et al. (2006) is included in these analyses, which combine Year 1 and Year 2 data across both studies, ³ = data was not included in final meta-analysis

Table 1.5. Effects of growth mindset intervention.

Citation	Group(s)	Outcome(s)	Effect Size (within Year 1)	Effect Size (end of Year 2)	Method
1. Blackwell et al., 2007	Whole sample	Course grades Math	.62**3		Calculated
2. Brougham, 2016 ¹	Whole sample Magnet school Traditional school	GPA GPA GPA	-.68* .54 ³ -.96**3		Calculated
3. Burnette et al., 2017	Whole sample	GPA	.06		Calculated
4. Eskreis-Winkler et al., 2016 – Study 4	Whole sample Low performers	GPA GPA	.21* .23**3		Calculated
5. Eskreis-Winkler et al., 2016 – Study 5	Whole sample Low performers	GPA GPA	.19 .21 ³		Calculated
6. Good et al., 2003	Females Males Whole sample	Standardized test scores Math Math Reading	1.31**3 .62**3 .51**3		Reported
7. Oyserman et al., 2006	Whole sample Whole sample	GPA Standardized test scores Proportion of tests passed	.25* .36**3	.30*	Reported

8. Paunesku et al., 2015	Low performers	GPA	.09 ^{*3}	Calculated
9. Wilkins, 2014 ¹	Whole sample Whole sample	Course grades Math Science	.02 .26 [*]	Calculated
10. Yeager et al., 2014 – Study 3	Whole sample Black students White students	GPA GPA GPA	.52 [*] .73 ^{*3} .34 ³	Reported but calculated ²
11. Yeager et al., 2016 – Study 2	Whole sample Low performers High performers	GPA GPA GPA	.09 ^{**} .10 ^{**3} -.03 ³	Reported

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation, ² = reported effect sizes were unable to be used and calculated effect sizes differed from reported effect sizes, ³ = data was not included in final meta-analysis

Table 1.6. Meta-analysis of stereotype threat intervention on GPA of threatened students.

Study	Time	Statistics Used	<i>d</i>	SE	Variance	CI Lower	CI Upper	Z	<i>p</i>
1. Borman et al., 2016	End of Year 1	Cohen's <i>d</i> , group sample sizes	0.09	0.06	0.00	-0.03	0.21	1.43	0.15
2. Cohen et al., 2009	End of Year 1	<i>t</i> , total sample size	0.28*	0.10	0.01	0.08	0.48	2.70	0.01
3. Hanselman et al., 2017	End of Year 1	Cohen's <i>d</i> , <i>SE</i>	0.00	0.04	0.00	-0.09	0.08	-0.05	0.96
4. Protzko & Aronson	End of Semester	<i>t</i> , total sample size	0.14	0.09	0.01	-0.04	0.33	1.48	0.14
5. Sherman et al., 2013 - Study 1	End of Year 1	Cohen's <i>d</i> , group sample sizes	0.29*	0.15	0.02	0.00	0.58	1.96	0.05
6. Sherman et al., 2013 - Study 2	End of Year 1	Cohen's <i>d</i> , group sample sizes	0.56***	0.17	0.03	0.23	0.89	3.37	<0.001
7. Shnabel et al. 2013	End of Year 1	<i>t</i> , total sample size	0.39***	0.11	0.01	.175	0.60	3.58	<0.001
8. Simmons 2013 ¹	End of Semester	Group means, group sample sizes, group <i>SDs</i>	-0.26	0.29	0.09	-0.83	0.32	-0.87	0.38
9. Wynne 2011 ¹	End of Year 1	<i>t</i> , total sample size	0.13	0.20	0.04	-0.27	0.53	0.66	0.51
Point Estimate			0.19**	0.06	0.01	0.06	0.31	2.97	0.01

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation

Table 1.7. Meta-analysis of growth mindset intervention on GPA of all students.

Study	Time	Statistics Used	<i>d</i>	SE	Variance	CI Lower	CI Upper	Z	<i>p</i>
1. Brougham, 2016 ¹	End of Semester	Group means, group sample sizes, group <i>SDs</i>	-0.68*	0.25	0.06	-1.18	-0.19	-2.72	0.01
2. Burnette et al., 2017	End of Year 1	<i>t</i> , total sample size	0.06	0.14	0.02	-0.20	0.33	0.47	0.64
3. Eskreis-Winkler et al., 2016 - Study 4	End of Year 1	<i>t</i> , total sample size	0.21*	0.10	0.01	0.02	0.41	2.19	0.03
4. Eskreis-Winkler et al., 2016 - Study 5	End of Year 1	<i>t</i> , total sample size	0.19	0.13	0.02	-0.06	0.44	1.46	0.14
5. Oyserman et al., 2003	End of Year 1	Cohen's <i>d</i> , variance	0.25	0.93	0.87	-1.58	2.08	0.27	0.79
6. Yeager et al., 2014 - Study 3	End of Semester	<i>t</i> , total sample size	0.53*	0.25	0.06	0.14	1.10	2.52	0.01
7. Yeager et al., 2016 - Study 2	End of Semester	<i>t</i> , total sample size	0.09*	0.03	0.00	0.02	0.15	2.66	0.01
Point Estimate			0.11	0.08	0.01	-0.05	0.27	1.34	0.18

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation

Table 1.8. One study removed sensitivity test for stereotype threat meta-analysis.

Study	<i>d</i>	SE	Variance	CI Lower	CI Upper	Z	<i>p</i>
1. Borman et al., 2016	.21**	.08	.01	.05	.37	2.64	<0.01
2. Cohen et al., 2009	.18**	.07	.00	.04	.31	2.56	<0.01
3. Hanselman et al., 2017	.23***	.06	.00	.10	.35	3.61	<0.001
4. Protzko & Aronson	.20*	.07	.01	.05	.34	2.71	<0.01
5. Sherman et al., 2013 - Study 1	.15*	.06	.00	.04	.27	2.61	<0.01
6. Sherman et al., 2013 - Study 2	.18*	.07	.00	.05	.31	2.64	<0.01
7. Shnabel et al. 2013	.16*	.06	.00	.04	.28	2.53	<0.01
8. Simmons 2013 ¹	.21***	.06	.00	.08	.33	3.19	<0.001
9. Wynne 2011 ¹	.19***	.07	.00	.06	.33	2.86	<0.001

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation

Table 1.9. One study removed sensitivity test for growth mindset meta-analysis.

	Study Removed	<i>d</i>	SE	Variance	CI Lower	CI Upper	Z	<i>p</i>
1.	Brougham, 2016 ¹	.14**	.05	.01	.05	0.24	2.88	<.01
2.	Burnette et al., 2017	.12	.10	.01	-.08	0.32	1.17	.24
3.	Eskreis-Winkler et al., 2016 - Study 4	.08	.11	.01	-.13	0.29	0.75	.45
4.	Eskreis-Winkler et al., 2016 - Study 5	.09	.10	.01	-.11	0.29	.91	.36
5.	Oyserman et al., 2003	.11	.09	.01	-.06	0.28	1.27	.20
6.	Yeager et al., 2014 – Study 3	.07	.08	.01	-.08	0.23	.97	.33
7.	Yeager et al., 2016 – Study 2	.11	.13	.02	-.15	0.36	.83	.41

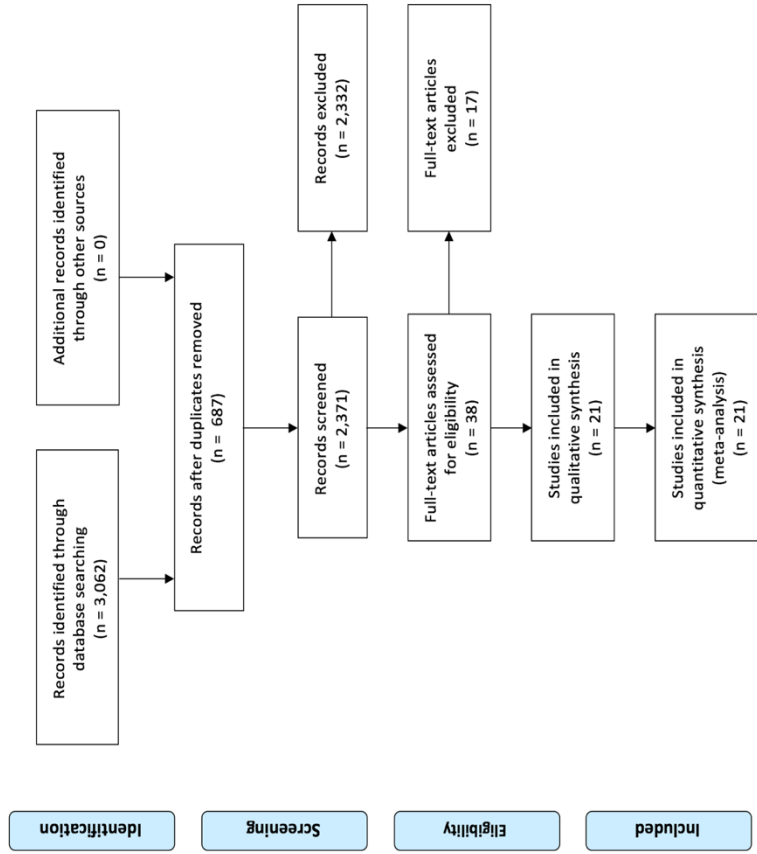
Note: * = $p < .05$, ** = $p < .01$, *** = $p < .00$, ¹ = dissertation

Table 1.10. One study removed sensitivity test for growth mindset meta-analysis, with Brougham (2016) removed.

	Study Removed	<i>d</i>	SE	Variance	CI Lower	CI Upper	Z	<i>p</i>
1.	Burnette et al., 2017	0.17**	0.06	0.00	0.04	0.30	2.65	<0.01
2.	Eskreis-Winkler et al., 2016 - Study 4	0.13*	0.06	0.00	0.01	0.25	2.18	0.03
3.	Eskreis-Winkler et al., 2016 - Study 5	0.15*	0.06	0.00	0.02	0.27	2.34	0.02
4.	Oyserman et al., 2003	0.15*	0.06	0.00	0.04	0.26	2.67	<0.01
5.	Yeager et al., 2014 – Study 3	0.10***	0.03	0.00	0.05	0.16	3.50	<.001
6.	Yeager et al., 2016 – Study 2	0.20**	0.06	0.00	0.07	0.33	3.11	<0.01

Note: * = $p < .05$, ** = $p < .01$, *** = $p < .001$, ¹ = dissertation

Figure 1.1. Identification of eligible studies.



Note: Flow diagram from Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group (2009)

Figure 1.2. Risk of bias assessment across stereotype threat studies.

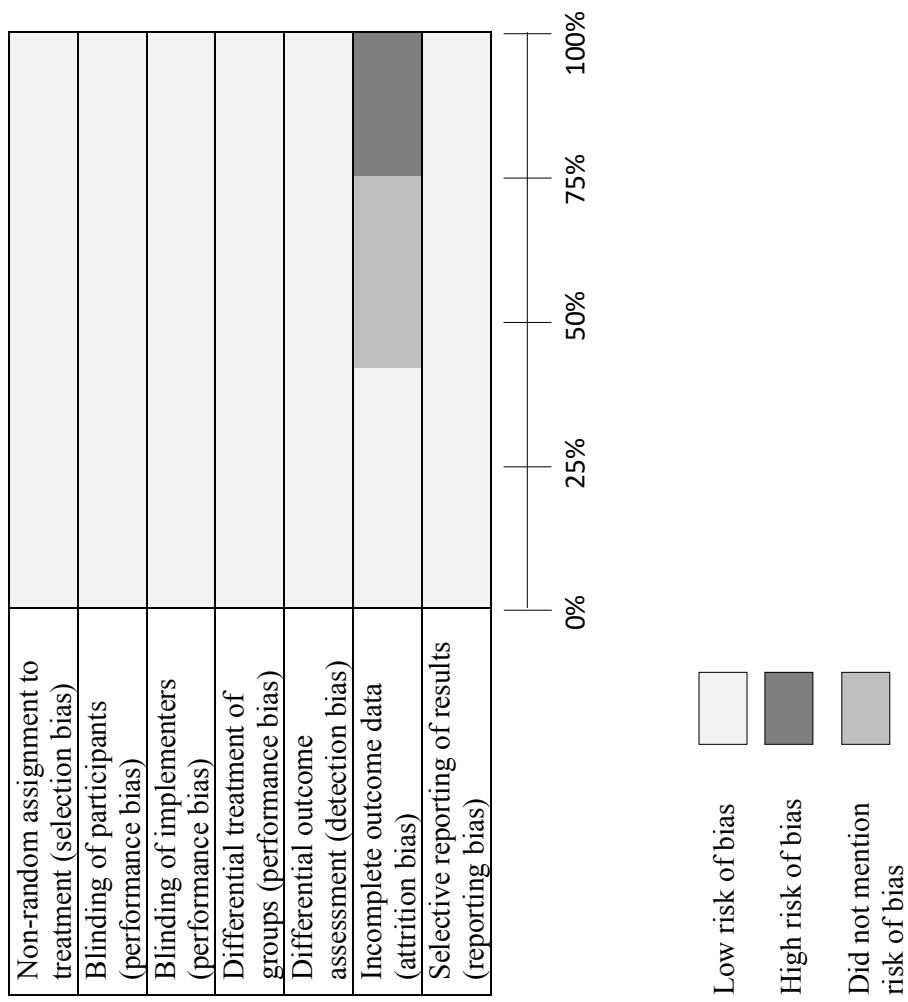


Figure 1.3. Risk of bias assessment across growth mindset studies.

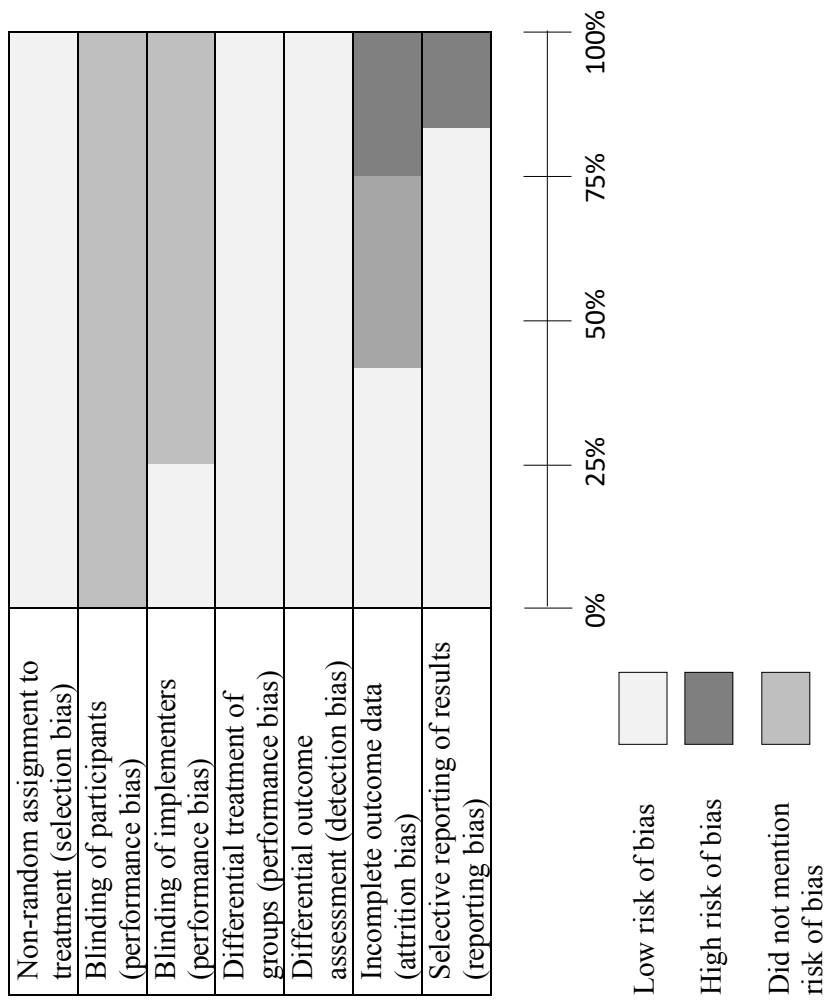


Figure 1.4. Summary of outcomes, populations and samples for stereotype threat studies.

STUDY	GPA	COURSE GRADE(S)	TEST SCORES	HIGH SCHOOL	MIDDLE SCHOOL	SAMPLE: THREATENED STUDENTS ONLY
Borman et al. (2016)						
Bowen et al. (2013)						
Bratter et al., (2016)						
Cohen et al. (2009)						
Dee (2015)						
Good et al. (2003)						
Hanselman et al. (2017)						
Protzko & Aronson (2016)						
Sherman et al. (2013) – Study 1						
Sherman et al. (2013) – Study 2						
Shnabel et al. (2012)						
Simmons (2013) ¹						
Wynne (2011) ¹						
Total	9	4	3	3	10	5

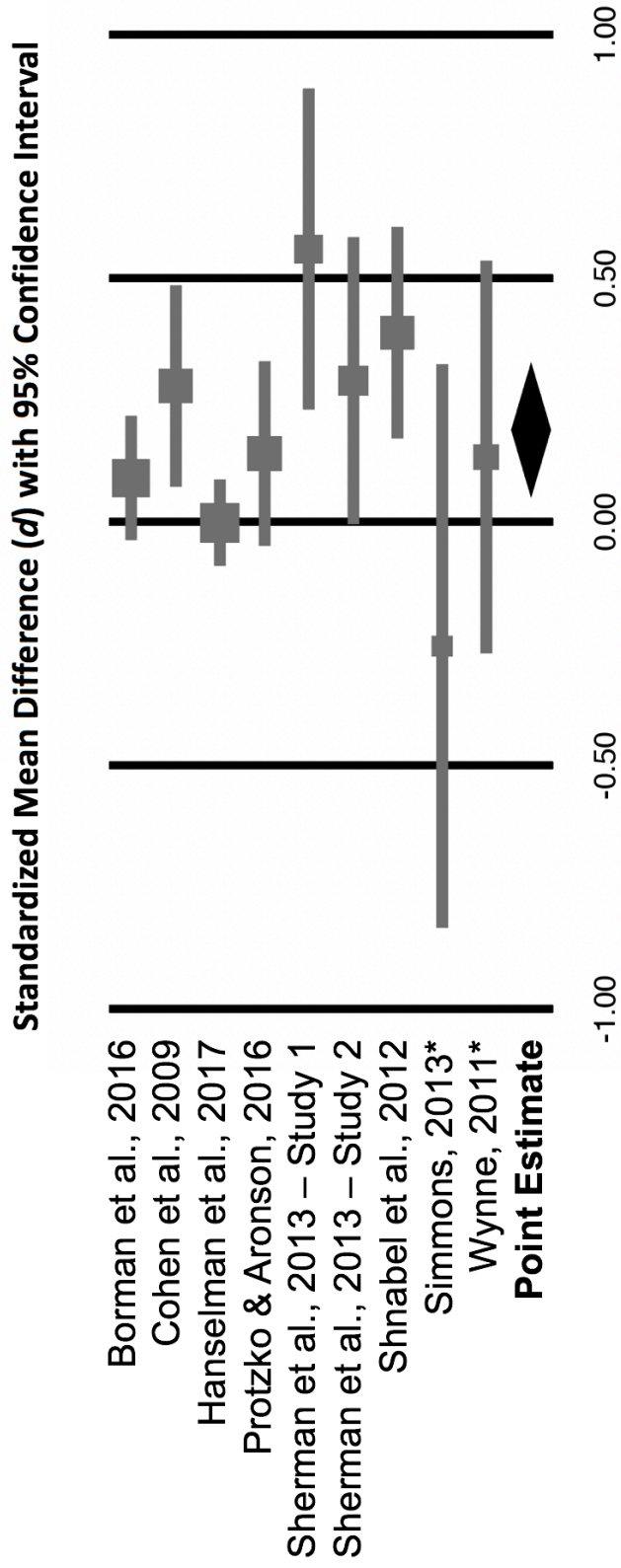
Note: ¹ = dissertation

Figure 1.5. Summary of outcomes, populations and samples for growth mindset studies.

GROWTH MINDSET STUDIES	GPA	COURSE GRADE(S)	TEST SCORES	HIGH SCHOOL	MIDDLE SCHOOL
Blackwell et al., (2007)					
Brougham (2016) ¹					
Burnette et al. (2017)					
Ekreis-Winkler et al. (2016) – Study 1					
Ekreis-Winkler et al. (2016) – Study 2					
Good et al. (2003)					
Oyserman et al. (2006)					
Paunesku et al. (2015) ²					
Wilkins (2015) ¹					
Yeager et al. (2014)					
Yeager et al. (2016)					
Total	7	2	2	5	6

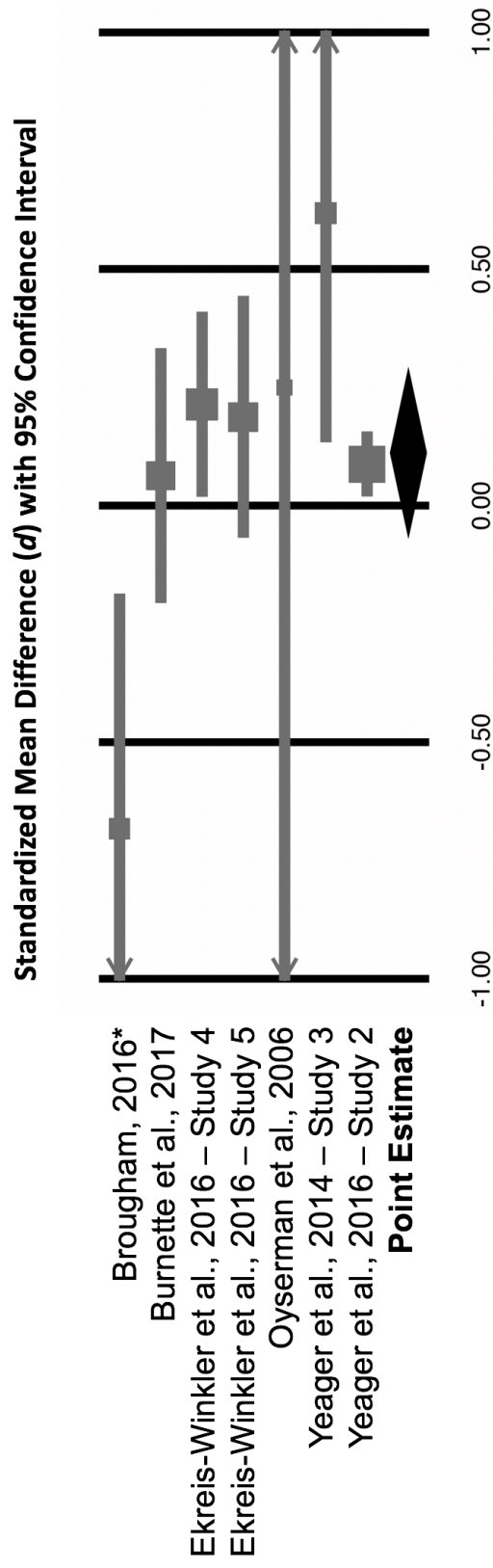
Note: ¹ = dissertation, ² = this study used GPA as an outcome, but was not included in the meta-analysis because it provided data on low performing students only, not the whole sample of students

Figure 1.6. Forest plot for meta-analysis of stereotype threat intervention on GPA.



