# SCALED-UP SOCIAL PSYCHOLOGY: INTERVENING WISELY AND BROADLY IN EDUCATION

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF PSYCHOLOGY

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

David Paunesku

August 2013

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Greg Walton) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Carol Dweck)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Geoff Cohen)

Approved for the University Committee on Graduate Studies

# Abstract

Over the last several decades, research has examined how students' beliefs about school and about their own abilities affect their academic goals, motivation, and achievement (for reviews see, Dweck, Walton, & Cohen, 2011; Farrington et al., 2012). It has also investigated how these beliefs and associated patterns of behavior can be influenced through interaction with others (Gunderson et al., 2013; Linnenbrink, 2005; Mueller & Dweck, 1998; Turner et al., 2002) and through precise, psychological interventions (J. Aronson, Fried, & Good, 2002; Blackwell, Trzesniewski, & Dweck, 2007; Cohen, Garcia, Apfel, & Master, 2006; Good, Aronson, & Inzlicht, 2003; Walton & Cohen, 2007, 2011; Wilson & Linville, 1982, 1985; Yeager & Walton, 2011). This research has focused on the educational context, but it has provided rich, generalizable insights. It has revealed a complex cross-section of the ways in which individuals' worries and motivations interact with their social environments to affect their behavior and major life outcomes.

The education context is in many ways ideal for the study of psychology. There are unambiguous, regularly-collected, socially meaningful outcomes; there are complex,

but consistently structured relationships; and there is diversity, yet commonality, in people's aspirations and concerns. The schoolhouse has always been a rich source of data for psychologists (Berliner, 2006; Davidson & Benjamin, 1987). However, as computers increasingly saturate education, the schoolhouse and its contemporary equivalents provide unprecedented opportunities for psychological researchers: Opportunities to make a measurable and socially meaningful impact on the lives of students and teachers; opportunities to display to society at large the benefits of a careful, psychologically-wise approach to solving social problems; and opportunities to learn about psychological theory by pushing its predictions to the limit in new contexts and at new scales of operation.

This dissertation investigates these opportunities from several different perspectives. Chapter 1 focuses on education as a context for psychological research: I elaborate on the factors that make education a rich context for psychological research, and I describe how researchers have used this context to apply and to further basic psychological theory. In Chapter 2, I focus on the the benefits, challenges, and methods of large-scale research. Chapters 3-5 each present data from a different, large-scale efficacy study. Chapter 3 presents a study of the robustness and generalizability of two social psychological interventions across a sample of over 1500 students from 13 sociodemographically heterogeneous schools. Chapter 4 describes the process of selecting

and customizing psychological interventions to address psychological obstacles to success in community college math. Chapter 5 tests the efficacy of psychologically wise encouragement in a sample of over 250,000 online learners. In Chapter 6, I explore how the local context influences students' responses to a growth mindset intervention and the relationship between individual students' mindsets and achievement. Finally, Chapter 7 reviews what we have learned about psychology through recent large-scale studies in education; it considers what new areas deserve exploration; it provides loose estimates for the economic impacts of psychological interventions in schools; and it discusses psychological interventions as a vehicle for large-scale social change.

# Acknowledgement

The ideas and results described in this document have been a central driving force in my life over the last three years. As a good friend recently observed, this work has been a labor of love, and I am extremely thankful and humbled to have shared in it with an incredible group of people who have taught me, challenged me, helped me, and laughed with me. I am glad the completion of this document marks not the end of our work together but the beginning of a new stage of collaboration.

There are far too many nice things to say about my mentors and colleagues at Stanford. Greg Walton, my doctoral advisor, has been an incredible source of support and intellectual stimulation over the last five years. More than anyone else, he has shaped the way I think about social psychology and about mentorship. My other great mentor at Stanford has been Carol Dweck. Her work on implicit theories of intelligence is what originally drew me to Stanford and