Note: * = dissertation

Figure 1.7. Forest plot for meta-analysis of growth mindset intervention on GPA.



Note: * = dissertation

PAPER II

MINDSETS AND STEREOTYPES OF INTELLIGENCE: TEACHER BEHAVIORS THAT MOTIVATE STUDENTS TO ACHIEVE

Nationally, an academic achievement gap persists between underprivileged students and their more advantaged peers (Dee & Jacob, 2011; Kena et al., 2016; Vanneman, Hamilton, Anderson, & Rahman, 2011). Historically, low income students and students of color – specifically, those who identify as black or Hispanic – have been labeled “at-risk” for academic underachievement. It is often difficult to extract the intersectionality of race, ethnicity and poverty because the lowest performing schools are segregated to the areas of highest poverty and are predominantly attended by students of color (Williams, 2014). Achievement gaps widen as students matriculate through secondary education, leading to high dropout rates and low rates of postsecondary degree acquisition among low income students and students of color (NAEP, 2015; Cahalan & Perna, 2015; Kena et al., 2016).

As a result, so called “at-risk” youth are common targets for academic intervention, commonly categorized as either cognitive interventions that improve students’ content knowledge (e.g., curriculum reform, test taking strategies) and non-cognitive interventions that address social and emotional determinants of learning (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). Especially for students who are vulnerable to low achievement, non-cognitive skills such as persistence and self-control have been identified as effective leverage points to promote school success (Heckman & Rubinstein, 2001). Not only are non-cognitive skills stronger predictors of academic performance than cognitive ability, but they are also highly malleable (Duckworth, Quinn, & Tsukayama, 2011a).

Promoting a Growth Mindset

Because non-cognitive skills are essential for academic success, it is important to determine how they are developed. Evidence suggests that students' persistence in the face of challenge may be related to underlying beliefs about the nature of intelligence and whether they see effort as a mechanism for increasing ability (Duckworth & Eskreis-Winkler, 2013). Dweck's (1999; 2006) work on implicit theories of intelligence defines two beliefs that students may possess regarding ability: (a) a "growth mindset" that views intelligence as malleable and gained through practice or (b) a "fixed mindset" that sees intelligence as an inherent, stable trait that cannot be modified, regardless of effort.

Whereas students with a growth mindset seek out challenges and focus on improvement, children with a fixed mindset avoid challenges and focus on appearing competent (Dweck & Leggett, 1988; Mueller & Dweck, 1998). These differences are most apparent in the response to failure: Students with a growth mindset frame mistakes as part of the learning process, while those with a fixed mindset perceive failure as an indicator of irreparably low ability (Dweck, 2006). Students with fixed mindsets are less likely to persist in the face of a difficult challenge, particularly after receiving negative feedback about their performance – a common occurrence in academic settings (Cimpian, Arce, Markman, & Dweck; Elliot & Dweck, 1988). Over time, mindset beliefs drive differences in knowledge and skill acquisition. Data from longitudinal randomized trials show that children who learn about the ability to "grow their brains" through effort show a better trajectory of academic achievement than fixed mindset peers of equal ability, and the effects are largest among low performers and students of color (Blackwell, Trzesniewski & Dweck, 2007; Good, Aronson

& Inzlicht, 2003; Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015; Yeager et al., 2014; Yeager et al. 2016).

Reducing Stereotype Threat

Both fixed and growth mindset beliefs are found among children of all backgrounds, but low income students and students of color face an additional barrier to developing a growth mindset in that stereotypes of intelligence suggest people of certain identities are inherently less intelligent than others. Although stereotypes are conceptually distinct from mindsets of intelligence, stereotypes about intelligence represent a specific type of fixed mindset beliefs (e.g., that students of color are not as smart as white students and there is little that can be done about it; Steele & Aronson, 1995; Good, Aronson, & Inzlicht, 2002). Stereotypes that link ability and intelligence can produce the phenomenon of stereotype threat, where an individual's awareness of a negative stereotype about their identity creates a cognitive burden that can result in underperformance (Steele, 1992; Steele and Aronson, 1995). Stereotype threat can influence students' academic achievement in several important ways. First, students of stereotyped identities may subscribe to negative beliefs about their ability, which removes high achievement from their academic self-concepts and creates low expectations for success (Aronson, Blanton, & Cooper, 1995; Pelham & Swann, 1989; Steele, 1992). Second, students under threat may attempt to compensate for or disprove stereotypes of their ability, creating stress and anxiety that, paradoxically, can lead to underperformance (Aronson, 1999; Aronson, Quinn, & Spencer, 1998; Blascovich, Spencer, Quinn, & Steele, 2001; Steele & Aronson, 1995). Finally, stereotype threat can create ongoing interference with the learning process that cognitively burdens students of

stereotyped identities, impeding content acquisition over time (Mangels, Good, Whiteman, Maniscalco, & Dweck, 2012; Rydell, Shiffrin, Boucher, Van Loo, & Rydell, 2010).

Like growth mindset beliefs, stereotype endorsement is malleable, and threat can be reduced by reaffirming equality across identities. For example, before administering an assessment, stating that students of all identities perform equally well on the exam can eliminate performance gaps (Steele & Aronson, 1995; Steele & Aronson, 2000). In contrast, other threat-reducing interventions do not aim to directly counter stereotypes but to reduce potential effects of stereotype threat in the classroom. One threat-reducing intervention, developed by Cohen and colleagues (2006), facilitates students' identity affirmation through a brief written reflection on personal values. Though students only engage in the writing activity a few times throughout the year, these reflections appear to generate a "cascade" of positive school experiences that leads to long term academic improvements, reducing achievement gaps by nearly half (Cohen, Garcia, Apfel, & Master, 2006). Self-affirming reflections have been successfully replicated with black, Hispanic (Cohen, Garcia, Purdue-Vaughns, Apfel, & Brzustoski, 2009; Sherman et al. 2013) and low income students (Bowen, Wegmann & Webber, 2013), with similar results. Subsequent large, multi-school replications have also demonstrated that effects are most profound in contexts where the potential for stereotype threat is highest – supporting the underlying link between students' identities and their academic achievement (Hanselman, Bruch, Gamoran, & Borman, 2014).

Increasing Student Motivation: Self-Determination Theory

As academic interventions, teaching students about the incremental nature of intelligence and reducing stereotype threat accomplish similar goals in that they promote students' sense of competence, which is a requisite for motivated behavior as defined by

Self-determination theory (SDT; Deci and Ryan, 1985; Ryan and Deci, 2000). Deci and Ryan (1985) describe that, in addition to competence, a students' sense of autonomy and relatedness are critical in creating intrinsically motivated behavior (e.g., engagement in school). In the school context, autonomy relates to students' level of choice and control over classroom tasks and assignments, while relatedness involves the relationships that form among teachers and students (Deci, Vallerand, Pelletier, & Ryan 1991). Studies have demonstrated that students of teachers who provide autonomy and develop relatedness show greater interest and motivation for their coursework, promoting positive engagement that is correlated with increases in academic performance (Deci et al., 1981; Grolnick and Ryan, 1987).

More directly, however, beliefs about the nature of intelligence are linked with motivation and performance through students' perceived competence (Deci, Vallerand, Pelliter, & Ryan, 1991; Ryan 1982). If students believe that intelligence is a product of effort, they may be more likely to feel that they are competent to master a difficult task. Path analysis conducted by Vallerand and Reid (1984; 1988) found that students' perceptions of competence mediated the relationship between teachers' positive feedback and students' intrinsic motivation. Another study by Law and colleagues (2012) showed that if students believe that they have fixed levels of intelligence, failure is a strong threat to competence and students may attempt to maintain competence through non-intrinsically motivated behaviors (i.e., cheating instead of studying; Law, Elliot and Murayama, 2012). In contrast, one benefit of believing that intelligence is cultivated through effort is that feelings of competence can be maintained despite failure: By shifting the focus to the importance of effort, students feel more competent and efficacious to produce desired outcomes (Dweck, 2006). Self-

determination theory would suggest that intelligence beliefs that promote students' sense of competence help them maintain persistence and effort, even when they have yet to experience the reinforcement of success (Elliot & Dweck, 1988; Cimpian et al., 2007). Indeed, students' perceived competence predicts their challenge-seeking behavior (e.g., sticking with a difficult challenge rather than seeking out easier options; Boggiano, Main, & Katz, 1988). As students internalize that they can achieve academically because their performance can increase with effort and is not predetermined by their belonging to a negatively stereotyped group, they feel more competent and that academic success is possible (Niemiec & Ryan, 2009).

Though intelligence beliefs are closely related to students' sense of competence, Deci and Ryan argue the three domains of self-determination rarely work in isolation, with any one domain affecting the others (Deci et al., 1991). For example, autonomy-supportive teachers who provide students with a level of choice and control in the classroom can be simultaneously communicating that they trust that students are competent to handle such freedoms – and, as a result, may develop better relationships with their students. After conducting detailed interviews with students in underperforming, underresourced schools, Chhuon and Wallace (2014) describe an “ordinary magic” as relatedness, autonomy and competence work in concert. Participating students emphasized that they feel more motivated with teachers who (a) avoid making negative assumptions about students' capabilities or circumstances regardless of prior performance, (b) make students “feel known” as individuals, and (c) provide specific instrumental support that reflects their understanding of students as individuals (Chhuon and Wallace, 2014).

Teachers as Intervention Agents

Teachers have an ongoing and powerful influence on students' intelligence beliefs (Dweck & Bush, 1976; Dweck, Davidson, Nelson, & Enna, 1978; Henderlong & Lepper, 2002; Mueller & Dweck, 1988; Rattan, Good, & Dweck, 2012). Findings from randomized studies show that praising children for their effort (“you worked hard to get good grades” or “with effort, you got better at math”) rather than their natural ability supports the development of a growth mindset and inoculates against feelings of helplessness that cause students to disengage from challenging tasks (Davis, Burnette, Allison, & Stone, 2011; Dinger, Dickhauser, Spinath, & Steinmayr, 2013; Henderlong & Lepper, 2002). Conversely, when students are given ability-focused praise (“you’re so smart” or “you’re a math person”), they respond negatively when they experience failure, they seek out opportunities to highlight existing mastery and they avoid future opportunities for learning (Dweck & Leggett, 1988; Mueller & Dweck, 1998).

Importantly, teachers often unknowingly resort to giving ability-based feedback to struggling students because it can be perceived as kind or comforting. Rattan and colleagues (2012) identify that “consoling” is a common way that teachers unwittingly support fixed mindset in their students, by saying things like “it’s okay – not everyone is a math person” or by suggesting unhelpful strategies such as dropping a difficult class. Teachers who were randomized to learn about growth mindset theory were less likely to endorse the use of these consoling behaviors (Rattan et al., 2012). The same study found that when students were randomized to a scenario where their teacher gave growth mindset feedback in response to a failed exam, students felt more supported and were more likely to attempt improvement than their peers who received consoling feedback. Together, these studies suggest that when students are made to feel that intelligence is an inherent trait, if they believe that they do not

possess the skills for a challenging task, further effort is futile and disengagement is a logical coping strategy (Dweck & Bush, 1976; Kamins & Dweck, 1999).

In addition to the critical role that teachers play in shaping students' beliefs about intelligence, they are ideal targets to implement academic interventions because evidence from school-based, non-cognitive programs indicates that effectiveness is bolstered by implementing the intervention in the same context where the effects are desired (Hawkins, von Cleve, & Catalano, 2001). Multilevel models from other classroom-based research in the area of achievement motivation have shown that between 5 and 35% of variation in student outcomes is attributable to classroom-level effects, highlighting the potential for teachers to affect youth outcomes (Meece, Anderman, & Anderman, 2006). Yet, despite the growing evidence base supporting the benefits of focusing on students' beliefs about intelligence, teachers have had very limited involvement in intervention implementation, and little empirical work has been dedicated to translating growth mindset and stereotype threat research into the classroom (Edwards, Esmonde, Wagner, & Beattie, 2017; Sparks, 2013). A review of major achievement motivation theories and their adaptation to the classroom revealed that well-evidenced theories were commonly suggested to teachers as useful strategies, with few supporting resources to help teachers integrate theory into practice (Turner 2010; Urdan & Turner, 2005). By nature, teaching requires responding to students' successes, struggles and failures (Henderlong & Lepper, 2002). Without support, teachers may rely on ineffective and potentially damaging strategies of discussing students' intelligence and academic performance (Rattan, Good, & Dweck, 2012; Turner, 2010).

Similarly, relatively few studies have examined how to best involve teachers in the process of reducing stereotype threat. Only one replication of the Cohen et al. (2006) values

affirmation intervention has examined how teachers' involvement affected student outcomes. Bowen and colleagues (2013) found that, among students who received the intervention, having teachers read their reflections resulted in stronger academic improvement compared to students who received the intervention but whose teachers did not read their reflections. While this study provides support that teachers play a critical role in implementation, and may similarly boost effects for other intervention strategies, it does not directly address the daily interactions that occur between teachers and students.

There are many other aspects of teachers' daily practice where they send messages about the nature of intelligence and stereotypes. For example, students pick up cues about their ability based on how teachers organize partner or group work. Ability-based grouping can signal to students not only that intelligence is fixed but that they can determine how intelligent they are based on which group they are assigned to (Boaler, 2013). However, when teachers facilitate activities such that all groups members can push the learning of others and/or are valued for their individual contributions to the group, students develop more positive beliefs about ability - in both their self-assessments and in their perceptions of others (Gehlbach, 2010). Similarly, presenting course assessments as part of the learning process rather than as objective indicators of ability provides frequent opportunities to build growth mindset beliefs and can drastically affect the engagement and learning of underperforming students (Smeding, Darnon, Souchal, Toczec-Capelle, & Butera, 2013). Because the link between students' beliefs about their intelligence and their classroom engagement and behavior is affected by teacher behaviors, it is critical to understand how teachers can best shape students' intelligence beliefs (Blackwell et al., 2007; Cohen et al. 2006; Cohen et al, 2009; Dinger et al., 2013). As teachers are better able to dispel

intelligence-based stereotypes and promote growth mindset beliefs in their daily instruction, students will be better set up for academic success.

The Student Perspective

As new research explores best practices for teachers to adapt growth mindset and stereotype threat strategies for use in the classroom, it must be built upon an understanding of how students make sense of intelligence and identity and how they naturally monitor their teachers' behaviors for clues about their own ability. Though much of the foundational mindset and stereotype threat research was conducted with undergraduate students, interventions developed from the research have often been implemented with middle school students, in part because students are more likely to hold growth mindset beliefs in elementary school but switch to holding fixed mindset beliefs during their middle school years (Eccles & Midgley, 1989; Meece, Anderman, & Anderman, 2006; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991). Compared to elementary school classrooms, the structure of the middle school classroom is more focused on achievement and grades, and middle school teachers report using growth mindset-aligned instructional practices much less frequently than elementary school teachers do (Eccles & Midgley, 1989; Midgley, Anderman, & Hicks, 1995)

It is logical to target middle school students for mindset and stereotype interventions, but, first, greater understanding of middle school students' thinking around ability and their perspectives of their classroom environments is needed. Youth's capacity to make sense of nuanced concepts like intelligence, identity, and stereotypes – in addition to their ability to make sense of others' behaviors – may change drastically during adolescence, making it difficult to assume with confidence that students' perceptions of how intelligence operates in

the classroom match our understanding of them (Eccles, Lord, & Midgley, 1991). Exploring students' experiences with how their teachers convey the nature of intelligence may point to new methods or strategies teachers can use to promote a growth mindset and to reduce stereotype threat in the classroom. Qualitative work is particularly important for understanding students' perspectives because it can more accurately capture critical details about students' thoughts, emotions and behaviors that would otherwise be overlooked in quantitative analyses (Bradburn, Sudman, Wansink, 2004; Berkwits & Inui, 1998).

This study aims to supplement the growth mindset and stereotype threat literatures by examining middle school students' stories and descriptions of how teachers convey the nature of intelligence and intelligence-based stereotypes to explore strategies for intervention. We pursued the following research questions:

1. What do middle school students believe about the nature of intelligence and intelligence-based stereotypes?
2. What are middle school students' perceptions of and experiences with teacher behaviors and messages regarding the nature of intelligence and intelligence-based stereotypes?

Methods

This study uses qualitative data from focus groups with middle school students. Data were analyzed using thematic analysis, an inductive approach (Braun & Clarke, 2006). Focus groups were audio recorded and discussions were coded, transcribed and analyzed by the primary researcher.

Participants

The sample for this study was drawn from three middle schools in a rural county in the Southeastern United States. All schools are in the same county: Schools 1 and 2 are public schools in the same school district, and School 3 is a public charter school that operates as part of its own school district. These three schools were selected because they predominantly serve students of color in an economically distressed community. The county population is slightly less than 25,000 people, over 20% of residents are living below the poverty line, and only 50% of residents over the age of 16 are employed. Although the racial demographics of the county are nearly 60% black and 40% white, the racial demographics of both school districts are disproportionately black – nearly 80% of students in the county school district and 75% of students in the charter school are black. In the charter school, 75% of students are considered low-income, as defined by participation in the Free and Reduced Lunch Program, compared to more than 98% of students in the county school district.

Historically, the county school district has demonstrated low academic performance. Currently, only 70% of its students graduate from high school within four years, and fewer than 25% of middle school students score proficiently in reading and math on state-wide standardized assessments. As a result, the charter network, under which School 3 is housed, was established in the county. The charter school's historical performance on standardized assessments indicates that, although it serves a similar population as the county school, students are making significant academic gains. School-wide, over half of students score proficiently in reading and math on state-wide standardized assessments. At many grade levels, students outperform state proficiency averages.

The sample for the study included 44 students from the three middle schools. Students' demographic information was aggregated using administrative data provided by

schools. The total sample included 18 boys (40.9%), 26 girls (59.1%), 31 black students (70.5%), 10 white students (22.7%), 2 Hispanic students (4.5%) and 1 multiracial student (2.3%). Thirteen students (29.5%) were in sixth grade, 15 students (34.1%) were in seventh grade and 16 students (36.4%) were in eighth grade. Overall, the sample includes 28 charter school students (63.6%) and 16 county school students (36.4%). See Table 2.1 for more detail on the overall sample.

Sampling

All students at each middle school were eligible for participation, unless they received services for severe cognitive impairment through the occupational course of study program (OCS). Non-OCS students who received special education services or Section 504 accommodations were eligible for participation. Recruitment in Schools 1 and 2 was conducted using simple random sampling. At each school, rosters were blocked by grade and alphabetized. Twenty students were sampled and given a consent packet to take home to their parents and return to school administrators. Four students consented to participate in School 1 and took part together in one focus group. Twelve of the 20 randomly selected students consented to participate in School 2 and were placed into three focus groups based on the order their consent forms were received. The first two focus groups had six and three students, respectively. The third focus group was conducted with three students at School 2 who were absent on the day that the first two groups took place.

In School 3, sampling was conducted using cluster random sampling because administrative staff preferred to have students participate in the focus groups during lunch, which was scheduled by grade. Fifteen students were sampled from each grade-level roster. Six sixth grade students, 11 seventh grade students and 11 eighth grade students consented to

participate. Five focus groups were conducted in School 3, and groups contained either five or six students. In all, one focus group was conducted in School 1, three groups were conducted in School 2 and five groups were conducted in School 3. See Table 2.1 for more detail on group composition.

The random sampling of students, though not critical in qualitative research, promotes credibility in these data because students were not recruited based on researcher or teacher bias (Shenton, 2004). Because the discussion focused heavily on students' school experiences and teachers' preferential treatment, to reduce bias, it was important to avoid teacher or administrator recommendations for participants.

Data Collection: Focus Groups

Data for this study were collected under IRB approval from the University of North Carolina at Chapel Hill, with students first obtaining parental consent and then providing their own verbal and written assent on the day of participation. Focus groups were chosen to create rich discussions among participants that may be more detailed and may present a wider range of responses than what might have been elicited in individual interviews (Morgan, 1997). To minimize disruption in the school day and loss of instructional time, when possible, students were interviewed during lunch. Focus groups started with a discussion of how students did not have to share anything that they were uncomfortable disclosing to the group. Students were assured that their responses would not be shared with teachers or school staff but were reminded that, while other participants in the group agreed not to share comments or details of the discussion, we could not guarantee total confidentiality. Students were asked to keep this risk in mind as they responded to questions. To promote privacy, focus groups were conducted with the door closed in non-instructional

spaces (e.g., administrator offices, conference rooms, etc.), reducing the likelihood of contact with teachers, staff, and other students. Names of participating students were not used in the focus groups.

Focus groups were conducted in Schools 1 and 2 during the spring of 2015 and were facilitated by two researchers. Focus groups were conducted in School 3 during the spring of 2016 and were facilitated by one researcher. To increase students' feelings of comfort and safety, all groups began with a brief conversation about students' interests and shared experiences to highlight commonalities and increase relatedness among the group (Morgan, 1993). All groups were audio taped, and recording began after the conclusion of these initial warmup discussions. After the completion of the first four groups, data were transcribed verbatim from audio recordings by the primary researcher. A second round of transcription was conducted after completion of the final five groups.

Focus group facilitators used a structured approach to interviews to guide discussion around the research questions and to promote consistency of discussion topics across groups (Morgan, 1997). Each focus group was asked a series of 10 questions (see Appendix A, Guide 1), which served as the scaffold of the interview, and facilitators used probes and follow-up questions to elicit more discussion, details and specific concrete examples. Interviewers facilitated conversation through accent probes (e.g., Wow!, That's interesting, etc.) to make students feel heard and to encourage them to continue talking. Facilitators used probes and follow up questions to elicit more details on the cognitive and affective content (e.g., What did you think about that? How does it make you feel when a teacher does that?) to encourage students to discuss the impact of classroom events or teacher behaviors rather

than to simply describe them. Finally, when necessary, interviewers used paraphrasing to repeat or clarify students' responses (Gorden, 1998).

The level of facilitation varied among groups (Morgan, 1997). Groups 1, 2 and 4 required high levels of facilitation and probing. Particularly in Group 4, some students were reticent to participate and gave short, general responses that lacked detail. Students in Group 4 needed encouragement to engage in the conversation and responded to prompts as individuals rather than responding to each other in a whole group discussion. In Group 2, two of the three students were hesitant to participate, while the third student was especially eager to participate. Group 2 was facilitated with the aim of balancing responses across students, and many probes were directed at specific students to encourage participation of all group members. Groups 3 and 5-9 required low levels of facilitation and probing. Students needed little encouragement to participate and commented frequently on others' thoughts and opinions, sharing explicitly whether they agreed or disagreed.

To avoid overrepresentation of individual students' experiences, discussions were facilitated by encouraging differences of opinion and by soliciting responses from those who were less dominant in the discussion (e.g., Do you have anything to add? I'd like to hear what so-and-so thinks about that; Morgan, 1993). Similarly, many students in the group had the same teachers (in the past or currently), and it was often the case that a student would point out, without further detail, that he or she had also had a teacher who was the topic of discussion. These students were probed further (e.g., What do you think about that? What was your experience like?) to foster discussion among students and to encourage students to share differences of opinion (Gorden, 1998).

During the first focus group, it became apparent that students' answers to questions were enriched when they were prompted to consider a specific teacher in mind or from the perspective of other students (e.g., Do you think everyone in your class feels that way?). The interview protocol was modified to reflect what was learned through this initial conversation and to promote richer discussion among students (see Appendix A, Guides 1 and 2, for comparison). The interview guide was revised to focus on specific teacher behaviors and also included language that students used to describe concepts (e.g., *fairness* or *favoritism* in addition to *stereotype*). Upon completion of data collection, focus group transcripts were analyzed to identify which questions were asked across focus groups. Appendix A includes questions that emerged from the conversations in the first focus group that were discussed consistently with all subsequent groups. A review of transcript data indicated that data pertinent to the research questions for this study were present in all focus group discussions; therefore, the full sample was retained for analysis.

Data Management and Analysis

All transcripts were managed using ATLAS.ti and analyzed inductively, using thematic analysis (Braun & Clarke, 2006). Thematic analysis is data driven, and data are encoded and organized before they are interpreted and further analyzed for patterns and themes (Boyatzis, 1998). Thematic analysis is appropriate for these data because it does not require the use of a theoretical framework and can be used to explore participants' experiences and descriptions of their reality related to a particular phenomenon (Boyatzis, 1998; Braun & Clarke, 2006; Daly, Kellehear, & Gliksman, 1997). As defined by Braun and Clarke (2006), thematic analysis includes six stages: (a) familiarization with data, (b) generating initial codes, (c) searching for themes, (d) reviewing themes, (e) defining and

naming themes, and (f) producing the report. Each stage will be discussed in greater detail below. At all stages, the primary researcher drafted analytic memos and created a detailed audit trail to document this analysis (Padgett, 2016; Panditt, 1996; Strauss & Corbin, 1994).

Familiarization with data. Familiarization with the data began during transcription, where the audio recordings were transcribed by the primary researcher. Multiple readings of the text that accompany manually transcribing data – especially data that have been collected by the analyst – promote close familiarity with the data and provide an initial opportunity to note interesting patterns and to generate ideas for further analysis (Bird, 2005; Lapadat & Lindsay, 1999; Riessman, 1993). After transcription, the primary researcher read the transcripts twice, to deepen familiarity with the data and to continue to note potential meanings and patterns (Braun & Clarke, 2006). Before coding began, transcripts were reviewed to identify data relevant to the research questions (Miles & Huberman, 1994). Data were selected for further analysis if students were discussing: (a) their beliefs about the nature of intelligence (i.e., whether you can get smarter), (b) teacher beliefs about the nature of intelligence, (c) teacher behaviors that communicate their beliefs about students’ ability, (d) what stereotypes are and what stereotypes students know about, (e) stereotypes teachers hold about students’ ability, or (f) teacher behaviors that convey stereotypes about students’ ability. Data were excluded if students were discussing teacher behaviors that were not specific to intelligence or stereotypes (e.g., being fun) or were describing classroom or school factors that were unrelated to students’ beliefs about intelligence, their own ability, and how teachers influence those beliefs (e.g., general complaints about lack of free time, homework, etc.). The following analytic questions guided coding and analysis of the data:

1. What do students believe about whether intelligence is malleable or fixed?

2. What intelligence-based stereotypes are students aware of, if any?
3. What do students believe about whether intelligence-based stereotypes are true?
4. What do teachers do and say that makes students feel smart?
5. What do teachers do and say that makes students feel not smart?
6. How do teachers address stereotypes of intelligence with students?
7. How do students tell whether teachers think some students are smarter than others and who those students are?

Generating initial codes. Initial rounds of line-by-line open coding were conducted to label students' statements with descriptive labels that, when possible, focused heavily on actions and feelings as described by participants (both students' and teachers'; Braun & Clarke, 2006). Line-by-line, open coding methods were critical for the analysis of these data because there is established theory related to growth mindset and stereotype threat and their influence on students' thoughts and behaviors. Open coding allows participants' experiences to drive analysis, which can help identify data that are disparate from or more nuanced than extant theory, and it can reduce the likelihood of over-synthesizing initial codes using prior knowledge (Charmaz, 2006). Best practices for initial coding include generating as many codes as possible; our process of open coding resulted in nearly 100 discrete descriptive codes (Braun & Clarke, 2006). The initial list of codes and their attached quotes was reviewed by a second researcher, who checked random excerpts of coded data to assess face validity and consistency of code use.

Throughout the coding process, brief comments were attached to codes to clarify their use and meaning. Initial codes were further developed and refined through an iterative process (Cohen, Manion, & Morrison, 2011; Creswell, 2009). Upon completion of an initial

round of coding all transcripts, all quotes and their attached data were compared to assess the consistency of the application of codes across transcripts (Glaser, 1992). To prevent drift, codes that changed in their application across transcripts were divided or recoded, and all transcripts were re-analyzed using the modified codes. After several rounds of comparative analysis, a list of 88 codes and their definitions was developed. This list was checked against data segments attached to each code to assess the accuracy and scope of code definitions.

Searching for and reviewing themes. The stages of searching for and reviewing themes involved moving to a broader view of the data by sorting codes into potential themes and evaluating and refining those themes (Braun & Clarke, 2006). At this point, the primary researcher began grouping related codes together to identify potential categories that emerged from the data. Potential categories were loosely grouped based on commonalities among the codes clustered within the category (e.g., “stereotypes students see”). During this stage, codes within a category were mapped together using linkages that defined the relationship among codes; for example, if one code is the *cause* of another code, if one code is *associated* with another code, if one code *contradicts* another code, etc. Codes were iteratively grouped and regrouped into different categories based on these maps. If after defining the relationships among several grouped codes in a category, it became clear that a code no longer belonged in the group or was a more accurate fit elsewhere, codes were reorganized. For example, codes such as *feeling embarrassed for being dumb* and *feeling bad about falling behind* were originally included under the category “student beliefs about intelligence,” but after mapping this category, it was clear that these codes did not connect as well to others in the category and fit better under a separate category called “how students respond to failure”.

Contradictory codes were noted to inform later negative case analysis (Corbin, 2017; Padgett, 2016). It was noted whether a data segment was in contradiction to the relationship among codes in the category or whether it was a non-example that further defined the category. For example, in the category “how teachers respond to failure,” the code *shaming students* was identified as an informative non-example of *making students feel safe*, but the code *being born smart* was noted as a direct contradiction and potential negative case of the code *not being born smart*. An initial set of categories and maps was developed and reviewed by the second researcher.

Defining and naming themes. The aim of this stage is to refine the meaning of emerging themes and to establish clear definitions and descriptions of each theme (Braun & Clarke, 2006). Thematic maps, their codes and the associated data were reviewed to assess whether there was coherence among the codes in each theme and whether there were clear boundaries among themes (Patton, 1999). For each theme, detailed descriptions were developed to document and summarize the analysis process. In tandem, themes were reorganized to increase specificity, and adjustments were made to theme descriptions based on continued review of the related codes and data. Contradictory codes or examples within a theme and across themes were explored further through negative case analysis, where all transcripts were reviewed several more times using the list of contradictory codes/concepts to check for their presence in the data (Padgett, 2016).

Producing the report. Finally, as suggested by Braun and Clarke (2006), the selection of illustrative quotes and examples and the synthesis of study results provides an additional analytic opportunity, where presented themes can be checked for coherence and validity and

modified as necessary. In line with this process, themes were solidified and tweaked while reporting the findings of the analysis.

Results

Analysis of student responses resulted in four themes. One theme, *everyone is smart*, emerged in response to our first research question regarding students' beliefs about the nature of intelligence and stereotypes. The remaining three themes, *the smart kids*, *instrumental support*, and *emotional support*, emerged in response to our second research question regarding teacher behaviors related to conveying the nature of intelligence and stereotypes. In the following section, individual participants are identified with two numbers, first with their focus group number, followed by their identification number within the focus group. For example, a student who participated in the first focus group who was given the ID #4 within the group will be referenced as 1-4. A student from the same group who was given the ID #2 will be referenced as 1-2.

Everyone is Smart

Participants were asked to describe where intelligence comes from and how intelligence factors into student outcomes (e.g., if a student were struggling in school). When directly asked about the nature of intelligence, the majority of students responded that no one is born smart and that intelligence develops over time. For example, students made statements such as, "I think people learn it over time from either going to school or reading or studying different things" (2-2), or "I don't like saying if somebody is smart or not cause we all have different levels of how much knowledge we know, and it's just a development over time" (4-1). These quotes demonstrate that some participants saw childhood as a time

when children “grow smarter” and learn to do things that make older children appear to be more intelligent.

Students also gave examples of skills that needed to be learned, as evidence that intelligence develops over time. One student (5-4) remarked:

It’s like, you go through school so you learn things. And the more you know, the more intelligence that you get...so it’s not like you’re just born smart, knowing how to do numbers, like multiplying and stuff like that. You’ve got to learn how to do that. That’s what makes you smart - learning.

However, in one group, a participant brought up a conflicting example of a classmate who excelled in school but did not appear to work hard. Participants discussed this classmate’s intelligence, but did not seem to believe that it was because she was naturally smarter than others:

2-4: In a way, I think, like, you’re kind of born like that some kind of way. Because I have a friend at lunch and she’s really smart but she tells me that she doesn’t study. She just remembers off the top of her head.

Interviewer: What do you think makes her that way?

2-4: I think, she get like a healthy diet or something, probably.

2-5: Another reason because [she] learned it over time. Nobody is born knowing, like, what’s 1,000 times 2,000. No one is born knowing that, you learn that over time by going through school. Or learning it from your parents or whoever is teaching you it.

2-2: I think people learn it over time from either going to school or reading or studying different things.

When asked to make judgments about the ability of struggling students, participants said that students may not do well because “they just lazy, they not doing they work” (2-1), “they never tried to see if they could” (1-3), or they never “pull the teacher to the side and tell her you’re struggling and get help” (6-4). As students in one group discussed:

Interviewer: Okay, so if they struggle a lot, what does that mean?

2-3: I guess you could have a learning disability.

2-5: Either that or you’re just not trying.

2-3: Or you’re not focused.

I: Okay.

2-1: Or you’re just not getting it.

2-4: Or not used to the material that you're learning.

I: Are they going to be stuck like that forever?

Multiple students: [*in unison*] Uh-uh. No.

2-2: They could ask the teacher for help.

2-3: They can ask for help or ask for tutoring if they need it.

Other participants described struggling students as smart but potentially not choosing the right behaviors or being affected by negative outside circumstances. These statements can be seen in the following exchange among participants:

3-1: I think they still smart, they just can't get the material. And I just think they probably just need more extra help on that material. So I still think that they're smart, they just need more help on that subject.

3-5: I think they're smart and it's like, you never know what their home situation could be and for like, our age, we have hormones that are jumping all over the place so we're focusing on our body and all this drama with ourselves and other people. So we're basically not focused on our work because we worried about something else going on.

3-3: It's like me. I know I struggle, but it's not...I know I am smart. It's just that like, either I don't get the materials or I don't ask for help when I need it.

I: Can they do as well as other students?

Multiple students: [*in unison*] Yes.

3-3: Cause we've seen it before.

I: What do you mean?

3-3: Like, [student names, outside of the focus group], they was like, failing just about all they classes and they still go for help for each of their teachers, and now they're doing a whole lot better with the teachers talking to them and trying to get them to do their work.

This excerpt shows that when asked to make decisions about the ability of others, participants often focused on explanations other than lack of innate intelligence, highlighting the importance of effort and getting help from the teacher.

The Smart Kids

Though most students believed that certain students were not smarter than others, many expressed that some teachers held stereotypes about which students were the "smart kids" whereas other teachers treated all students equally. Students' experiences varied by teacher, as exhibited in the following exchange among students:

I: What kinds of things do your teachers do and say about intelligence and ability in the classroom?

2-1: See, they tell us that everybody's smart.

2-5: Sometimes.

2-3: And some [teachers] be like "why can't you be like [student name] or [student name]?"

2-5: Yeah.

I: What does that look like?

2-5: They're sometimes mad and, well, usually they are mad when they're saying it...like, "why aren't you more like, [student name] or [student name]?" when we're in class learning, when we're trying to think.

Students in another group felt that their current teachers did not unfairly judge students but mentioned prior experience with teachers who did:

I: Do you think that there are teachers who stereotype against their students?

8-4: Around [this school]? No.

8-1: If they are, they're very good at hiding it. I feel like at my old school, it was more about stereotyping because I had this one teacher that always used to call me names.

Participants in another group described teachers who did not hold stereotypes about students' ability. One student cited her band teacher as an example:

1-3: Like what our band teacher was talking about in band today...everybody you sitting beside is just as important as you, everybody here is just as important as you. Nobody is smarter or better than you are.

Similarly, one student (9-1) mentioned a teacher from a different grade level who briefly observed his class and told students "she could see the ones who's gonna pass the [end of grade test] and the ones who's not." In another group, two students described a shared teacher who told students in her honors class that they are smarter than other students. On one occasion, this teacher was disappointed that her honors students were not performing as well as expected in class, citing their honors status as evidence that they should be doing better than students in her regular course. As these students recounted:

8-1 Student: Yeah, she asked a question and she be like, 'You're the smart class, you should know how to do this.'

8-3: When she said, ‘Oh, you supposed to be the smart group,’ she was putting us up as a higher ranking than all the other kids. They smart too, they might not catch on as fast as we do, but they still smart. You don’t need to say that.

Students in one group discussed that teachers’ biases about the smart kids could be damaging, because others may adopt the teachers’ beliefs of their ability over their own perceptions of their ability:

4-2: I think it affects their self-esteem. If they have a lower self-esteem then it make them feel like they can’t perform at their best.

4-6: I guess they feel like, yes, if the teacher tells you you can’t do it, that’s what it means because they feel like the teacher knows best. But I feel that if you believe in yourself, you can do it – no matter who tells you you can’t.

4-1: To agree with her, it’s like...some people may feel like if the teacher say something bad about you or...it’s the way they say stuff, like, you can’t do it and stuff like that. It’s gone make that person feel like they just don’t want to do anything anymore. They may want to drop out of school.

Students monitored their teachers’ behavior to see which students were helped the most, which provided cues as to who the smart kids were. Many of the stories that students discussed involved times when they were waiting to be called on or to be assisted, but their teacher ignored their hands in favor of the smart kids – calling on smart kids frequently (or exclusively) in class. One student (1-1) said, “[Teachers] always call on them, every time they put their hands up. They probably don’t even know the answer and they still call on them.”

Students in another group discussed how they identify the smart kids:

I: If students are ever treated differently in the classroom, are there patterns that you see?

3-1: Yes because, you know, how some [teachers] call on the same people all the time...and you don’t get a chance to speak. Like they just keep on calling that same person.

I: Okay.

3-6 Student: I agree. One day I was in science class and, um, the teacher she called on [student’s name] the whole entire class period. It was our last period in our last class. She really didn’t give no one else a chance to talk, she just kept calling on that one person.

Students in another group stated:

5-5: Like another example is with [student name], I can tell she be confused, and when she ask a question, [the teacher] don't answer her question completely, she still be confused.

5-3: Sometimes, [the teacher] just walk away and don't answer the question.

5-2: If like, I'm sitting there raising my hand, we doing work, partner work, I'm raising my hand trying to get some help, she'll look at you and she'll walk away.

5-1: Go to somebody else.

As another student (3-1) described it:

I can be sitting with my hand up, knowing that I know the answer, and then she'll pick on them and then [to us] be like, "None of y'all know the answer." So it's like, they automatically go to the smarter person, the more intelligent person.

Another student recounted:

8-3: This is what I don't understand. They pick the good students that raise they hand, but when somebody else trying to do their work, [teachers] act like they don't wanna call on them.

These students' stories illustrate their perceptions that teachers interact with students differently, calling on and helping the smart kids more than others.

Students provided very few examples of times that they saw teachers stereotype "the smart kids" by race/ethnicity or gender, often speaking about favoritism with some uncertainty as to the cause. One student shared that teachers in his old school, which was composed almost entirely of black students, only taught students the bare minimum based on what would be on standardized tests, but he did not confirm that this was specifically because of their race/ethnicity:

2-1: Well, comparing schools to this school, at [old school name], it's a lot of, like, my color [black] that go to that school, then the teachers, they only teach them what the [end of grade test] says, but this school, they teach us extra.

I: Why do you think that is?

2-1: I really don't know.

Students in other groups pointed out that they sometimes saw differential treatment by gender but that it was related to other factors, not teachers' preference. A male student in one group discussed that female students were able to answer questions more often than male students, but pointed to the fact that this may be because the teacher used a "popcorn" strategy of letting students "pass" participation to others:

1-1: I can sometimes see it, but like...say if the teacher called on one girl – no offense to y'all, no offense – then after that, they say to pass it, [the girls] gone pass it to other girls and then the boys have they hands up forever.

Similarly, in another group, students pointed out that teachers may treat female students better because male students are more likely to misbehave.

I: How might [teachers] decide who they treat better than others?

9-1: They'll treat the girls better than the boys.

I: Do you girls think so?

9-3: I mean them boys act up more than the girls do, so...

[*boys nod*]

I: I see the boys nodding yes.

More often, students shared that teachers commonly judged students because they come from a rural, low income community:

I: What types of stereotypes do you feel like you encounter?

4-3: Coming out of [our town], a lot of people might think we are not as smart as them.

4-1: Or the community, not just like [our town] but like the suburban versus rural areas or something like that.

Students in another group discussed that teachers may assume that students who appear poor will not do well in school:

I: What do you think about how the teacher decides which kids are smart and which kids aren't?

7-2: Like who looks the neatest or who looks the cleanest.

7-1: Mmhmm. I'm not trying to mean right now, but this boy...he don't come to school with no bookbag, no pencils. He try to ask other children if they've thrown their bookbag away in the trash, he ask them can he get their bookbag...so [teachers] just let him go to sleep in their class. Tell you the truth, they don't help him. And that

mean that he ain't learning...[Teachers] think the ones that live in the ghetto and the hood and stuff, [teachers] think that they gone already come to school and do bad.

Students also described that teachers may make assumptions that students who appeared to be poor or from “bad neighborhoods” would be unruly, violent, and dumb. As one student described it:

1-1: They think that they all so smart and then the people that is in, like, rural communities is always bad. Will kill you...will drop out of school, have babies early and stuff.

One student recounted an instance of a teacher doubting that she would “make it out” of her town, explicitly telling her that she would not be successful in the future:

8-3: One time when I told him I was going to this college, [he] talking bout, ‘No you not.’ Like, you can’t doubt us...we might be better than you one day, we might be in your chair that you sittin’ in one day. You can’t be doubting us like that, like we not gone make it.

I: Teachers are telling students that they doubt what they’re able to do?

All students: [*in unison*] Yes!

8-3: They doubt you, and you can’t doubt nobody. You never know what they gone do in life.

These excerpts demonstrate that students felt not only that they faced a stereotype of being less intelligent because they come from a low income community but also that they felt that some teachers used this stereotype to predetermine which students would be successful in the classroom.

Instrumental Support

Participants were asked to describe teachers who made them feel smart and capable and those who did not. In their descriptions, students reported that instrumental support was a key indicator of a teachers’ beliefs about their ability and made students feel that they could “be smart” or “get smarter” in that teachers’ subject. Students described two important aspects of teachers’ help: first, that teachers would offer to spend extra time with students,

providing tutoring or other help; and second, that teachers would offer students the opportunity to make up work or redo failed assignments.

One student (6-1) described the teachers who make her feel smart as those who “help me with the work...if I raise my hand, I’ll ask the teacher for help and then she or he will come to help me with the classwork.” Students in another focus group described a teacher who made students feel smart by offering extra help and making sure all students got the help they needed, whether it was during or outside of class:

2-3: Yeah, like if you don’t understand...and if you don’t understand and you don’t say nothing, she know you don’t understand. She’ll pull you to the side and tell you to come to her class during your elective or lunch.

2-5: And even if she’s at her desk and we’re working in groups, she’s working and looking around seeing what we’re doing and if we’re confused. She sometimes calls us up individually, seeing if we need help or anything.

Other students stated:

5-1: Like if you would like to stay after school for extra help for the people who...[trails off]

5-3: If you need help, some of the teachers, they’ll come to you at lunch.

Several participants in a different focus group shared stories about a teacher who made them feel smart, describing how she not only provided individual help to students during and outside of class (including on the weekends), but she also gave students the opportunity to redo their work and helped them complete their makeup work:

3-6: She gives them their work that they need to make up and then she just tell them, like, ‘you’ll get through it’ and then she be like, ‘you need some help?’ She ask you if you need help and stuff like that.

3-1: I was thinking [the same teacher] because I struggle a lot in math and when I stay after school, she kind of pull me to the side and she help me out with stuff that I need help with.

3-4: I can agree with [3-1] because in February, I was struggling with math and she called my parents to tell them, to ask them can I come to school on a Saturday [so she can] help me out.

3-3: I can agree too because the last two quarters, I been failing her class completely. But when I got my report card, I noticed that all my grades were like Cs...She calls

you up and tell you what you can work on and what you need to do to help pull your grade up. To me, it's better because some teachers, they won't even tell you what you doing wrong, they just wait 'til you get your report card.

In contrast, multiple students described teachers who denied help through explicit policies where they would only make one copy of a worksheet or assignment and that students would "get what they get" in terms of grades. One group of students described a teacher who did not make them feel they would do well in class:

6-2: She's like, 'y'all gotta work, like, this problem, this problem, this problem.'

6-4: And don't lose your paper, cause she ain't making a second copy.

6-2: Yeah, we don't get no second copies.

6-3: Yeah, she said, 'y'all not getting a second copy.'

6-1: And turn it in, finished or not.

Students in another group compared teachers who made them feel smart and teachers who did not, based on their willingness to help and whether they provided makeup work:

8-1: That's how [teacher name] was in my sixth grade class so, like, any time I need help, she'll help me, but other teachers, they, like, get an attitude every time you ask for extra work or help to pick your grade up.

8-2: Yeah, like yesterday, she asked a question and she was pressuring us. She was like, 'You're the smart class, you should know this' and then [another student] was like, 'Well, maybe you taught it, but we might not understand it.'

8-3: Like, [teacher name]...I be like, '[Teacher name] – can you help me?' and she be like, 'Well you should have got your notes and you should know how to do it yourself.' And all that.

If a teacher was willing to help, students felt that teacher thought they had the capacity to understand lesson material and were therefore smart. One student (1-3) reported feeling more comfortable in these teachers' classes:

When you need help, they'll help you, and I know some people want to go to certain classes because that's the subject they feel most comfortable with because of things the teacher talk about, they can relate to or they get one-on-one help with the teacher.

Another student (2-3) reflected on a teacher who made her feel smart, even when she struggled in class:

Like say, when she called [on] me and I didn't get it, she would stop what she was doing...she would go over it on the board and ask me if I knew and then she could come to me after class and she could tell me, 'You can come to my room tomorrow, if you still need help.'

In contrast, when teachers refused to help, students wondered whether they possessed the ability to do well, as described by this student (1-4):

My [social studies] teacher, she would call on me in class and when I asked her to help me with my work, she won't help me...I don't like social studies...it make me think I don't really get it cause she ain't really helping me with it.

These excerpts illustrate that students looked to a teachers' willingness to help as a key indicator of their ability in the class.

Emotional Support

In addition to helping and encouraging behaviors, teachers' warmth and care was also seen as an indicator of students' intelligence. Students described teachers who made students feel "comfortable" in class and were able to identify when students were having a bad day or in need of emotional support, not simply struggling with class material:

1-3: It's like, they don't try to say anything that might make you upset or that might not make you want to learn. But when you need help, they'll help you.

I: What types of things might make you feel uncomfortable?

1-2: If the teacher doesn't know, like, how I feel and she tells me to be quiet or something like that.

A student in another group stated:

I: What makes you really feel like you can do well even when it's challenging?

2-4: Like, I think teachers should show more love towards the students because sometimes I do bad because I get fussed out for something. If I don't get fussed out, I do better than this.

One student (3-2) described her teacher's care as having a direct effect on her feelings of competence in class:

Usually my science teacher...right when I'm in class and I get upset at myself, she usually comes by me and pats me on the back and asks if I'm okay. And, like, asks

me if I want to go to the hallway to talk about it...it just boosts up my confidence to do my work. Cause usually the materials is really hard...so yeah, it boost up my confidence to do it.

In other groups, participants shared that teachers both showed a lack of care and made their beliefs about the smart kids very obvious by calling out struggling students in front the class:

7-1: Sometimes a teacher, they'll stereotype the children.

I: How so?

7-1: Like they might say, like some teacher will pick on you and try to call you names and say that 'you always know how to talk, but you don't know how to do your work!'

Students in another group shared similar experiences, describing how teachers would embarrass struggling students with what one student called "slick comments":

6-2: Some teachers, they like to make slick comments and like to say stuff.

I: So what type of comment might a teacher say?

6-4: Like uh, about the grades. When people be laughing and [teachers] be like, 'I don't know why you laughing when you need to be looking at your grades!'

6-2: Right.

6-3: That's what my reading teacher do.

6-4: It's disrespect.

6-5: Like she be telling us, 'You need to take your medicine!' and then, she don't know if we're on [medicine] or not, so how she know?

Other students shared:

3-4: I mean, some of the teachers, some of the stuff they say about you, like in front of the whole [class] is really not the [whole class'] business, it's just yours and the teachers.

3-5: I agree with [3-4] because, like, sometimes you can be playing in the classroom and the teacher gets so aggravated with you they be like, 'Well, I don't think you should be sitting there talking or sitting there playing around because last time I checked, your grades isn't like...' - they'll shout out another person - 'Your grades isn't like [student's name] or [student's name]. So you shouldn't be talking, you should be paying attention to what I'm saying.'

I: How does that feel?

3-5: It make you feel kind of uncomfortable because you're putting us on the spot saying our grades aren't as high as somebody else's. Well, you can't really compare our grades to theirs because they could be more in an advanced level of learning than we are.

Taken together, these excerpts highlight that teachers' care made students feel smarter in that teachers' class, while a lack of care conveyed that teachers' had low expectations about students' ability.

Discussion

Overall, themes derived from these data provide important insight into the influence of teacher behaviors on students' beliefs about intelligence and motivation to achieve. Our first research question involved students' beliefs about the nature of intelligence. Participants in our study tended to believe that every student can be smart and that intelligence increases with effort. Our second research question involved students' perceptions of teacher behaviors that message teachers' beliefs about intelligence and stereotypes. Participants felt that some teachers believe certain students are smarter than others and that characteristics like students' socioeconomic status factored into teachers' stereotypes. Teachers' instrumental support and emotional support made students feel smart, and students used disproportionalities in teachers' provision of both types of support to determine which students the teacher believed to be smartest.

Many students in the sample endorsed growth mindset beliefs, providing specific examples of how intelligence develops over time (e.g., math). Students tended to believe that intelligence is earned through effort, and many did not personally draw a connection between needing help and not being smart, even if they were struggling or failing a class. Many participants in this sample saw their inability to understand course material as temporary, and they made the same judgements of their classmates. They did not believe that struggling students were inherently less intelligent than others or that intelligence varied based on gender, race/ethnicity or other aspects of identity. However, participants did recognize

individual differences in ability, noting that some students can succeed with less effort than others, some students have learning disabilities and some students need more time or teachers assistance to master course material.

In several focus groups, such as the group where students wondered whether it was a healthy diet or prior exposure that made a classmate appear naturally smarter, students' discussion focused on bridging their beliefs that everyone is smart with their observations that students require varying levels of effort to succeed. Though students wondered about these individual differences, it appeared that the belief that everyone is smart superseded beliefs about fixed intelligence. For example, participants frequently ascribed differences in ability to differences in controllable behaviors, such as asking for help, paying attention in class and studying – or to uncontrollable external factors, such as a difficult home life. It is possible that students felt endorsing growth mindset beliefs was the socially acceptable response, and in the focus group setting, were less likely to share conflicting beliefs about the fixed nature of intelligence. Or, it is also possible that, in these schools, teachers have made efforts to promote growth mindset beliefs, reducing the tendency of students to adopt growth mindset beliefs as they get older (Eccles & Midgley, 1989; Meece, Anderman, & Anderman, 2006; Wigfield et al., 1991).

Though students tended to believe that everyone is smart, they did express that they began to doubt their ability in a subject when the teacher did not provide instrumental or emotional support. Students commonly referenced typical helping behaviors that teachers used to provide instrumental support, like tutoring outside of class, but they also highlighted the impact of teachers knowing when students needed help and providing help when needed. Studies of teachers' instrumental support show similar patterns, where either total lack of

support or excessive, unnecessary support from teachers creates the impression that students are not capable and reduces student motivation and engagement (Urdan & Schoenfelder, 2006). Interestingly, students tended to believe that if they received a lot of help from the teacher, it was because the teacher thought they were smart and worth helping – not that they were “slow” and needed a lot of help to learn. In the same way, if a teacher demonstrated care, it was directly related to the teachers’ beliefs about their ability: Teachers showed care to help students achieve in class because they believed students had the capacity to achieve. When teachers seemed unwilling to provide help, denied “retakes” and did not provide emotional support – for example, openly shaming students for underperformance – students saw these behaviors as cues that the teacher did not believe they could succeed.

Our findings support education research based in self-determination theory, which has shown that teachers’ interpersonal involvement and relatedness with students (i.e., building relationships and providing emotional support) is linked with students’ sense of competence and intrinsic motivation (Connell & Wellborn, 1991; Ntoumanis, 2005). Turner and colleagues (2002) found that teachers who were most supportive of growth mindset development in their students maintained high expectations for student work but paired expectations with frequent social and emotional support. Similarly, our findings reflect a model tested by Taylor, Ntoumanis, and Standage (2008), who identify instrumental help and support (e.g., providing help and guidance for students to improve) and gaining understanding of students (e.g., fostering meaningful affiliations) as two critical motivational strategies teachers can use to foster students’ sense of competence and relatedness. Students in our sample echo the elements of self-determination theory, showing that teachers’ help and

care influence not just students' happiness but their beliefs about their own academic abilities.

In addition, students identified the ability to redo assignments or complete extra work as a sign that teachers thought they could improve in a class. Much research has been dedicated to studying the impact of grading structures on students' motivation. "Retake" policies align with research on grading and assessments that indicates that in comparison to competitive grading structures, which emphasize discrepancies in student performance, formative assessments – including the opportunity to resubmit assignments – help convey that students have the ability to grow and develop mastery over time, supporting growth mindset beliefs (Boaler 2016; Cohen, Steele, & Ross, 1999). Other evidence suggests that when teachers emphasize the importance of learning and making improvement, students place more importance on the learning process and are more likely to develop growth mindset-aligned goals (Maehr & Midgley, 1996). Linnenbrink (2005) found that when teachers recognize improvement through their grading practices rather than use evaluations as a metric for comparing students to each other, students are more likely to endorse growth mindset beliefs than students whose teachers publicly recognize high performing students. Because students' sense of competence is influenced by their capacity to do what is required of them, when teachers create value solely around the process of scoring highly on assessments, with little value on the learning process, struggling students do not feel competent to meet expectations; when teachers create value around the process of working hard, any student can feel competent to meet expectations, increasing motivation and engagement (Urduan & Schoenfelder, 2006). As students in this sample discussed, retakes can be one way that teachers can emphasize the importance of improvement in the learning

process rather than the value of normative performance grades (Urda & Schoenfelder, 2006).

Much research has also been dedicated to the mechanism through which students' sense of competence and growth mindset-aligned goals influence their behavior and achievement. Dinger and colleagues (2013) showed via structural models that students' perceived competence is related to greater hopes for success and lower fears of failure. In tandem, competence and high hopes lead to an increase in intrinsically motivated behavior (i.e., mastery goals, learning for the sake of learning) and, subsequently, increased academic achievement. A similar connection between students' beliefs about their ability and their behavior has shown a connection between students' beliefs about ability (i.e., that they can improve with effort), decreased feelings of helplessness, an increase in the use of productive strategies (e.g., working hard, studying) and improved academic performance (Blackwell et al., 2007). Paired with findings from the self-determination literature, the Dinger and Blackwell models support a mechanism through which teacher behaviors that influence students' intelligence beliefs effect differences in their academic performance.

Though participants believed that everyone is smart, they did feel that some teachers believe that certain students are naturally smarter than others and other teachers believe that all students can be smart. Several groups gave examples of teachers who explicitly communicated positive beliefs about students' ability, stating that all students are equal, in contrast to their descriptions of teachers who communicated negative beliefs that students should "just be like" their more intelligence classmates. More commonly, however, students identified the smart kids based on differences in teacher treatment, typically referring to this behavior as *favoritism* rather than stereotypes. They identified discrepancies in teacher help

and attention (i.e., which students get called on in class) as a way to identify who teachers believed were the smartest, perhaps because participants also believed that teachers' instrumental support is linked to students' intelligence. Students tended to remember, in great detail, instances where they felt overlooked when raising their hands for help or to answer a question in class.

Yet, though students felt that some teachers held opinions about “the smart kids”, they did not believe that these opinions were tied to gender or race/ethnicity but were based on socioeconomic status. Students felt that teachers may stereotype students because they came from a rural, low-income community – including relative comparisons based on where students are situated within the same community (e.g., lowered expectations for students from “the hood”). A student communicated this stereotype powerfully when he shared the anecdote of a classmate who appeared poor and was therefore not provided much help by teachers – highlighting that teachers' judgments prevented this student from learning. Some work has been dedicated to exploring the role of intelligence beliefs in low income, rural samples, including a randomized controlled trial with adolescent girls that demonstrated increased grades via an increase in growth mindset beliefs and learning motivation (Burnette, Russell, Hoyt, Orvidas, & Widman, 2017). Although stereotype threat based on socioeconomic status is not commonly assessed in school-based interventions, our findings suggest that rurality and poverty may be two areas where students – including white students – may experience suppressed achievement due to stereotype threat. Indeed, Bowen and colleagues (2013) explain that conducting a threat-reducing values affirmation in a low income school may explain significant increases in achievement for both black and white students. And qualitative evidence on the intersectionality of low socioeconomic status and

rurality indicates that both aspects of identity are salient to students in relation to their academic self-perceptions, increasing fear, anxiety and concern about future academic success (i.e., graduation from a four-year college or university; Morton, Ramirez, Meece, Demetriou, & Panter, 2018).

The Middle School Context

Taken together, these findings provide support for the trend that many students come to middle school with a growth mindset and shift to fixed mindset beliefs over time, particularly in response to teacher behavior (Eccles & Midgley, 1989; Meece, Anderman, & Anderman, 2006; Wigfield et al., 1991). When asked to think about intelligence generally, students in this sample tended to endorse growth mindset beliefs but, when thinking of specific interactions with teachers, students demonstrated an awareness that growth mindset beliefs may not be shared by all teachers, which caused them to wonder whether they had low ability in certain subjects. Over time, the evaluative nature of secondary schools, the pressure to perform academically, and students' "teacher knows best" attitudes may cause students to supplant their own beliefs with what they perceive to be their teachers' better informed evaluations. The power of an individual teacher is highlighted in the fact that participants' assessments of their abilities were often subject-specific, based on the behavior of that teacher. One student, for example, stated that her social studies teacher's failure to help led her to believe that she was not smart in social studies, but not necessarily that she was incapable overall. A similar trend has been found in quantitative work that shows students start to form domain specific attributions about ability in middle school (Meece, Anderman, & Anderman, 2006). Relatedly, mindsets can also become domain specific, particularly in math, where students are more likely to have fixed mindset beliefs compared

to other subjects, where they may still hold growth mindset beliefs (Buehl, Alexander, & Murphy, 2002; Jonsson, Beach, Korp, & Erlandson, 2012; Yeager & Dweck, 2012).

Our findings also reflect developmental trends in adolescence where middle school students become increasingly aware of individual differences and grow in their capacity to consider causal effects and to make judgments about others (Eccles, 1999; Wigfield, Eccles, & Pintrich, 1996). However, our data lack consistency regarding students' ability to think critically about intelligence and identity, to perceive nuance and to make judgments about others' behavior. When some students discussed instances where teachers treated students differently by gender – disproportionately disciplining boys or calling on girls – they were able to consider that it may not be caused by teacher stereotypes about gender but by differences in student behavior by gender (i.e., boys being more active in class) or in classroom procedures (i.e., students “popcorning” participation to other students of the same gender). Some students demonstrated a level of nuance in determining their reasons; for example, the student who believed that a potentially “gifted” classmate who did little work but excelled in class may have simply learned that material elsewhere.

Other times, students used very concrete and fixed criteria for processing the differential treatment different their classmates received, without thinking much about the circumstances or context. For example, students tended to believe that if a teacher did not call on a certain student in class, it was always because the teacher did not think the student knew the correct answer. They also tended to believe that when teachers denied help, it was always because the teacher thought the student was not smart. Students considered what the teacher was thinking about students' ability, but tended to make the same attributions any time they or another student was not called upon. Of course, it is possible that the teacher

may simply be calling on students at random or may have chosen not to call on a student because it was already clear they *did* know the correct answer, but students in this study did not appear to think too deeply about the possible motivations behind this specific teacher behavior. They also tended to share this attribution with helping behaviors, monitoring inequity in teachers' help as evidence of teachers' beliefs about students' intelligence.

Interestingly, though students were able to describe individual differences in ability and understood that students required varying levels of help, at the same time, they looked to whether teachers called on, helped and cared about all students equally as an indicator of whether the teacher thought all students were intelligent. The conflict between these two beliefs reflects adolescents' growing ability to perceive and make attributions about the behavior of others, particularly with moral judgements about what constitutes the "right" thing to do (Crone & Dahl, 2012). In general, adolescence marks a time when youth progress from more rigid justifications about behavior (e.g., importance of compliance with set rules, only one right way to act, etc.) and more sophisticated justifications about behavior (e.g., situation-specific determinations requiring multiple viewpoints; Beauchamp, Dooley, & Anderson, 2013). However, adolescents' development of increased sophistication in their perspective-taking and reasoning is not always synchronous and can vary widely based on youth's age, gender and cognitive capacities (i.e., executive function; Van der Graaff, Branje, De Wied, Hawk, Van Lier, & Meeus, 2014; Vera-Estay, Dooley, & Beauchamp, 2015). Future studies with data that are not pooled across all middle school grade levels would be more appropriate to investigate how student characteristics affect their decision making about teacher behaviors related to the nature of intelligence.

Implications for Practice

Thus far, the literature on the development of growth mindset beliefs in children has focused heavily on what is *said* to students. Some scholars have criticized classroom implementation of the theory as setting teachers up to encourage students to simply “have a growth mindset” without changing any elements of the classroom context that impact students’ mindset development (Edwards, Esmonde, Wagner, & Beattie, 2017). Most research has focused on the direct impact of telling students that they can grow their brains or that they can get smarter through hard work (Blackwell et al., 2007; Paunesku et al., 2015; Yeager et al., 2014; Yeager et al., 2016). Other work has focused on the impact of praise, showing that it is more motivating to praise students for working hard than for being smart (Cimpian et al., 2007; Henderlong & Lepper, 2002; Rattan et al., 2012). Though these strategies are certainly important, our study contributes to the literature by demonstrating that, in the middle school classroom, students may also attend heavily to the things that teachers *do*, even if teachers never directly discuss intelligence. One specific strategy identified by students in our sample is the ability to make improvements on their work via resubmission or “retakes”. When given the opportunity to resubmit work, students felt that teachers believed that they could improve with effort. In addition, teachers may need to be mindful that middle school students may be monitoring which students teachers help or call on most for cues about their own ability. In processing their observations, students may use inconsistent logic to extract meaning about their own abilities. In our sample, students showed a strong capacity to understand individual differences in learning, but they still maintained rigid expectations for totally equal treatment. Because meeting all students’ needs requires equitable rather than equal treatment, it may be necessary for teachers to pair discussions of the growth mindset with conversations about individual differences and by

openly address their decision making, to reduce the likelihood that students mistake their behavior as an indicator of fixed, low ability.

Likewise, teachers should be mindful that providing emotional support may help shape students beliefs about their ability in a particular subject. What's more, the tendency of students in this sample to endorse domain-specific beliefs about their ability highlights that it is not sufficient for only a few teachers to be supporting positive intelligence beliefs; a growth-mindset aligned social studies classroom may do little to support students' growth mindset beliefs about their ability in math. To optimally support students' sense of competence, all teachers in a building must work to create growth-mindset aligned classrooms. Future research should explore these and other teacher behaviors to assess their influence on students' intelligence beliefs, classroom engagement, and academic achievement.

Finally, in implementing strategies to reduce stereotype threat, teachers and school administrators should be aware that gender and race/ethnicity are not the only identities that may expose students to stereotype threat. Students in our sample felt most affected by stereotypes around socioeconomic status and living in a rural community; however, it is important to note that in other school contexts, students may perceive vastly different stereotypes at work, including traditional stereotypes about race/ethnicity and gender. Understanding the influence of a school's unique context, including the outside community, is critical for identifying which stereotypes may need to be dispelled to promote student achievement. Although the values affirmation task does not address any aspect of identity specifically, and may reduce threat broadly from a number of stereotypes, the impact of this intervention on identities other than race/ethnicity has yet to be tested.

Limitations

Although this study informs the literature about how mindsets and stereotypes of intelligence affect students' motivation and achievement, it is not without limitations. First, it relied on a small sample of students from one county, limiting the generalizability of the work. In addition, the nature of focus group interviews does not lend itself to ascertaining every students' opinion on every question. There may have been important viewpoints that were not expressed during focus group discussions and are therefore not represented here. Groups were not able to be structured homogeneously by students' race/ethnicity or gender, which may have limited students' willingness to disclose (Morgan, 1993). Similarly, social desirability may have led students to avoid discussing issues like race/ethnicity or gender, and students may have been more likely to claim to endorse growth mindset beliefs if they felt this was the desired response from the facilitator (Morgan, 1993).

In addition, only some focus groups were able to be organized by grade level, which limits our ability to make conclusions about students' developmental capacities to understand and make sense of the nature of intelligence and teacher behaviors – and how these beliefs change as they get older. Future work could improve upon our study by using individual interviews with students or by structuring focus groups by race/ethnicity, gender and age to better understand how these factors affect students' experiences. Finally, it is unknown whether schools in this sample were actively engaged in efforts to promote growth mindset beliefs or to reduce stereotype threat. The degree to which students were previously exposed to growth mindset and/or stereotype threat is also unclear. It could be that students in the sample were supported in their understanding of the nature of intelligence and intelligence-

based stereotypes and may be able to talk about them differently than students in other schools.

Throughout the study, the primary researcher engaged in a reflexive process and must acknowledge several perspectives that may influence the findings of the study. The first author is a former secondary education teacher in the public county school district from which a portion of the sample was drawn. Though prior experience provides greater familiarity with the context, it may bring assumptions that bias interpretation of students' responses (Padgett, 2016). Through coaching other teachers, the researcher had worked in the classroom with some of the students who were selected to participate, and others knew her because she taught older siblings or relatives. Familiarity with the facilitator may have put students at ease or increased their willingness to share, but it may have also encouraged students to affirm the concepts discussed in the groups. The potential for the facilitator to bias focus group discussion was considered when deciding to randomly sample students; at the very least, students were not chosen because of prior relationships. However, because the primary researcher was the single coder of these data, it is possible that prior experience or personal bias influenced the analysis of these data. Future work would be strengthened by the use of a second or multiple coders who may help reduce likelihood that findings are influenced by the personal biases of a single coder (Padgett, 2016).

Finally, the sample had an overrepresentation of students from the charter school, which may bias results. The charter school requires that students and their families complete a detailed application to attend the school, which may draw a self-selected population of students who may be different (e.g., more motivated, more engaged) or may have different environmental circumstances or supports (e.g., more parental involvement). Though there

was a high level of accord between statements of students from the county schools and the charter school, findings may vary in future studies with students from a wider variety of schools.

Conclusion

Despite its limitations, this study provides important insights on the use of growth mindset and stereotype threat theories in schools - particularly for low income students and students of color. These findings support existing research on growth mindset and stereotype threat theories, showing that teachers influence students' beliefs about the nature of intelligence and their own ability. In addition, while extant research has focused heavily on the effects of explicating the nature of intelligence with students, our findings suggest that students also attend to other aspects of teacher behavior to guide the development of their intelligence beliefs. Though the results of this study do not generalize to other settings, they indicate that future research on the implementation of growth mindset and stereotype threat interventions should explore students' perspectives to understand more about how students see stereotypes and the nature of intelligence at work in the classroom, how students' interpret teacher behaviors, how students perceive stereotypes at work in their unique school context and how students feel that they are affected by stereotypes about intelligence. In sum, this study informs ongoing research for two popular academic interventions and identifies new areas of focus for intervention strategies that show promise for promoting the academic achievement of our highest need youth.

REFERENCES: PAPER II

- Aronson, J. (1999). The effects of conceiving ability as fixed or improvable on responses to stereotype threat. Unpublished manuscript, New York University.
- Aronson, J., Blanton, H., & Cooper, J. (1995). From dissonance to disidentification: Selectivity in the self-affirmation process. *Journal of Personality and Social Psychology*, 68(6), 986. doi:10.1037/0022-3514.68.6.986
- Aronson, J., Quinn, D., & Spencer, S. (1998). Stereotype threat and the academic underperformance of women and minorities. In J. Swim & C. Stangor (Eds.), *Stigma: The target's perspective*. New York: Academic Press. doi: 10.1016/B978-012679130-3/50039-9
- Beauchamp, M. H., Dooley, J. J., & Anderson, V. (2013). A preliminary investigation of moral reasoning and empathy after traumatic brain injury in adolescents. *Brain Injury*, 27(7-8), 896-902. doi:10.3109/02699052.2013.775486
- Berkwits, M., & Inui, T. S. (1998). Making use of qualitative research techniques. *Journal of General Internal Medicine*, 13(3), 195-199. doi:10.1046/j.1525-1497.1998.00054.x
- Bird, C.M. (2005). How I stopped dreading and learned to love transcription. *Qualitative Inquiry*, 11, 226-248. doi:10.1177/1077800404273413
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, 12(3), 225-229. doi:10.1111/1467-9280.00340
- Boaler, J. (2013). Ability and mathematics: The mindset revolution that is reshaping education. *FORUM*, 55(1), 143–152. doi:10.2304/forum.2013.55.1.143
- Boaler, J. (2016). *Mathematical mindsets: Understanding students' potential through creative math, inspiring messages and innovative teaching*. San Francisco, CA: Jossey-Bass.
- Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preference for challenge: The role of perceived competence and control. *Journal of Personality and Social Psychology*, 54(1), 134. doi:10.1037/0022-3514.54.1.134

- Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2013). Enhancing a brief writing intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology, 105*(2), 427-435. doi:10.1037/a0031177
- Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.
- Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design--for market research, political polls, and social and health questionnaires*. John Wiley & Sons.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa
- Buehl, M. M., Alexander, P. A., & Murphy, P. K. (2002). Beliefs about schooled knowledge: Domain specific or domain general? *Contemporary Educational Psychology, 27*(3), 415-449. doi:10.1006/ceps.2001.1103
- Burnette, J. L., Russell, M. V., Hoyt, C. L., Orvidas, K., & Widman, L. (2017). An online growth mindset intervention in a sample of rural adolescent girls. *British Journal of Educational Psychology*. doi:10.1111/bjep.12192
- Cahalan, M., & Perna, L. (2015). Indicators of Higher Education Equity in the United States: 45 Year Trend Report. Pell Institute for the Study of Opportunity in Higher Education.
- Charmaz, K. (2006). Constructing grounded theory: A practical guide through qualitative analysis. Sage.
- Chhuon, V., & Wallace, T. L. (2014). Creating connectedness through being known: Fulfilling the need to belong in U.S. high schools. *Youth & Society, 46*(3), 379-401. doi:10.1177/0044118X1436188
- Cimpian, A., Arce, H. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science, 18*(4), 314-316. doi:10.1111/j.1467-9280.2007.01896.x
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*, 1307-1310. doi:10.1126/science.1128317
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*(5925), 400-403. doi:10.1126/science.1170769

- Cohen, L. M., & Manion, L., & Morrison, K. (2011). *Research Methods in Education*. UK: Routledge.
- Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, 25(10), 1302–1318. doi:10.1177/0146167299258011
- Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. *Minnesota Symposia on Child Psychology*, 23, 43-77.
- Corbin, J. (2017). Grounded theory. *The Journal of Positive Psychology*, 12(3), 301-302. doi:10.1080/17439760.2016.1262614
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches (3rd ed.)*. London, UK: Sage.
- Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social–affective engagement and goal flexibility. *Nature Reviews Neuroscience*, 13(9), 636. doi:10.1038/nrn3313
- Daly, J., Kellehear, A., Gliksman, M., & Daly, K. G. (1997). *The public health researcher: A methodological guide*. Melbourne: Oxford University Press.
- Davis, J. L., Burnette, J. L., Allison, S. T., & Stone, H. (2011). Against the odds: Academic underdogs benefit from incremental theories. *Social Psychology of Education*, 14(3), 331-346. doi:10.1007/s11218-010-9147-6
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, 26(3-4), 325-346. doi:10.1080/00461520.1991.9653137
- Dee, T. S., & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446. doi: 10.1002/pam.20586
- Dinger, F. C., Dickhäuser, O., Spinath, B., & Steinmayr, R. (2013). Antecedents and consequences of students' achievement goals: A mediation analysis. *Learning and Individual Differences*, 28, 90-101. doi:10.1016/j.lindif.2013.09.005
- Duckworth, A. and Eskreis-Winkler (2013). True grit. *Observer*, 26(4). Retrieved from: <http://www.psychologicalscience.org/index.php/publications/observer/2013/april-13/true-grit.html>

- Duckworth A., Quinn, P. and Tsukayama, E. (2011a). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized test scores and report card grades. *Journal of Educational Psychology*, 104(2), 439-451. doi:10.1037/a0026280
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Dweck, C. S., & Bush, E. S. (1976). Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. *Developmental Psychology*, 12(2), 147-156. doi:10.1037/0012-1649.12.2.147
- Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental Psychology*, 14(3), 268. doi:10.1037/0012-1649.14.3.268
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256-273. doi:10.1037/0033295X.95.2.256
- Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, 3, 139-186.
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents? The impact of educational contexts on early adolescents. *American Journal of Education*, 99(4), 521-542. doi:10.1086/443996
- Edwards, A. R., Esmonde, I., Wagner, J. F., & Beattie, R. L. (2017). Learning mathematics. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 57–80). New York, NY: Routledge.
- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5-12. doi:10.1037/0022-3514.54.1.5
- Gehlbach, H. (2010). The social side of school: Why teachers need social psychology. *Educational Psychology Review*, 22(3), 349-362. doi:10.1007/s10648-010-9138-3

- Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs forcing*. Sociology Press.
- 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*(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Gorden, R. L. (1998). *Basic interviewing skills*. Waveland Press.
- Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology, 52*, 890-898.
- Hawkins, J. D., Von Cleve, E., & Catalano, R. F. (1991). Reducing early childhood aggression: Results of a primary prevention program. *Journal of the American Academy of Child & Adolescent Psychiatry, 30*(2), 208-217. doi:10.1097/00004583-199103000-00008
- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education, 87*(2), 106-124. doi:10.1177/0038040714525970
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149. doi:10.1257/aer.91.2.145
- Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin, 128*(5), 774-795. doi:10.1037/0033-2909.128.5.774
- Jonsson, A.-C., Beach, D., Korp, H., & Erlandson, P. (2012). Teachers' implicit theories of intelligence: Influences from different disciplines and scientific theories. *European Journal of Teacher Education, 35*(4), 387-400. doi:10.1080/02619768.2012.662636
- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., ... & Barmer, A. (2016). The Condition of Education 2016. NCES 2016-144. *National Center for Education Statistics*.
- Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology, 35*(3), 835. doi:10.1037/0012-1649.35.3.835
- Lapadat, J.C. and Lindsay, A.C. 1999: Transcription in research and practice: From standardization of technique to interpretive positionings. *Qualitative Inquiry, 5*, 64-

86. doi:10.1177/107780049900500104

- Law, W., Elliot, A. J., & Murayama, K. (2012). Perceived competence moderates the relation between performance-approach and performance-avoidance goals. *Journal of Educational Psychology, 104*(3), 806-819. doi:10.1037/a0027179
- Linnenbrink, E. A. (2005). The dilemma of performance-approach goals: The use of multiple goal contexts to promote students' motivation and learning. *Journal of Educational Psychology, 97*(2), 197. doi:10.1037/0022-0663.97.2.197
- Maehr, M. L., & Midgley, C. (1996). *Transforming School Cultures*. Westview Press.
- Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience, 7*, 230–241. doi:10.1093/scan/nsq100
- Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487-503. doi:10.1146/annurev.psych.56.091103.070258
- Midgley, C., Anderman E.M., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: a goal theory approach. *Journal of Early Adolescence, 15*(1), 90–113. doi:10.1177/0272431695015001006
- Miles, M. B., Huberman, A. M., Huberman, M. A., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Morgan, D. (Ed.). (1993). *Successful focus groups: Advancing the state of the art (Vol. 156)*. Sage Publications.
- Morgan, D. L. (1997). *Focus groups as qualitative research (Vol. 16)*. Sage publications.
- Morton, T. R., Ramirez, N. A., Meece, J. L., Demetriou, C., & Panter, A. T. (2018). Perceived barriers, anxieties, and fears in prospective college students from rural high schools. *The High School Journal, 101*(3), 155-176. doi:10.1353/hsj.2018.0008
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*(1), 33-52. doi:10.1037/0022-3514.75.1.33
- NAEP (2015). 2015 Mathematics & Reading Assessments [Website]. Retrieved from: <http://www.nationsreportcard.gov/>.
- Niemiec, C. & Ryan, R. (2009) Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*, 133-134.

- Ntoumanis, N. (2005). A prospective study of participation in optional school physical education based on self-determination theory. *Journal of Educational Psychology, 97*, 444-453.
- Padgett, D. K. (2016). *Qualitative methods in social work research (Vol. 36)*. Sage Publications.
- Pandit, N. R. (1996). The creation of theory: A recent application of the grounded theory method. *The Qualitative Report, 2*(4), 1-15.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research, 34*(5), 1189.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017
- Pelham, B. W., & Swann, W. B. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology, 57*(4), 672. doi:10.1037/0022-3514.57.4.672
- Rattan, A., Good, C., & Dweck, C. S. (2012). “It's ok — not everyone can be good at math”: Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*(3), 731-737. doi:10.1016/j.jesp.2011.12.012
- Riessman, C. K. (1993). *Narrative analysis (Vol. 30)*. Sage.
- Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450-461.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68. doi:10.1037/0003-066X.55.1.68
- Rydell, R. J., Shiffrin, R. M., Boucher, K. L., Van Loo, K., & Rydell, M. T. (2010). Stereotype threat prevents perceptual learning. *Proceedings of the National Academy of Sciences, 107*, 14042–14047. doi:10.1073/pnas.1002815107
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63-75. doi:10.3233/EFI-2004-22201
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., . . . Cohen, G. L. (2013). Deflecting the trajectory and changing the

- narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, 104(4), 591-618. doi:10.1037/a0031495
- Smeding, A., Darnon, C., Souchal, C., Toczek-Capelle, M. C., & Butera, F. (2013). Reducing the socio-economic status achievement gap at university by promoting mastery-oriented assessment. *PLoS One*, 8(8), e71678. doi:10.1371/journal.pone.0071678
- Sparks, S. (2013). Growth mindset gaining traction as school improvement strategy. *Education Week*, 33(3), 1.
- Steele, C. M. (1992). Race and the schooling of Black Americans. *The Atlantic Monthly*.
- Steele, C. M., & Aronson, J. (2000). Stereotype threat and the intellectual test performance of African Americans. In C. Stangor, & C. (Stangor (Eds.), (pp. 369-389). New York, NY, US: Psychology Press.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811. doi:10.1037/0022-3514.69.5.797
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of Qualitative Research*, 17, 273-85.
- Taylor, I.M., Ntoumanis, N. and Standage, M. (2008). A self-determination theory approach to understanding antecedents of teachers' motivational strategies in physical education. *Journal of Sport and Exercise Psychology*, 30(1), 75-94. doi:10.1123/jsep.30.1.75
- Turner, J. C. (2010). *Unfinished business: Putting motivation theory to the “classroom test”*. *The decade ahead: Applications and contexts of motivation and achievement*. Emerald Group Publishing Limited.
- Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, A. M., Kang, Y., & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology*, 94, 88–106. doi:10.1037/0022-0663.94.1.88
- Urduan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, 44(5), 331-349. doi:10.1016/j.jsp.2006.04.003
- Urduan, T., & Turner, J. C. (2005). Competence motivation in the classroom. *Handbook of Competence and Motivation*, 297-317.

- Vallerand, R. J., & Reid, G. (1984). On the causal effects of perceived competence on intrinsic motivation: A test of cognitive evaluation theory. *Journal of Sport Psychology*, 6, 94-102. doi:10.1123/jsp.6.1.94
- Vallerand, R. J., & Reid, G. (1988). On the relative effects of positive and negative verbal feedback on males and females' intrinsic motivation. *Canadian Journal of Behavioral Sciences*, 20, 239-250. doi:10.1037/h0079930
- Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology*, 50(3), 881. doi:10.1037/a0034325
- Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2011) Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Vera-Estay, E., Dooley, J. J., & Beauchamp, M. H. (2015). Cognitive underpinnings of moral reasoning in adolescence: The contribution of executive functions. *Journal of Moral Education*, 44(1), 17-33. doi:10.1080/03057240.2014.986077
- Wigfield A., Eccles J.S., Mac Iver D., Reuman, D., & Midgley C. (1991). Transitions at adolescence: changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, 27, 552-565. doi:10.1037/0012-1649.27.4.552
- Williams, J. (2014). How to best address the needs of African American children vulnerable to the continual widening of the achievement gap. *Social Work Research*, 38(3), 131-134. doi:10.1093/swr/svu022
- Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302-314. doi:10.1080/00461520.2012.722805
- 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(3), 374. doi:10.1037/edu0000098
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143(2), 804. doi:10.1037/a0033906

Table 2.1. Demographics of focus groups.

Group	n	Female	Male	Black	White	Hispanic	Multiracial
1	6	4	2	4	1	0	1
2	3	1	2	2	1	0	0
3	3	3	0	3	0	0	0
4	4	2	2	4	0	0	0
5	5	3	2	3	2	0	0
6	6	2	4	4	1	1	0
7	6	3	3	4	2	0	0
8	6	3	3	4	1	1	0
9	5	5	0	3	2	0	0
Total	44	26 (59%)	18 (41%)	31 (71%)	10 (23%)	2 (.05%)	1 (.02%)

Table 2.2. Summary of themes.

Theme	Student Beliefs
Everyone is Smart	<ul style="list-style-type: none"> • All students are smart • Intelligence is something that you have to work at • intelligence develops over time
The Smart Kids	<ul style="list-style-type: none"> • Some teachers believe that some students are smarter than others • Teachers help the smart kids more • Teachers call on the smart kids more • Students are stereotyped based on where they live • Some teachers believe that students in low income and/or rural areas are less intelligent than students in more affluent communities
Instrumental Support	<ul style="list-style-type: none"> • If a teacher helps you, it's because he/she believes you are smart • If a teacher allows resubmission of work, he/she believes you can improve with more effort
Emotional Support	<ul style="list-style-type: none"> • When a teacher shows that they care about you, it also shows that they believe you can do well in their class

Appendix A

Guide 1: Original Interview Guide for Students

1. What do you think the word “stereotype” means?
2. What types of stereotypes do you think exist?
3. What types of stereotypes do you think people might have about you?
4. What types of stereotypes do you think people might have about other students in your school?
5. Do you think teachers have stereotypes about you?
6. Do you think teachers have stereotypes about other students in your school?
7. Do you think stereotypes can affect how well a student does in school?
8. Do you think people are either born smart or they aren't? Or do you think people can change how smart they are?
9. Why do you think people can't change how smart they are? OR Why/how do you think people can change how smart they are?
10. Do you think your teachers think that students can change how smart they are?

Guide 2: Actual Interview Questions with Students

1. Can you think of a teacher who makes you feel that you can do well, even if the work is hard?
 - a. What is he/she like?
 - b. What does he/she do?
 - c. What does he/she say?
 - d. How does that make you feel?
 - e. How does that make you feel about [subject]?
2. If someone is doing really well in school, why is that?
3. If someone is not doing well in school, why is that?
 - a. Can they do better? How?
4. How do your teachers respond when you make a mistake?
 - a. How does that make you feel?
5. Have you ever heard the word stereotype? What do you think it means?
 - a. What are some stereotypes you know about?
6. Do stereotypes happen in school? How?
 - a. Are all students treated the same?
 - b. How do your teachers show fairness / favorites?
 - c. How do you know who the “smart kids” are?
7. Do teachers talk about stereotypes?
 - a. What kinds of things do teachers do and say?

PAPER III

TRANSLATING GROWTH MINDSET AND STEREOTYPE THREAT THEORIES TO THE CLASSROOM: A QUALITATIVE ANALYSIS

A pervasive achievement gap exists such that low income and minority students routinely underperform on measures of academic performance (Vanneman, Hamilton, Anderson, & Rahman, 2011; Dee and Jacob, 2011). There is little dispute that the achievement gap is a serious concern, but education reform strategies often do not address the root causes of underachievement (McGuinn, 2012). Most visibly, legislative policies such as No Child Left Behind (NCLB) and Race to the Top (RttT) have been enacted to reduce achievement gaps, yet they have proven ineffective at reaching students of the highest need (Krieg, 2008; Rothstein & Jacobson, 2006). Certainly, the achievement gap has remained central to the discussion of education reform, but attempting to boost cognitive skills without addressing underlying non-cognitive issues can exacerbate underperformance, and heightened accountability can cause schools to “push out” low performers – however unintentionally – by focusing on testing data (Brown, 2013; Duckworth, Quinn, & Tsukayama, 2011; Glennie et al., 2012). In contrast, research on the non-cognitive, social and emotional determinants of learning has shown promise for increasing students’ engagement, motivation, and academic performance (Heckman & Rubinstein, 2001). Two well-evidenced theories – growth mindset and stereotype threat – tie students’ beliefs about the nature of intelligence to increased achievement and have been applied as non-cognitive interventions

in schools (Blackwell, Trzesniewski, & Dweck, 2007; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Good, Aronson, & Inzlicht, 2003; Sherman et al., 2013).

Growth Mindset and Stereotype Threat as Interventions

Growth mindset and stereotype threat are distinct but related social phenomena. Broadly, both theories involve students' perceptions of school and of themselves, including how they make sense of their abilities and experiences, such as interactions with teachers (Ross & Nisbitt, 1991). Conceptualized as an intervention conducted in schools, teaching students about the growth mindset and the malleability of intelligence has been linked with increased performance on standardized exams and with better course grades (Blackwell et al., 2007). Iterations of this intervention have involved supporting students with mentors trained in growth mindset theory (Good et al., 2003); teaching students to view critical teacher feedback as indicative of the teachers' expectations that students can improve with effort (Yeager et al., 2014); teaching students about the value of pursuing challenging academic goals (Oyserman, Bybee, & Terry, 2006); and providing various supporting information on the growth mindset to bolster intervention effects (e.g., celebrity endorsement, scientific data on the negative effects of a fixed mindset; Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015). All of these strategies have been linked with increased student achievement, and importantly, effects were often strongest among low performers and students of color (Pauneksu et al., 2015; Yeager et al., 2014; Yeager et al., 2016).

In a similar vein, stereotypes about intelligence represent a specific type of fixed mindset belief, where students may link stable, low ability with an aspect of their identities – commonly, race/ethnicity or gender (Good et al., 2003; Steele and Aronson, 1995). If students internalize these stereotypes, they may feel that their identity precludes academic

achievement and may disengage as a result (Aronson, Blanton, & Cooper, 1995). The cognitive burden of stereotype threat can also impede students' ability to learn, negatively affecting academic performance over time (Mangels, Good, Whiteman, Maniscalco, & Dweck, 2012). School-based interventions to reduce stereotype threat have involved brief written affirmations where students reflect on their personal values just prior to events that are evaluative and most likely to elicit threat (e.g., at the end of the grading period, days before a standardized test). Conducting several brief affirmations throughout the year appears to disrupt the negative effects of threat by affirming students' identities in the school context. The values affirmation task has been associated with long-term changes in academic performance, boosting black and Hispanic students' grades for up to three years and cutting achievement gaps nearly in half (Bowen, Webber, & Wegmann, 2013; Cohen et al., 2009; Sherman et al., 2013).

Teacher Influence on Students' Intelligence Beliefs

Cohen and colleagues (2009) believe the values affirmation task has shown effects on students' academic achievement because it increases their sense of belonging in school, which may be under threat from negative stereotypes about their ability. Identity affirmation may be an especially critical strategy for teachers in historically low performing schools and/or those with diverse student populations, where students of color are likely to experience a high amount of threat (Hanselman, Bruch, Gamoran, & Borman, 2014). Yeager and colleagues (2014) identify that, for students of stereotyped identities, the connection between their sense of belonging and mindset beliefs are doubly affected by their interactions with teachers; for example, students of color may interpret a teachers' low expectations as

evidence that the teacher believes they cannot get smarter *and* as evidence that their teacher harbors a negative stereotype about their identity.

In contrast, when students form trusting and positive relationships with teachers, they are less likely to interpret interactions (e.g., critical feedback about mistakes) as an indication of their teachers' negative biases about their ability, promoting sense of belonging, positive intelligence beliefs and academic performance (Bryk & Schneider, 2002; Yeager et al., 2014). Based on evidence from growth mindset and stereotype threat interventions, Yeager and colleagues (2014) developed the strategy of "wise" feedback, which conveys that a teacher has both high expectations for a students' ability (reducing stereotype threat) and believes that students can improve (promoting growth mindset). For example, students received critical feedback on an essay with a note from their teacher that contained either wise feedback ("I'm giving you these comments because I have very high expectations and I know that you can reach them.") or a generic comment ("I'm giving you these comments so that you'll have feedback on your paper.") Implementation of wise feedback in low income schools has shown increased achievement for black but not white students, reducing achievement gaps (Yeager et al., 2014). Disparate intervention effects in student achievement by race/ethnicity indicate that students of color process academic feedback from their teachers through the lenses of both mindsets and stereotypes of ability. Importantly, prior work that has tested the effects of either message separately (i.e., high standards or ability to reach them) compared to their combined effects shows that both elements are necessary for affecting school engagement among stereotyped students (Cohen & Garcia, 2008; Cohen, Steele, & Ross, 1999, Cohen & Steele, 2002).

Implementation in Schools

Despite a growing body of research on growth mindset and stereotype threat strategies, which details the importance of student-teacher interactions in promoting academic achievement, teachers have yet to be leveraged as the primary implementers of intervention strategies. Both theories have largely been tested in schools with activities or programs that are separate from regular classroom instruction. Growth mindset interventions have been implemented by research teams (e.g., Blackwell et al., 2007; Good et al., 2003) or computer programs (e.g., Paunesku et al., 2015; Yeager et al., 2016) rather than with teachers as implementers, and in stereotype threat interventions, though teachers often distribute the values affirmation task, they are relatively removed from the intervention process. Often, teachers are blinded from the intervention process entirely (Blackwell et al., 2007; Borman, Grigg, & Hanselman, 2016; Bowen et al., 2013; Cohen et al., 2009; Pauneksu et al., 2015; Protzko & Aronson, 2016; Sherman et al., 2013; Yeager et al., 2014; Yeager et al., 2016).

Limited supporting resources are available to guide classroom implementation, and little research has explored how teachers generate their own classroom strategies when exposed to information on growth mindset and stereotype threat theories. In particular, growth mindset theory has been criticized for being pitched to classroom teachers as simply telling students to “have a growth mindset,” rather than focusing on the classroom processes and specific behaviors that contribute to mindset development (Edwards, Esmonde, Wagner, & Beattie, 2017). At the same time, only a small number of empirical studies have looked at the effects of growth mindset training on teachers’ behaviors. Rattan and colleagues (2012) found that teachers who learned about growth mindset theory reported that they would be less likely to give students fixed mindset feedback (e.g., “it’s okay – not everyone is a math person”) or to suggest disengagement for struggling students (i.e., dropping the class), but

these findings were based on teachers' predictions of behavior rather than direct observation. A recent study by Sun (2018) provides some guidance on growth mindset strategies through a literature-based framework for classroom observations of growth-mindset aligned behaviors. She identifies potential strategies for future investigation, including mixed-ability grouping; encouraging and normalizing mistakes; and grading policies, such as the ability to submit work multiple times. Though Sun (2018) does not assess the influence of teachers' strategy choice on student engagement and achievement, she reports that teachers had developed their own "best practice" strategies for implementing growth mindset theory (Sun, 2018).

Factors Affecting Teachers' Implementation

As more research is dedicated to identifying classroom methods that promote growth mindset and reduce stereotype threat, it is important to consider that implementation may be more complicated than simply providing teachers with a defined curriculum or list of potential strategies. Schools have a complex ecology that influences teachers' strategy use and their fidelity and consistency of intervention implementation (Durlak, 2015).

Frameworks from implementation of school-based interventions outline a host of critical factors that affect whether and how successfully teachers can implement preventative strategies within the school context. For example, Hall (1978; 2013) developed the Concerns Based Adoption Model (CBAM), a framework that identifies three major implementation concerns teachers experience: (1) self-concern (e.g., Can I do it? What will I need to learn?); (2) task-concern (e.g., Will it take up all of my time to do it?) and (3) impact-concern (e.g., Will my students benefit from what I'm doing?).

In the stage of self-concern, teachers must feel that they understand their role in implementing the intervention, as teachers are more likely to continue strategy use if they feel that they have clarity over what successful implementation looks like (Hall, 2013). Yet, direct instruction and training (e.g., professional development) has shown mixed results in promoting teachers' internalization of implementation procedures and processes, often failing to promote teachers' knowledge and confidence for enacting strategies in their classrooms (Fixsen, Naoon, Blase, & Friedman, 2005; Noell, Duhon, Gatti, & Connell, 2002; Wickstrom, Jones, LaFleur, & Witt, 1998). In contrast, research that has focused on leveraging teachers' autonomy has shown promise for increasing teachers' self-efficacy, fidelity and consistency of implementation. Including teachers in the decision making process through a collaborative model – where teachers contribute to the intervention selection, design and rollout – including addressing potential barriers – increases the likelihood not only that teachers are familiar with strategies but are confident in how to use them (Frey, Lee, Small, Seeley, Walker, & Feil, 2013; Kelleher, Riley-Tillman & Power, 2008; Ringwalt, Ennett, Johnson, Rohrbach, Simons-Rudolph, Vincus, & Thorne, 2003). What's more, teachers have extensive knowledge of their school and classroom contexts, and they can often offer ideas to tailor interventions in ways that *increase* effectiveness, simultaneously promoting teachers' confidence, perceptions of effectiveness and commitment to implementation (Frey et al., 2013; Durlak, 2015).

Related to the task-concern domain of the CBAM framework (i.e., do I have time?), teachers face increasingly varied work demands - often including the implementation of multiple interventions simultaneously - which can amplify stress and reduce their effectiveness at implementing any one strategy in particular (Dorman, 2003; Lasky, 2005;

Woods, 1999). Indeed, teachers report working nearly 20 hours a week outside of school to keep up with professional demands (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). Lack of time and high levels of stress can cause teachers to fatigue in their strategy use over time and can be strong limiting factors in whether teachers will agree to take on additional responsibilities in the classroom (Betoret, 2009; Santavirta, Solovieva & Theorell, 2007). The issue of burnout is often exacerbated in struggling schools, where teachers often expected to “do more with less” - supporting high need students with few resources or support staff (Durlak, 2015; Jennings & Greenberg, 2009). As Hall (1978; 2013) outlines, if teachers feel that new strategies will substantially increase their workload, they are less likely to commit to and maintain quality implementation over time, and the reverse is also true.

Finally, in the stage of impact-concern, teachers’ sustained implementation is also contingent upon how teachers perceive the strategies’ effectiveness, whether they feel their efforts are paying off, and whether the strategies align with their personal beliefs (Collie, Shapka, & Perry, 2013; Hall, 2013; Helfrich, Weiner, McKinney, Minasian, & 2007). To consider implementing an intervention, teachers must believe that the strategies will be effective for their students, if implemented well (Von Brock & Elliott, 1987). Though there is evidence that teachers are more likely to implement a strategy when presented with evidence supporting its effectiveness (von Brock & Elliott, 1987), the feedback loop between teachers’ implementation and effects in their own classrooms is also be a powerful reinforcer for their continued implementation; in other words, if teachers feel that they are seeing positive results from their strategy use, they are more likely to continue or intensify their efforts, creating a positive “up spiral” of intervention effects (Collie, Shapka, & Perry, 2013). Hall (2013) describes that change is a personal process, where teachers attend to the consequences of

how strategies play out for themselves – and for others – as they make ongoing evaluations about the worth of an intervention.

Though research specific to teachers' strategy use related to intelligence beliefs is limited, in recent qualitative research about the application of growth mindset theory in the classroom, teachers reported that, after being exposed to information about the theory, they developed their own strategies for implementation and tweaked their strategies over time, based on their relationships with students and which strategies they felt were working (Bethge, 2018). For example, some teachers felt students fatigued from hearing about the growth mindset too much or that talking about the growth mindset was not as successful as modeling it through their own behavior or relationships with students (Bethge, 2018). It is likely that self-, task- and impact-concerns such as teachers' beliefs, workload, stress and autonomy in strategy use also affect how teachers choose to implement growth mindset and stereotype threat strategies in their classrooms, and exploration of these factors will provide insight for how to best translate theory into practice.

The Current Study

To gain a better understanding of how growth mindset and stereotype threat theories are used by classroom teachers, more research is needed that examines teachers' strategy use in the classroom and explores the effects of their implementation. It is especially important to consider how teachers apply these interventions in high need schools, where both fixed mindset beliefs and stereotype threat are likely to be highest (Claro, Paunesku, & Dweck, 2017; Hanselman et al., 2014). Also, teachers' strategy use in high performing schools may be altogether different from those in high need schools, and well-evidenced school-based interventions often are poorly implemented and fail to replicate positive effects in

underresourced schools (Gottfredson, Jones, & Gore, 2002; Tolan, Gorman-Smith, & Henry, 2004; Ransford, Greenberg, Domitrovich, Small, & Jacobson, 2009). Yet, these schools have the greatest need for intervention, and more research must be dedicated to understanding successful implementation of growth mindset and stereotype threat interventions – specifically when teachers are targeted as implementers. To fill this gap, this study will use classroom observations and qualitative interviews with middle school teachers in a high need school district that has provided teachers with training on promoting growth mindset and dispelling intelligence-based stereotypes. It will also investigate teachers’ opinions and use of these strategies. This study began with a focus on the following research questions:

1. How do teachers apply growth mindset and stereotype threat strategies in their classroom and in their school?
2. What are teachers’ perceptions and attitudes about using growth mindset and stereotype threat strategies in their classrooms and in their school?

However, through the process of teacher interviews, it became clear that, in discussing the implementation of growth mindset and stereotype threat strategies, teachers were focused primarily on the outcomes of using these strategies – for both themselves and for students. Thus, our third research question was modified to be:

3. What do teachers believe are the outcomes of using growth mindset and stereotype threat strategies in their classrooms and in their school?

Methods

Population

The sample for this study was drawn from core content middle school teachers (i.e., math, science, English/language arts and social studies) at a public charter school in a rural

county in the Southeast United States. The school is positioned in a low-income county of fewer than 25,000 people, 20% of whom are living below the poverty line. The student population is 75% black and nearly 25% white, with few children identifying as Hispanic, Asian or other racial/ethnic groups. Approximately 75% of students qualify for the Free and Reduced Lunch Program, an indicator of poverty. While the public schools in the county have historically struggled with underperformance on standardized tests and other measures of school accountability, this charter school has demonstrated considerably better student achievement – with students outperforming the state average in many subjects and grade levels. The charter school belongs to a larger, national network of charter schools which share common structure in school procedures, policies and norms, including character traits and classroom procedures.

Recruitment and School Context

Data for this study were collected and analyzed under approval of the Institutional Review Board at the University of North Carolina at Chapel Hill. This charter school was chosen to obtain a purposive sample of teachers who work in a high need community and who received ongoing resources and professional development training around growth mindset and stereotype threat (Padgett, 2016). The school's mission includes equipping students with the skills and character necessary for college success. During initial conversations with the primary researcher, school administrators shared that, as part of fulfilling the school's mission, teachers were actively encouraged by administrative staff to emphasize the malleability of intelligence (e.g., promote a growth mindset) and discuss social issues and dispel stereotypes in class (e.g., address stereotype threat). In follow-up conversations, administrators shared that before the 2017-2018 school year began, all

teachers on staff participated in a week of professional development training at a national summit hosted by the charter network and a week of professional development led by the school's regional executive director. Throughout their training, all staff read and discussed five articles related to growth mindset and stereotypes, covering topics such the intersectionality of identity, how students' effort is affected by inequity in schools, and the disproportionate suspension, expulsion and incarceration of black males. All staff also read "Letter from a Birmingham Jail" by Martin Luther King Jr., as well as an excerpt from *The Second Sex* by Simone de Beauvoir – two texts that discuss racial and gender inequity. In addition, all staff participated in a film study of the documentary *13th*. One administrator shared that staff watched the documentary together and paused at various points to discuss "the impact of our school systems on the identity and trajectory of black and brown students". In addition to the training provided by national and regional leaders, all teachers received two days of professional development training led by school administrators, which focused heavily on role playing scenarios of responding to students in ways that supported growth mindset and dispelled racial stereotypes. Strategies were listed in training materials as "affirming, encouraging, inclusive and collaborative".

In term of growth mindset, this school emphasized that effort is the key to learning and to future success (Blackwell et al., 2007). All classrooms featured signs referencing the related concept of grit, which involves working hard and overcoming obstacles in the pursuit of a goal (Duckworth, Peterson, Matthews, & Kelly, 2007; Duckworth & Seligman, 2005). Classrooms also had consistent signage that reinforced the importance of effort, such as "there are no shortcuts" and "we are never done". Across classrooms, students and teachers used consistent language that reflected growth mindset beliefs. For example, they frequently

used the word “earning” to describe the connection between effort and outcomes (privileges were earned through students’ effort, grades were earned through students’ effort, etc.). Teachers also often used the term “with quality” to make clear, for example, that homework was not just to be completed but completed *with quality*. The cafeteria was decorated with pennants from a variety of universities and showcased T-shirts that depicted the graduation dates of all prior classes. In addition, an entire wall was painted with a mural of a mountain climber placing a flag at the summit, as a reflection of the hard work necessary to gain acceptance to a four-year college or university.

In addition to visual messaging, to reinforce growth mindset messaging around the importance of effort, administrators had set a school policy around makeup work, where teachers were required to offer students the opportunity to improve their grades by redoing assignments or retaking assessments. Teachers had freedom to determine how “retakes” would operate in their classrooms and to deny students the opportunity based on their own explicit policies (e.g., no homework retakes after the unit test, retakes on quizzes and homework but not exams, etc.), but they were required to provide some opportunity for students to make improvements on prior work. Teachers were also expected to make themselves available to students outside of class, and they often started class by reminding students of the days they would be available during lunch or after school. Because the school day runs from 8 a.m. to 5 p.m., staying after school was a considerable commitment for both students and teachers.

Sample

In total, the school had 12 core content teachers in grades 6-8, and all 12 teachers were invited to participate. Teachers were recruited through an initial informational email

sent by the primary researcher, who stressed that participation was voluntary. Five teachers expressed interest in participation, and observations and interviews were scheduled via email. Remaining teachers were contacted in person on their planning periods. An additional three teachers consented to participate after the in-person discussion with the primary researcher, and observations and interviews were scheduled. In total, eight of the twelve teachers (67%) consented to participate. All eight teachers participated in the classroom observation. However, due to inclement weather prior to a holiday break, only seven teachers were able to be interviewed (58%). The sample includes both beginning and experienced teachers across all grades and content areas, as well as male and female teachers who identified as either black or white. Because this middle school has only one teacher in each content area at each grade level, to reduce the likelihood of deductive disclosure, the sample will not be discussed in further detail.

Data Collection: Observations and Individual Interviews

Teachers were observed for approximately two hours, on two separate occasions, each for a full class period over the course of a week and a half. Teachers were informed that the intention of observations was for the researcher to gain greater context for their classrooms and that the purpose of the study was related to motivational strategies for students. In scheduling the observations, teachers frequently commented about how certain classes were more difficult than others or that students were particularly sluggish or excited during certain times of day (e.g., after lunch, last period). To increase the likelihood of seeing the full range of teachers' responses to students, observations were conducted during different class periods and, when possible, during different times of day. To ground classroom observations in research on growth mindset and stereotype threat theory, field

notes were guided by the use of the Sense of Competence subscale of the Effective Classroom Checklist, developed by Dr. Natasha K. Bowen, a substantive expert in growth mindset and stereotype threat (see Appendix A).

Follow-up interviews occurred during teachers' planning periods and lasted for approximately an hour, depending on teachers' availability. One teacher had a prior obligation that limited the interview time to thirty minutes. Interviews were conducted using a semi-structured interview guide (see Appendix B). After the completion of the interviews, teachers were informed that the intention of the study was to learn specifically about how they apply growth mindset and stereotype threat theories in their classrooms. Teachers were given the opportunity to revoke their consent if they were uncomfortable with the full details of the study, but no teachers requested to do so. A second consent form was signed after the debriefing process.

Data Management and Analysis

Data collection resulted in 293 minutes of interview data and 30 pages of field notes, including 16 classroom observation checklists (two checklists per teacher), 11 pages of additional classroom observation notes, and three additional pages of general notes about the school. Interview data were professionally transcribed verbatim from audio recordings of individual teacher interviews. Transcripts and observational data were managed in ATLAS.ti (v8; Friese, 2018) and primarily analyzed inductively, using thematic analysis (Braun & Clarke, 2006). The use of the observation checklist introduces a deductive approach, relying on existing research on teacher behaviors to guide classroom observations. However, observations were not limited to the strategies included on the checklist. Often, classroom observations provided specific examples of broader strategies listed on the checklist. For

example, growth mindset strategies listed on the checklist include praising effort and rewarding improvement. If a teacher acknowledged effort or improvement through praise, their specific language was noted. If a teacher implemented a more formal procedure (e.g., changing grades), details of their strategy choice were also noted. Observation data were analyzed using the codes that emerged from the strategy checklist and from teacher interviews, as analysis of observation data occurred after the analysis of teacher interview data. After observation data were coded, data were organized such that interview segments and observational data with the same codes were matched together and linked via teacher ID number. To triangulate among data sources, it was noted when teachers' observational data corroborated data from their interviews. Analysis was focused on capturing the range of teacher responses to identify the variety of strategies used by teachers and their experiences with implementing growth mindset and stereotype threat theories.

Thematic analysis is suitable for these data because it can be used to develop patterns and themes inductively from participants' experiences, without the need for additional theory-building steps of grounded theory analysis (Boyatzis, 1998; Braun & Clarke, 2006; Daly, Kellehear, & Gliksman, 1997). Thematic analysis includes six stages: (1) familiarization with data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes and (6) producing the report (Braun & Clarke, 2006).

Familiarization with data. Because these data were transcribed professionally and not by the primary researcher, it was important to ensure adequate familiarization with the data before proceeding with further analysis. The primary researcher listened to recordings of all interviews, checking transcripts against audio recordings, adjusting any errors in the

transcripts and noting initial impressions of the data. Transcripts were read through again to deepen familiarity and note any additional ideas or patterns that emerged (Bird, 2005; Lapadat & Lindsay, 1999; Riessman, 1993). Through familiarization with that data, it was not immediately clear how the strategies teachers discussed aligned with the research on growth mindset and stereotype threat theories; therefore, data reduction was not conducted, so that all data could be analyzed in more detail (Miles & Huberman, 1994).

Generating initial codes. Next, data were analyzed through initial rounds of line-by-line coding, resulting in nearly 150 separate codes (Braun & Clarke, 2005). Initial codes were assessed for consistent use across transcripts, with frequent reference to the data associated with each code (Glaser, 1992). Codes were modified iteratively and recoded, combined or divided, based on their use across transcripts (Cohen, Manion, & Morrison, 2011; Creswell, 2009; Punch, 2009). To help prevent drift, all transcripts were reviewed again and recoded as necessary, resulting in a reduced list of 131 codes. A list of reduced codes and their definitions was developed, and all transcripts were checked using constant comparative analysis to assess whether definitions were sufficiently well-developed and descriptive (Glaser, 1992).

Searching for and reviewing themes. The second and third stages of thematic analysis involve an iterative process of searching for themes emerging from the data, reviewing the data to check the accuracy of emerging themes and modifying themes as necessary. During these stages, related codes were grouped together into initial categories. Codes within categories were organized into thematic maps, focusing on the relationships among codes (e.g., codes that contradict each other, codes that are related to each other, codes that are subcomponents of other codes; Braun & Clarke, 2006). Through the thematic

mapping process, and by returning to associated data segments, codes were reorganized and recategorized. A second set of maps were developed to organize the relationships among codes in each category.

Defining and naming themes. During this stage, codes, thematic maps and data were reviewed to develop rich descriptions of themes (Braun & Clarke, 2006). Codes were reorganized and modified based on developing descriptions of themes and their associated data segments. Codes that appeared to be disparate from theme descriptions were explored further during negative case analysis – a process that involves searching for data that do not support or contradict emerging themes or patterns (Padgett, 2016). Negative case analysis can be used to improve the rigor of qualitative analysis by promoting the accuracy of developing themes and refining or revising themes based on negative cases that may arise (Allen, 2017).

Producing the report. The final stage of thematic analysis occurs during the generation of the research report, where themes may be modified and further developed through the process of selecting illustrative quotes and stories and writing up study findings (Braun & Clarke, 2006). In accordance, findings in this study were adjusted while writing the study report. Specifically, third level themes were generated to consolidate and organize themes on a broader level.

Rigor and Trustworthiness

Several other measures were taken to promote the rigor and trustworthiness of the findings presented here. First, interview data were triangulated with field notes from classroom observations. Observation notes were analyzed in ATLAS.ti (v8; Friese, 2018), separately from interview transcripts, using the same codes applied to interview data. While

reviewing and defining themes, teachers' coded interview data were compared with observational data, which revealed considerable consistency between (a) what teachers did in their classrooms and what they discussed in their interviews and (b) among what was observed across teachers in their classrooms. Corroborating data from classroom observations were organized with theme codes and descriptions and were used to further develop theme descriptions. After these data were organized into broader themes, all participating teachers were contacted for member checking. One teacher was unable to be reached because the school email address had been disconnected. No further contact information was sought for this teacher, to preserve confidentiality of participation. Two study participants consented to member checking and commented on the validity of study themes and descriptions (Padgett, 2016). While no themes changed as a result of member checking, adjustments and additions to the explanations of themes were made based on participant feedback. Finally, the primary researcher documented all stages of analysis through analytic memos and a detailed audit trail, to track patterns in the data, to more clearly define codes and to synthesize the relationships among codes, code families and negative cases (Padgett, 2016; Pandit, 1996; Strauss & Corbin, 1994).

Results

Analysis of interview and observational data resulted in two categories of themes regarding (a) what strategies teachers used to support positive intelligence beliefs in their classrooms and (b) the consequences of implementing these strategies. The major theme related to strategies that teachers used was to support *productive struggle*. A consequence of promoting productive struggle was that teachers felt they needed to spend a lot of time *managing retakes* – the school-wide assignment retake policy that was designed to encourage students'

effort and improvement. See Table 3.1 for the organization and definition of these themes. Each of these themes will be described in greater detail in the following section, starting with a definition of the theme, followed by examples of strategies teachers' used in their classrooms and their experiences with implementing these strategies. Throughout this section, teachers will be referenced by an identifying number. Data from classroom observations will be indicated first with the teacher's identifying number and second with the observation number – to indicate whether the data were collected during the first or second classroom observation. For example, if observational data were collected during the second observation of Teacher 6, the data will be referenced as (6-2).

Strategy: Productive Struggle

The theme of *productive struggle* summarizes some common strategies teachers used to implement growth mindset theory in their classrooms. To support students' productive struggle, teachers used three primary strategies: They provided challenging work, encouraged students to make mistakes, and – most importantly – encouraged students to learn from their mistakes. Teacher 2 described productive struggle as a process of getting students comfortable with not immediately mastering concepts. As Teacher 5 stated, “It's like, we're not there *yet*, you don't understand *yet*, you're not where you need to be *yet* [emphasis added].” Teacher 3 described that she spent a lot of time at the beginning of the school year encouraging students to be comfortable with productive struggle:

One of my bigger struggles at the beginning of this school year was the productive struggle - getting kids to be okay with being wrong. It's like, it's okay if we fail, it's okay if we're not right the first time. Even now, [students are] like, ‘[Teacher 3 name] am I right? Am I right?’ And I'm like, ‘I don't know, are you?’

Other teachers also stressed that it was important to teach students that it is okay to make mistakes, so long as they work hard to learn from their mistakes. Teacher 1 said, “I think that

something I get students to do is trusting the process of, it's all right right now for making mistakes, but what's not all right is lack of effort.” In describing productive struggle, teachers often paired the discussion of mistakes with words like “intentional” and “purposeful”.

Teacher 2 stated:

If you're just making mistakes, if you're woefully making mistakes, then it's just like, you're not becoming a better person...It's creating that culture of: this is a place where I can purposefully make mistakes and grow from them. So, I think for me, it's just that growth mindset of like, being able to come into the environment and make purposeful mistakes that will enable you to be better and try to build yourself back up. I think that's a big part of keeping students motivated.

Teacher 2 developed a classroom strategy to support the process of productive struggle (2-1). He had students complete a pre-test that closely resembled their upcoming unit test. When students struggled with questions, he reminded them that the purpose of the pre-test was to make mistakes and that, to prepare for the test, they would spend the next few days reviewing their work and fixing their errors. In his interview, Teacher 2 explained that students were given the opportunity to self-grade their pre-tests so that they did not have to share their mistakes with others if they were not comfortable doing so. In class, he encouraged students to make mistakes and said, “mistakes are part of the learning process” (2-1).

Teacher 7 explained that she developed a classroom acronym to encourage students to learn from their mistakes:

I use an acronym [suggesting mistakes are good] - as long as you learn from it. I just build a lot of that into my culture as far as, like, mistakes are cool, mistakes shouldn't be something that you should be afraid of. But then the last piece is - as long as you learn from them.

Teacher 6 collected several student papers while her class was working independently and projected the papers on the board using a document camera (6-1). She requested the

attention of all students and reviewed several mistakes with the whole class. In her interview, she referenced this process as a way to help students learn from their mistakes:

So, [if] they do problems incorrectly, we tend to just talk about what went wrong in the process. Sometimes we'll put their work on a document camera and we'll all give feedback. I don't get mad, I just [say], 'Okay where is the mistake? Class, can we get this person feedback?'

Similarly, Teacher 4 (4-1) had students exchange papers to provide feedback on introductory paragraphs and thesis statements. She encouraged students to help their peers become stronger writers and had several students write their thesis statements on the board to be evaluated and improved by the whole class.

In addition to informal opportunities to help students learn from their mistakes, teachers described that the school-wide policy around assignment "retakes" was implemented as a way to model the process of productive struggle. Students were allowed to complete and resubmit assignments multiple times in order to learn from their mistakes and, accordingly, improve their grades. Teacher 1 described how each teacher could "tweak" the policy but that it was an expectation that they would provide retakes:

So there's a school-wide thing that every grade and every teacher kind of tweaks to their specific content in their specific class. The expectation for all of the teachers is staying after school at least once a week to help with those retakes and to address the makeup work.

Teacher 3 discussed how she purposely graded students quite harshly on their first submission of assignments so that any mistake immediately reduced students' grade to a 50%, regardless of their actual score. She explained that this policy was intended to encourage students to complete their work as accurately as possible the first time around:

The first week they hear about it they're like, 'No, are you kidding me, a 50%?' It makes them mad because they're like, 'I got one thing wrong', and I'm like, 'Yeah, but you're going back and correcting.' And I see less and less mistakes from those students throughout the year.

Teacher 2 described his retake procedure:

So like the test can't be retaken, but quizzes if you got under an 89, you can retake it and then what happens is you take the average of the two scores and that's your final grade. And I kind of like it that way because it makes a student think like, okay, if I'm at like an 88, if I don't do better, then I may not want to retake this quiz because I could score lower. So, I think it's a really neat strategy because [for students] it's like, do I want to do it? I'm gonna do it, let me make sure that I'm studying where I need to, what I need to study to be ready for it.

In contrast, Teacher 7 allowed students to retake tests and quizzes, but required that students first complete corrections on missed test questions:

They know they have to do corrections. They have to get feedback on their corrections, on their previous quiz. And then they have to stay for some additional help. Now that additional help may be with me, it may be with another student, but they have to get some extra help [for me] to know that they are working.

These excerpts demonstrate that to support productive struggle, teachers encouraged students to feel comfortable learning from mistakes while emphasizing that mistake-making should be purposeful – creating an opportunity for students to learn from their errors. In addition to messaging the acceptability and utility of mistakes, teachers often created opportunities for students to analyze mistakes by reviewing their own work or by debriefing student work as a class. Correcting and resubmitting assignments or retaking assessments was another common strategy teachers used to create productive struggle in their classrooms.

Consequence: Managing Retakes

The theme of *managing retakes* summarizes strategies that teachers use to combat unintended consequences of implementing the school-wide retake policy – namely, instances where the policy encouraged student behavior antithetical to productive struggle. Teachers discussed retakes as a “Catch-22” when it came to promoting motivation and engagement. On one hand, teachers wanted struggling students to know that they could improve and fix

their mistakes. On the other hand, teachers spoke with great frustration about students abusing the policy, turning in “trash” work to meet a deadline, knowing that they would have time to improve their grades later. Teacher 1 stated:

Focusing back on motivation, there is somewhat of a Catch-22 with students being like, well I can just make it up. Or I'll get around to that at the end of the quarter. And it's a very tough kind of line to draw about here's the makeup work and then how you're addressing students not putting in effort in class, did you earn a retake, so to speak.

Teachers 4, 6 and 7 stated that they only implemented retakes because it was school policy.

Teacher 6 said:

I just feel some type of way about retakes in general. But I've always felt that way. It's just kind of hard as a teacher because you know you supposed to be on the bandwagon, and, I'm not really on that bandwagon. I mean, I get it, but it kinda gives them a crutch.

As Teacher 4 described it:

We have a policy as far as retakes, you can set the circumstances for your retakes, but we should allow them to retake. I'm like, no. You don't get to retake everything and, especially with this [group] right now, they have the mentality of I'm gonna fly through this because I know you gonna let me retake it...So if you already know in your mind that, oh [Teacher 4] ain't playing, she ain't gonna let me do [retakes] no more. If you already come into this with that, you gonna go ahead and put your best foot forward to begin with, versus the lemme just give you some trash so I can say I turned in something, so I can get the opportunity to redo it...it becomes a habit of let me not give 100 percent first because I can put the rest to it later. That's what's frustrating.

Teacher 4 went on to say that, although the retake policy encouraged some students to slack off, she believed that students who worked hard should be given the chance to improve. She discussed a student who was far below grade level in reading and writing:

She behind, but for her, it's not from lack of effort, it's not lack of work ethic, it's because of her comprehension problem. But because of her effort I be like, you know what baby, go back and rethink this, look over it, re-fix it. Those are the people who I feel like deserve the second chances.

Teacher 7 said:

The only reason I even allow retakes is we as a school has said we will allow retakes, although I personally disagree with the ideology of retakes. [Students] purposefully don't prepare, or they purposefully don't do an assignment because they know at the end of the quarter or later on down the road, they understand that there's this opportunity for a second chance, whether it be Imma give you a second assignment, Imma give you an opportunity to retake. So I personally disagree with it. I only comply with it because it is a school policy that we do.

Teacher 3 said:

I think that's the biggest question I have a lot of the time. The effort the first time is generally good, although they're all like, 'Oh, can I correct this?' And I'm like of course you can, you can always correct it. I think it gives them more of like, okay it's not the end of the world if I don't get it right. Which I like them having - it's okay to have that mindset. I get more worried about the students who don't do the homework because they're like, 'Oh I can turn it in whenever', and that's where I still struggle with what should I do in this sense.

Teachers described various strategies that they used to reduce the likelihood of students abusing the system, including: averaging the original and retake score, allowing retakes on homework and classwork but not exams, and accepting the retake score regardless of whether it was higher or lower than the first score, among others. Teacher 2 recounted an incident of having to justify his policy of averaging the original and retake scores:

I just had a conversation with one of my students for about a week and I'm like, that first grade still counts and she's like, 'Well why? Why don't I have a hundred?' And I'm like, because that first grade still counts. You can't just knock it off.

Teacher 7 required that students lay out a detailed plan for how they would prepare for the retake (e.g., stay after school for tutoring, make flashcards, etc.), and only gave retakes to those students who had evidence that they met the criteria of their plan. During class (7-1), while students worked on independent practice, she conferenced with some students individually to draft and approve their retake plans. If students could not elaborate on their strategy, they were sent back to their desks to develop the plan in greater detail. Similarly, if students did not have evidence that they had implemented the plan, they did not

earn the retake. During her interview, she described that these steps were designed to make retakes labor-intensive, to discourage students from taking advantage:

I make it worth my while for me having to do the extra work behind it. So, if I'm gonna have to give retakes, then that means one, you're gonna prepare for a retake. If I gotta create a second assessment, grade a second assessment, then I make them prepare ahead of time. They just can't come and say, 'Oh I wanna take a retake.'

Other teachers developed different strategies to prevent students from abusing the retake system. Some teachers explained that they did not “broadcast” their retake policy, but offered it individually to students who they believed earned a retake by working hard on their first attempt. Teacher 6 stated, “And then sometimes you have the situation, you'll give them a test and they'll be like, 'Oh I don't know this, I'll just retake it.' And that is the most annoying thing I've ever seen in my life. So I just don't tell them. I don't tell them.”

For many teachers, successfully managing the retake process was critical, as retakes intensified the amount of grading and paperwork teachers needed to process. Some teachers reported feeling overwhelmed with the amount of work needed to implement the retake policy, as they graded assignments multiple times and had to keep track of which students completed retakes and what their grades should actually be. Indeed, Teacher 1 and Teacher 3 were interrupted during their interviews by students who were requesting work for retakes. Teacher 3 directed students toward a filing system where materials were organized by assignment, reminding them that it was their responsibility to procure and complete retake materials. Teacher 1 paused the interview to disseminate materials to students and explain his expectations for the assignments. Afterward, he described the workload created by the retake policy:

For me personally, I find that the organization and the [grading] cycle of that to be really overwhelming. And to find the time...who's missing what, what time a deadline would be, how would a student absence factor into that...then, on top of that, [there

is] the actual grading – grading things the first time – planning ahead for future lessons, parent conferences with students...it's like, how is that done in a good way? That's a challenge for me. Doing the retake thing in a systematic way, to support students.

Teacher 3 pointed out that with approximately one hundred students, if she assigned thirty assignments in a semester, she was grading a minimum of three thousand pieces of student work before retakes. Retakes intensified this workload, creating what she called a “tremendous” amount of paperwork to process. She described the importance of only allowing retakes on certain assignments, to reduce the grading burden:

Right now, I just don't have...the system would be crazy if I were to try and like remember who was late for an assignment and do everything like that. And I know we can put that into our gradebook but I think it would make me crazy. I'm gonna look at a better system for that next year.

These excerpts highlight that the school-wide retake policy had several downsides. Its implementation had an unintended effect of decreased student motivation, where students did not put forth their best effort, banking on the opportunity to retake assignments to improve their grades. In addition, teachers felt it was important to dedicate time to reducing the likelihood that students would abuse the policy, not only to encourage students' motivation but because the retake policy created a considerable amount of extra work for teachers.

Discussion

This study explored how growth mindset and stereotype threat strategies are applied by middle school teachers in a public charter school in a rural, low-income community. The setting for this study provides a unique environment to explore how teachers translate these strategies for classroom use because, though it is a high need school, its context includes many factors that promote teachers' implementation (Domitrovich et al., 2008; Hall, 2013). The school's mission is strongly aligned to promoting positive intelligence beliefs (Datnow,

2002; Kallestad & Olweus, 2003). Teachers not only receive resources and training on growth mindset and stereotype threat concepts, but they are actively encouraged by administrators to apply them to their classrooms - though they are given flexibility for how they implemented mission-aligned policies, such as retakes (Ringwalt et al., 2003). In addition, the school had recently implemented the school-wide retake policy to support growth mindset development. At the time teachers were interviewed, they were responding to outcomes of strategy use in their classrooms.

Supporting Positive Intelligence Beliefs

Teachers emphasized the importance of effort through *productive struggle*, where students were not only expected to work hard but were encouraged to make and learn from mistakes. In class, teachers encouraged students to review their own mistakes, held whole-group discussions about mistakes in student work and provided ample, structured time where students could fix their errors as a learning opportunity (e.g., Teacher 2 using pre-tests before exams). Teachers' positive messages about mistakes are in line with seminal research that links similar statements to increased growth mindset beliefs in students (Dweck, Davidson, Nelson, & Enna, 1978; Dweck & Bush, 1976; Dweck & Leggett, 1988; Mueller & Dweck, 1998). Our findings also align with more recent research from classroom observations that found encouraging and normalizing mistakes to be a common growth mindset strategy used by teachers (Sun, 2018). To further support productive struggle and growth mindset development in students, the school in our study had adopted a retake policy for assignments and assessments, which was introduced to reinforce the growth mindset messaging that mistakes are part of the learning process and that students can improve their grades with continued effort. Retakes were also a growth mindset strategy found by Sun (2018) in her

classroom observations. Though retakes are not part of the large-scale growth mindset interventions conducted in schools (Blackwell et al., 2007; Good et al., 2003; Pauneksu et al. 2015, Yeager et al., 2014; Yeager et al., 2016), teachers in the present study believed that their retake policy was growth mindset-aligned. Similarly, as argued by Sun (2018), retakes make actionable the growth mindset messaging that intelligence develops over time, with continued effort applied to a learning task.

In addition, teacher's emphasis on productive struggle aligns with more recently identified complexities of growth mindset theory. Reflecting on several decades of research in this area, Dweck (2015) explains a common pitfall in the implementation of her research, where educators praise and encourage effort without providing much truth about whether students are actually learning. She clarifies that sheer effort, in and of itself, may not be sufficient to boost achievement. Students may need to ask for help or try a variety of new approaches rather than simply increase their persistence with strategies that have been unsuccessful in the past (Dweck, 2015). Teacher behaviors in this study reflected such nuance: the emphasis on productive struggle was honest about failure, encouraged students to feel comfortable with and to learn from their mistakes and messaged learning as a process. Teachers also frequently used the phrase *with quality* to convey high expectations about how students should complete their work. Some teachers told students that they would not accept retakes unless they were done with quality, suggesting that it was not solely students' continued effort but their thoroughness and attention to detail that would result in improved grades.

Dweck's (2015) sentiments were also reflected in the language teachers used across classrooms about the intentionality of effort and the importance of purposeful mistake-

making. For example, Teacher 2's statement about "woefully making mistakes" describes a student who is not thinking with intention about what mistakes have been made or what can be learned from them. The emphasis on productive struggle was modeled by Teachers 6 and 4, who used student work that contained mistakes to facilitate whole group conversations about using feedback to make improvement. Productive struggle was also reflected in strategies such as those used by Teacher 7, who required students to develop and receive feedback on a detailed plan that outlined *how* their effort would translate to improvement in their grades. In their emphasis on productive struggle, teachers in this study contained elements of "wise" feedback that, in other research, has been linked with students' increased motivation ("I'm giving you these comments because I have very high expectations and I know that you can reach them"; Yeager et al., 2014). Teachers' acceptance of mistakes and their willingness to provide retakes conveys a similar message that students' grades are feedback to indicate their current level of learning and that students can become smarter by processing their mistakes and showing continued improvement on their assignments.

Implications for Practice

The current study provides several implications for the implementation of growth mindset and stereotype threat strategies with middle school students. Though quick and low-lift intervention strategies (e.g., one 45-minute computer activity, 15-minute written reflection, wise feedback) have shown effects on student achievement (e.g., Cohen et al., 2009; Paunesku et al., 2015; Yeager et al., 2014), it may require much more effort for teachers to integrate these theories into their daily classroom practice. As Hall (2013) describes, school change is an ongoing and emotional process, where teachers grow in their

confidence and competence, and their strategy use “morphs” over time in response to their given context.

Teachers in this sample echoed elements of the CBAM framework (Hall 2013). In addressing teachers’ self-concerns (i.e., Can I do it?), this school provided teachers with several weeks of training dedicated to implementing growth mindset and stereotype threat strategies. Beyond training, however, teachers were given flexibility to adapt strategies to their individual classrooms – and many teachers chose to make adjustments based on ways that would alleviate both task (i.e., do I have time?) and impact (i.e., will it work?) concerns. For example, several teachers in this study established policies that made assignment and test retakes very labor intensive, reducing the likelihood that students abused the system and, in turn, reducing the additional grading burden on teachers. Some teachers shared impact-concerns that students may abuse growth mindset strategies, taking advantage of second chances or the impermanence of failure as an opportunity to exhibit effort selectively.

In addition, several teachers in this study indicated that the retake policy did not align with their personal beliefs and/or values and did not feel the policy would ultimately help their students, giving students a “crutch” to postpone their best effort until the last minute. These teachers stated that they only offered retakes because it was a school-wide policy. Although mandatory policies can be one way to force compliance among teachers, they undermine the autonomy that can create more intrinsically motivated implementation (Deci & Ryan, 1985; Domitrovich et al., 2008). Frey and colleagues liken autonomy-supportive implementation – which involves teachers in the development and planning around intervention strategies – to the use of motivational interviewing (MI; Miller and Moyers, 2006), a process of probing and questioning that can increase the likelihood participants’ will

exhibit new behaviors. When an MI-type framework is applied to the school context, administrators solicit teacher feedback as a primary step in enhancing teachers' confidence and in cultivating their commitment to the intervention (Frey et al., 2013). Including teachers in an ongoing conversation around the impact of intervention strategies can also illuminate ways that the strategies are successful and can address teachers' concerns that the intervention is ineffective or backfiring, promoting consistency and fidelity of implementation (Collie, Shapka, & Perry, 2013; Dorman, 2003; Frey et al., 2013; Hall 2013; Lasky, 2005; Woods, 1999). Soliciting teacher input can also help identify ideas to improve existing strategies, increasing the likelihood of successful implementation (Durlak, 2015). Though assessing collaborative interactions between administrators and teachers were beyond the scope of the present study, teachers in our sample did exhibit frustration with implementing a mandatory policy. It is possible that, at the very least, being given autonomy to customize a mandatory policy for their classroom use may have helped maintain implementation among teachers who did not wholly agree with strategies selected by school leaders (Ringwalt et al., 2003). To help teachers feel more fully invested in implementation, administrators could consider working more collaboratively with teachers, using an autonomy-supportive framework (e.g., Frey et al., 2013).

Implications for Research

This study indicates that when exposed to research theory and evidence, teachers and school leaders developed their own strategies for implementation. This study also provides several trailheads for future research in the areas of supporting positive intelligence beliefs. Future work could evaluate the effects of assignment retakes on students' intelligence beliefs, effort and achievement. Additional work could investigate the efficacy of various grading

strategies used to promote rather than reduce student engagement via retakes, such as averaging scores, using only the retake score, etc. Evaluation methods such as the rapid A/B testing used by Paunesku and colleagues (2015) can provide guidance for quickly testing program components against each other to identify critical intervention elements.

Strengths & Limitations

This study is not without limitations. These findings do not generalize to other schools and communities, where contexts may be vastly different. It is difficult to extract the influence of school culture on both teacher and student behavior, as the school context itself can have important implications for supporting student achievement and teachers' willingness to try new strategies (Durlak, 2015). In particular, the mission of this charter school was unusual in that it was explicitly and strongly aligned to fostering positive intelligence beliefs among students, which can make it more likely that teachers will engage in any one particular strategy or will buy in to strategy use in general (Domitrovich et al., 2008). It is possible that, absent a school context that supports the use of growth mindset and stereotype threat strategies, teachers in other schools may not develop their own strategies to integrate these theories in their classrooms, may develop different strategies or may be less likely to implement strategies that are provided to them (Domitroich et al., 2008). In a school with strong consistency of mission-aligned behaviors across classrooms, such as the school used for this study, teachers may also not be entirely aware of individual strategies that they are using in their classrooms or may not be aware that particular strategies are intended to influence students' intelligence beliefs, making it harder to discuss the interview questions for this study. The nature of the charter school's mission may also draw a self-selected population of teachers who are passionate about addressing issues like ability and identity,

promoting their comfort and investment with growth mindset and stereotype threat strategies. It must also be acknowledged that teachers who declined to participate in this study may differ in their opinions and use of growth mindset and stereotype threat strategies and may also have different perceptions of their students, school leadership and community context. Indeed, participating teachers alluded to the fact that not all of their colleagues support the school's procedures and policies or the school's mission. Only two teachers participated in member checking, and it is possible that teachers who declined to look at study findings may have disagreed with or had additional suggestions for modifications to these themes.

What's more, the primary researcher must acknowledge her positionality within the community and within the school (Foote & Bartell, 2011). The primary researcher is white, was not educated in a rural or low income community and was a secondary education teacher in the public school system in the same county as the charter school. These experiences may create bias in the researcher's worldview or prior beliefs about the influence of identity, the interactions among teachers and students, the role of educators or the nature of this schools' community context. In this study, the researchers' positionality may have influenced the lens through which classroom observations were conducted, the types of questions that were chosen for teacher interviews, the facilitation of interview discussions, the analysis of observational and interview data and the presentation of findings. Study findings may also have been affected by teachers' perceptions of the primary researcher, particularly in terms of prior experience as an educator in the county: Through continued work in community, the primary researcher is familiar with school administrators. Though great effort was taken to maintain privacy and confidentiality so that administrators were blind to which teachers chose to participate, teacher perceptions of relationships among the researcher and school

administrators may have promoted social desirability in their responses or reduced their likelihood of disclosing certain information. It is also possible that teachers' perceptions of relationships among the researcher and school leaders influenced their decisions about participation in general, creating bias in the sample for this study.

In addition, the sample for this study is small, and one teacher was not able to be interviewed following observation. Future studies could improve upon our work by targeting larger schools and/or schools in different communities with students of different backgrounds. However, the sample includes a majority of the core content teachers in the school and spans every subject in every grade level. It includes both veteran and beginning teachers of different races/ethnicities. Some teachers grew up and were educated in this community, and some teachers were transplants from across the country. To promote trustworthiness and rigor, data were triangulated between teacher interview and multiple classroom observations and findings were reviewed by participants, who provided feedback via member checking.

Conclusion

Despite its limitations, this study provides insight into how teachers operate in their classrooms when they are exposed to research on growth mindset and stereotype threat. To guide the dissemination of these theories into the field of education and among educators, we explored teachers' experiences and opinions of applying these theories, as well as outlined potential strategies that could be investigated in future research. In sum, this work informs ongoing research on the implementation of growth mindset and stereotype threat theories in schools, with teachers as implementers, in service of students in greatest need of intervention.

REFERENCES: PAPER III

- Allen, M. (2017). *The sage encyclopedia of communication research methods* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc. doi:10.4135/9781483381411
- Aronson, J., Blanton, H., & Cooper, J. (1995). From dissonance to disidentification: Selectivity in the self-affirmation process. *Journal of Personality and Social Psychology, 68*(6), 986. doi:10.1037/0022-3514.68.6.986
- Bethge, J. (2018). *The Power of Transformation: A Grounded Theory Study of Cultivating Teacher Growth Mindset Towards Student Intelligence*.
- Betoret, F. D. (2009). Self-efficacy, school resources, job stressors and burnout among Spanish primary and secondary school teachers: a structural equation approach. *Educational Psychology, 29*(1), 45-68. doi:10.1080/01443410802459234
- Bird, C.M. (2005). How I stopped dreading and learned to love transcription. *Qualitative Inquiry, 11*, 226-248. doi:10.1177/1077800404273413
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Borman, G. D., Grigg, J., & Hanselman, P. (2016). An effort to close achievement gaps at scale through self-affirmation. *Educational Evaluation and Policy Analysis, 38*(1), 21-42. doi:10.3102/0162373715581709
- Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2013). Enhancing a brief writing intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology, 105*(2), 427-435. doi:10.1037/a0031177
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. doi:10.1191/1478088706qp063oa
- Brown, E. (2013). No child left behind and the teaching of character education. *ABNF Journal, 24*(3), 77-82.
- Bryk, A., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.
- Claro, S., Paunesku, D., & Dweck, C. S. (2015). Mindset equals income as a predictor of achievement. *Unpublished manuscript, Stanford University, Stanford, CA*.

- Cohen, G. L., & Garcia, J. (2008). Identity, belonging, and achievement: A model, interventions, implications. *Current Directions in Psychological Science*, 17, 365–369. doi:10.1111/j.1467-8721.2008.00607.x
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, 324(5925), 400-403. doi:10.1126/science.1170769
- Cohen, L., Manion, L., & Morrison, K. (2011). Surveys, longitudinal, cross-sectional and trend studies. *Research Methods in Education, 7th edition*. Abingdon: Routledge, 261-4.
- Cohen, G. L., & Steele, C. M. (2002). A barrier of mistrust: How negative stereotypes affect cross-race mentoring. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 303–327). San Diego, CA: Academic Press. doi:10.1016/B978-012064455-1/50018-X
- Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, 104(4), 1189. doi:10.1037/a0029356
- Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor’s dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin*, 25(10), 1302–1318. doi:10.1177/0146167299258011
- Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, 104(4), 1189.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative and mixed methods approaches (3rd ed.)*. London, UK: Sage.
- Daly, J., Kellehear, A., Gliksman, M., & Daly, K. G. (1997). *The public health researcher: A methodological guide*. Melbourne: Oxford University Press.
- Datnow, A. (2002). Can we transplant educational reform, and does it last? *Journal of Educational Change*, 3(3-4), 215-239. doi:10.1023/A:1021221627854
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Dee, T. S., & Jacob, B. (2011). The impact of no child left behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446. doi:10.1002/pam.20586

- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., ... & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion, 1*(3), 6-28. doi:10.1080/1754730X.2008.9715730
- Dorman, J. P. (2003). Relationship between school and classroom environment and teacher burnout: A LISREL analysis. *Social Psychology of Education, 6*(2), 107-127. doi:10.1023/A:1023296126723
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087-1101. doi:10.1037/0022-3514.92.6.1087
- Duckworth A., Quinn, P. and Tsukayama, E. (2011a). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized test scores and report card grades. *Journal of Educational Psychology, 104*(2), 439-451. doi:10.1037/a0026280
- Duckworth, A. and Seligman, M. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939-944. doi:10.1111/j.1467-9280.2005.01641.x
- Durlak, J. A. (2015). *Handbook of social and emotional learning: Research and practice*. Guilford Publications.
- Dweck, C. (2015). Carol Dweck revisits the growth mindset. *Education Week, 35*(5), 20-24.
- Dweck, C. S., & Bush, E. S. (1976). Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. *Developmental Psychology, 12*(2), 147-156. doi:10.1037/0012-1649.12.2.147
- Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Developmental Psychology, 14*(3), 268. doi:10.1037/0012-1649.14.3.268
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273. doi:10.1037/0033-295X.95.2.256
- Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2005). Implementation research: A synthesis of the literature.
- Edwards, A. R., Esmonde, I., Wagner, J. F., & Beattie, R. L. (2017). Learning mathematics. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 57–80). New York, NY: Routledge.

- Foote, M. & Bartell, G. (2011). Pathways to equity in mathematics education: How life experiences impact researcher positionality. *Educational Studies in Mathematics*, 78, 45-68.
- Frey, A. J., Lee, J., Small, J. W., Seeley, J. R., Walker, H. M., & Feil, E. G. (2013). The Motivational Interviewing Navigation Guide: a process for enhancing teachers' motivation to adopt and implement school-based interventions. *Advances in School Mental Health Promotion*, 6(3), 158-173. doi:10.1080/1754730X.2013.804334
- Friese, S. (2018). *ATLAS.ti 8 Windows User Manual*. ATLAS.ti Scientific Software Development GmbH, Berlin.
- Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs forcing*. Sociology Press.
- Glennie, E., Bonneau, K., vanDellen, M., & Dodge, K. A. (2012). Addition by subtraction: The relation between dropout rates and school-level academic achievement. *Teachers College Record*, 114(8), 1-26.
- 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(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Gottfredson, G. D., Jones, E. M., & Gore, T. W. (2002). Implementation and evaluation of a cognitive-behavioral intervention to prevent problem behavior in a disorganized school. *Prevention Science*, 3(1), 43-56. doi:10.1023/A:1014671310038
- Hall, G. E. (1978). *Concerns-Based In-service Teacher Training: An Overview of Concepts, Research and Practice*
- Hall, G. E. (2013). Evaluating change processes: Assessing extent of implementation (constructs, methods and implications). *Journal of Educational Administration*, 51(3), 264-289. doi:10.1108/09578231311311474
- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education*, 87(2), 106-124. doi:10.1177/0038040714525970
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, 91(2), 145-149. doi:10.1257/aer.91.2.145
- Helfrich, C. D., Weiner, B. J., McKinney, M. M., & Minasian, L. (2007). Determinants of implementation effectiveness: adapting a framework for complex innovations.

- Medical Care Research and Review*, 64(3), 279-303.
doi:10.1177/1077558707299887
- Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, 79(1), 491-525. doi:10.3102/0034654308325693
- Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In *Advances in experimental social psychology* (Vol. 28, pp. 281-388). Academic Press.
- Kallestad, J. H., & Olweus, D. (2003). Predicting teachers' and schools' implementation of the Olweus bullying prevention program: A multilevel study. *Prevention & Treatment*, 6(1), 21a. doi:10.1037/1522-3736.6.1.621a
- Kelleher, C., Riley-Tillman, T. C., & Power, T. J. (2008). An initial comparison of collaborative and expert-driven consultation on treatment integrity. *Journal of Educational and Psychological Consultation*, 18(4), 294-324.
doi:10.1080/10474410802491040
- Lapadat, J.C. and Lindsay, A.C. 1999: Transcription in research and practice: From standardization of technique to interpretive positionings. *Qualitative Inquiry*, 5, 64-86. doi:10.1177/107780049900500104
- Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education*, 21(8), 899-916. doi:10.1016/j.tate.2005.06.003
- Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience*, 7, 230–241. doi:10.1093/scan/nsq100
- McGuinn, P. (2012). Stimulating reform: Race to the top, competitive grants and the Obama education agenda. *Educational Policy*, 26(1), 136-159.
doi:10.1177/0895904811425911
- Miles, M. B., Huberman, A. M., Huberman, M. A., & Huberman, M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Sage.
- Miller, W.R., & Moyers, T.B. (2006). Eight stages in learning motivational interviewing. *Journal of Teaching in the Addictions*, 5, 3 –18. doi: 10.1300/J188v05n01_02
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33-52. doi:10.1037/0022-3514.75.1.33

- Noell, G. H., Duhon, G. J., Gatti, S. L., & Connell, J. E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review, 31*(2), 217.
- Oyserman, D., Bybee, D., & Terry, K. (2006). Possible selves and academic outcomes: How and when possible selves impel action. *Journal of Personality and Social Psychology, 91*(1), 188. doi:10.1037/0022-3514.91.1.188
- Padgett, D. K. (2016). *Qualitative methods in social work research (Vol. 36)*. Sage Publications.
- Pandit, N. R. (1996). The creation of theory: A recent application of the grounded theory method. *The Qualitative Report, 2*(4), 1-15.
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017
- Punch, M. (2009). *Police corruption: Deviance, accountability and reform in policing*. Routledge.
- Protzko, J., & Aronson, J. (2016). Context moderates affirmation effects on the ethnic achievement gap. *Social Psychological and Personality Science, 7*(6), 500-507. doi:10.1177/1948550616646426
- Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review, 38*(4), 510.
- Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok — not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*(3), 731-737. doi:10.1016/j.jesp.2011.12.012
- Riessman, C. K. (1993). *Narrative analysis (Vol. 30)*. Sage.
- Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., Vincus, A., & Thorne, J. (2003). Factors Associated with Fidelity to Substance Use Prevention Curriculum Guides in the Nation's Middle Schools. *Health Education & Behavior, 30*(3), 375-391. doi:10.1177/1090198103030003010
- Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York, NY: McGraw-Hill.

- Rothstein, R., & Jacobsen, R. (2006). The goals of education. *The Phi Delta Kappan*, 88(4), 264-272.
- Santavirta, N., & Solovieva, S. (2007). The association between job strain and emotional exhaustion in a cohort of 1,028 Finnish teachers. *British Journal of Educational Psychology*, 77(1), 213-228. doi:10.1348/000709905X92045
- Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S., . . . Cohen, G. L. (2013). Deflecting the trajectory and changing the narrative: How self-affirmation affects academic performance and motivation under identity threat. *Journal of Personality and Social Psychology*, 104(4), 591-618. doi:10.1037/a0031495
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811. doi:10.1037/0022-3514.69.5.797
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of Qualitative Research*, 17, 273-85.
- Strizek, G. A., Pittsonberger, J. L., Riordan, K. E., Lyter, D. M., & Orlofsky, G. F. (2006). Characteristics of Schools, Districts, Teachers, Principals, and School Libraries in the United States: 2003-04 Schools and Staffing Survey. ED TAB. NCES 2006-313. National Center for Education Statistics.
- Sun, K. (2018). The role of mathematics teaching in fostering student growth mindset. *Journal for Research in Mathematics Education*, 49(3), 330-355. doi:10.5951/jresmetheduc.49.3.0330
- Tolan, P., Gorman-Smith, D., & Henry, D. (2004). Supporting families in a high-risk setting: proximal effects of the SAFE Children preventive intervention. *Journal of Consulting and Clinical Psychology*, 72(5), 855. doi:10.1037/0022-006X.72.5.855
- Vanneman, A., Hamilton, L., Anderson, J. B., & Rahman, T. (2011) Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Von Brock, M. B., & Elliott, S. N. (1987). Influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology*, 25(2), 131-144. doi:10.1016/0022-4405(87)90022-7
- Wickstrom, K. F., Jones, K. M., LaFleur, L. H., & Witt, J. C. (1998). An analysis of treatment integrity in school-based behavioral consultation. *School Psychology Quarterly*, 13(2), 141. doi:10.1037/h0088978

- Woods, P. (1999). Intensification and stress in teaching. Understanding and preventing teacher burnout: A sourcebook of international research and practice, 115-138.
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143(2), 804. doi:10.1037/a0033906
- 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(3), 374. doi:10.1037/edu0000098

Table 3.1. Final summary of themes.

Implementing Growth Mindset Theory		
	Theme	Description
Strategies	Productive Struggle	<p>Teachers encourage effortful practice</p> <ul style="list-style-type: none"> • Teachers message doing work with quality and earning success, privileges • Teachers tell students it's okay to make mistakes • Teachers encourage a culture where mistake making is welcome and celebrated • Teachers message purposeful mistake making • Teachers provide retakes so students can make improvements on past assignments
Unintended Consequences	Managing Retakes	<ul style="list-style-type: none"> • Some students abuse the retake policy by not putting forth effort the first time around • Teachers have to develop policies to reduce the likelihood of students taking advantage of retakes • The retake policy created a lot of additional work for teachers, especially grading and paperwork • Teachers must keep track of students' make up work and frequently changing grades

Appendix A

Effective Classroom Strategies Checklist – Sense of Competence Subscale
from N.K. Bowen, 2011

SENSE OF COMPETENCE

Amount of evidence	Extensive	Moderate	Minimal	None	NA during this observation
Evidence of the Growth Mindset					
Conveys assumption of success					
Rewards effort, perseverance					
Explicitly refers to or teaches growth mindset ideas					
Invites questions					
Praises improvement					
Discourages fixed mindset comments or behaviors					
Asks students to help each other					
Helps students problem-solve					
Models growth mindset					
Student comments or behaviors demonstrate growth mindset					
Recognizes multiple domains of ability					
Other evidence of growth mindset					
Evidence of Countering Stereotype Threat					
Explicitly discusses stereotypes & their negative consequences					
Makes explicit efforts to counter common stereotypes					
Statements refute or challenge stereotypes					
There are signs of respect for difference in classroom					
Models experiences with or learning about stereotypes					
Interactions with students from stereotyped groups actively counter stereotypes					
Other evidence					

Appendix B

Interview Guide for Teachers

1. How do you motivate your students to work hard and do well in your class?
2. How do you handle it when students make a mistake?
3. How do you handle it when students are failing?
4. What differentiates students who are learning from those who are not?
 - a. Do you use different strategies based on the student you're working with?
5. How do you deal with issues around fairness or favoritism?
6. How do you address stereotypes around intelligence?

If necessary:

7. What do you think the term “growth mindset” means?
 - a. How do you feel about incorporating growth mindset strategies into your daily teaching practice?
8. What do you think the term “stereotype threat” means?
 - a. How do you feel about incorporating stereotype reducing strategies into your daily teaching practice?

SUMMARY

Despite a growing body of research that has investigated the effects of implementing growth mindset and stereotype threat interventions in schools, little research exists to support school administrators and classroom teachers with the integration of these intervention strategies into their daily practice (Edwards, Esmonde, Wagner, & Beattie, 2017). These dissertation papers not only address this gap in the literature but, taken together, they also reiterate the importance of student-teacher interactions in shaping how students think about themselves and their academic abilities.

Paper 1: Results and Implications

Paper 1 is the first systematic review and meta-analysis of growth mindset and stereotype threat interventions, a particularly informative contribution to the literature in these areas. The results of Paper 1 indicate that both growth mindset and stereotype threat interventions have demonstrated impact in improving student achievement - albeit, on average, small effects. In addition, significant findings were inconsistent across studies in both areas, suggesting that school characteristics play a role in the success of either intervention. In terms of stereotype threat, it appears that the racial composition of the student body has strong implications for how stereotype threat manifests in schools, and threat-reducing interventions may only boost student performance insofar as performance is currently suppressed by threat. Although the factors that contribute to a high threat school context have been studied very little in reference to identity affirming intervention strategies (e.g., Hanselman, Bruch, Gamoran, & Borman, 2014), it appears that diverse student bodies

that contain both white and non-white students may have the highest potential for threat, while a student body that is composed entirely of students of color (e.g., Simmons, 2013) may elicit less threat, based on race. More research should be dedicated to understanding how threat manifests in schools so that practitioners are able to identify whether the characteristics of their schools are suitable for an intervention strategy like the values affirmation reflection. As the values affirmation reflection is tested in more schools, it may also be possible to conduct moderator analyses in a future meta-analysis that could more definitively assess the impact of school context on intervention effects. To facilitate this type of analysis, studies that pool data across schools could also analyze effects separately by school to avoid the potential for suppression effects by school context (see Hanselman et al., 2014 for an example).

Comparisons across growth mindset studies were complicated by the lack of a consistent intervention strategy. Studies varied in terms of the format, length and intensity of activities, making it difficult to isolate effective intervention components. In the same way, it was not possible to compare the effects of similar strategies across various student populations. Future research could use the results of Paper 1 to identify strategies that have shown significant impact on student achievement for use in more targeted replications that compare intervention elements and differential effects across students based on race/ethnicity, gender, socioeconomic status and prior academic performance (see Paunesku, Walton, Romero, Smith, Yeager, & Dweck, 2015, for an example). In sum, Paper 1 supports the importance of continued research on growth mindset and stereotype threat interventions as promising strategies to close achievement gaps. Paper 1 also identifies trailheads for

exploring aspects of each intervention that may lead to more clarity around which school and student populations may benefit most from growth mindset and stereotype threat strategies.

Paper 2: Results and Implications

Paper 2 elevates the perspectives of students in a low income, rural school – students whose voices infrequently reach academic literature. Findings from Paper 2 suggest that middle school students are likely to endorse growth mindset beliefs themselves, but they also attend to their teachers' behavior and take cues about what their teachers believe about the nature of intelligence. Students also monitored their teachers' behavior for clues as to which students they felt their teachers believed were the smartest, focusing on concrete and observable behaviors such as whom teachers chose to help or call on in class. Of note, students also made a link between teachers' emotional support and beliefs about students' intelligence, where students believed that teachers showed care to students they believed could achieve. Although students did not feel that teachers stereotyped students based on their gender or race/ethnicity, students in our sample were aware that some teachers may be judging students based on socioeconomic status, showing bias against low income students. Our findings highlight the ability of middle school students to observe and to understand abstract concepts about intelligence, identity, ability and stereotypes and the role that teachers play in shaping students' beliefs about these concepts.

Findings from Paper 2 contribute to the literature in several ways. First, Paper 2 establishes a basis for future work that explores students' perceptions of their teachers, providing a set of behaviors that that could be considered in the creation of measures of teachers' mindset-supportive or threat-reducing behavior, which could be assessed through further classroom observation. Second, the salience of stereotypes based on students'

socioeconomic status supports other research that has identified that low income students – of various races/ethnicities, including white students – may also be subject to suppressed achievement as a result of stereotype threat (Bowen, Wegmann, & Webber, 2013). The potential for low income white students to benefit from threat reducing intervention remains unexplored and should be a focus of future research.

Importantly, the use of focus groups, while helpful for including as many students as possible in our sample, was not ideal for discussing sensitive topics such as gender, race/ethnicity and socioeconomic status. Future qualitative research would be improved through the use of individual student interviews or by structuring focus groups more strategically, organizing groups homogeneously by identity age, etc. In addition, because we grouped students across multiple grade levels, we were not able to draw conclusions about the progression of students' beliefs over time, which would be useful knowledge for teachers and other school practitioners. Middle school is a time of particularly accelerated change in students' abstract thinking, perspective-taking abilities and beliefs about intelligence (Eccles, Lord, & Midgley, 1991; Meece, Anderman, & Anderman, 2006; Midgley, Anderman, & Hicks, 1995; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991), and analyzing patterns in students' responses by age would be a fruitful avenue for future research. Combined with findings from Paper 1, Paper 2 suggests that middle school students would attend to other types of classroom strategies to support growth mindset and reduce stereotype threat, in addition to those that have been evaluated in RCTs, providing support for continued research to further develop both interventions.

Paper 3: Results and Implications

Finally, Paper 3 makes a novel contribution to the literature by exploring how teachers integrate growth mindset and stereotype threat theories into their classrooms, including factors affecting teachers' implementation of strategies. Although limited research has investigated teachers' organic implementation of these theories, findings from Paper 3 corroborate prior work (Sun, 2018) that has shown resubmission of assignments (referred to as "retakes" by teachers in our sample) as a growth mindset-supportive strategy that has not been incorporated into large-scale growth mindset interventions (e.g., Blackwell, Trzesniewski, & Dweck, 2007; Good, Aronson, & Inzlicht, 2003; Paunesku et al., 2015; Yeager et al., 2014; Yeager et al., 2016). Retakes are a strategy that could be rigorously evaluated through future randomized controlled studies to better assess their influence on student behavior and achievement.

Findings from Paper 3 also contribute to the knowledge base around the integration of mindset and stereotype threat theories into schools by identifying several unintended consequences of using these strategies. Teachers felt that implementing a growth mindset strategy like retakes, for example, was not perfect at fostering increased effort and growth mindset beliefs among students and in fact sometimes produced the opposite effect of decreased student engagement. Paired with the increased workload created by retakes, lack of student effort made some teachers wary of implementing the strategy. Yet, importantly, students in Paper 2 also identified retakes and resubmission of work (as part of teachers' instrumental support) as indicators of teachers' beliefs about their ability, highlighting that understanding the full impact of any one strategy from both the perspectives of both teachers and students is important for developing a more complete understanding of its effects. Learning more about students' perspectives is helpful for generating new potential strategies,

but without teachers' perspectives – for example, on the drawbacks of retakes – it may be easy to assume they are universally positive strategies. As researchers and school leaders are encouraged by studies such as those included in Paper 1 that show the promise of growth mindset intervention for promoting student achievement, findings from Papers 2 and 3 provide context for what implementation may look like when integrated more fully into school practice.

In other school contexts, teachers may experience similar unintended impacts when implementing growth mindset and stereotype threat theories or may experience wholly different impacts, depending on their school context and study body. It is also likely that teachers in other schools have developed different strategies to support growth mindset development and to reduce stereotype threat. Much more research will need to be conducted in a variety of schools to explore how teachers develop strategies for daily use in their classrooms and how these strategies influence students and their families. As evidenced by the teachers in the sample for Paper 3, regardless of strategy, school leaders should consider giving teachers flexibility in their implementation and in how they respond to the impacts of their strategy use; ideally, teachers should be leveraged in the generation of strategies, and their feedback should be solicited in evaluating the success of implementation (Hall, 2013).

Final Conclusions

Overall, this dissertation builds upon the evidence base supporting two increasingly popular interventions for increasing student achievement and addressing achievement gaps (Aronson, Cohen, McColskey, Montrosse, Lewis, & Mooney, 2009). Findings from Paper 1 consolidate RCT evaluations of growth mindset and stereotype threat interventions, showing small, positive effects. Papers 2 and 3 demonstrate that teachers may use different strategies

to promote positive intelligence beliefs among students than those strategies explored in the studies from Paper 1. Papers 2 and 3 also identify additional strategies that could be included in future RCTs testing growth mindset and stereotype threat strategies, which could be summarized through future meta-analyses. All three papers highlight implications for teachers and school leaders, as well as researchers who study students' beliefs about intelligence and the impact on their academic success. This dissertation shows that supporting positive intelligence beliefs among students is a promising direction for closing achievement gaps, but that for these strategies to be maximally effective, more consideration should be made for both students' and teachers' experiences with these concepts in the classroom, particularly the drawbacks of strategy use. In sum, this dissertation advances our knowledge of how students' intelligence beliefs can be leveraged to improve their academic performance and how teachers can best support students in developing positive intelligence beliefs. It informs future research that can help improve academic outcomes for all students and, especially, for students who have historically been underserved by the schools they attend.

REFERENCES: SUMMARY

- Aronson, J., Cohen, G., McColskey, W., Montrosse, B., Lewis, K., & Mooney, K. (2009). Reducing stereotype threat in classrooms: A review of social-psychological intervention studies on improving the achievement of Black students (Issues & Answers Report, REL 2009–No. 076). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2013). Enhancing a brief writing intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology, 105*(2), 427-435. doi:10.1037/a0031177
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to early adolescents? The impact of educational contexts on early adolescents. *American Journal of Education, 99*(4), 521-542. doi:10.1086/443996
- Edwards, A. R., Esmonde, I., Wagner, J. F., & Beattie, R. L. (2017). Learning mathematics. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 57–80). New York, NY: Routledge.
- Hall, G. E. (2013). Evaluating change processes: Assessing extent of implementation (constructs, methods and implications). *Journal of Educational Administration, 51*(3), 264-289. doi:10.1108/09578231311311474
- Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education, 87*(2), 106-124. doi:10.1177/0038040714525970
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365 – 371. doi:10.1111/1467-9280.00272
- Inzlicht, M., & Ben-Zeev, T. (2003). Do High-Achieving Female Students Underperform in Private? The Implications of Threatening Environments on Intellectual Processing. *Journal of Educational Psychology, 95*(4), 796. doi:10.1037/0022-0663.95.4.796

- 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*(6), 645-662. doi:10.1016/j.appdev.2003.09.002
- Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487-503. doi:10.1146/annurev.psych.56.091103.070258
- Midgley, C., Anderman E.M., & Hicks, L. (1995). Differences between elementary and middle school teachers and students: a goal theory approach. *Journal of Early Adolescence, 15*(1), 90–113. doi:10.1177/0272431695015001006
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*(6), 784-793. doi:10.1177/0956797615571017
- Simmons, C. M. (2013). *Reducing stereotype threat in academically at-risk African American students: A self-affirmation intervention.* (2013-99010-209).
- Sun, K. (2018). The role of mathematics teaching in fostering student growth mindset. *Journal for Research in Mathematics Education, 49*(3), 330-355. doi:10.5951/jresmetheduc.49.3.0330
- Wigfield A., Eccles J.S., Mac Iver D., Reuman, D., & Midgley C. (1991). Transitions at adolescence: changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology, 27*, 552–565. doi:10.1037/0012-1649.27.4.552
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., ... & Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General, 143*(2), 804. doi:10.1037/a0033906
- 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*(3), 374. doi:10.1037/edu0000098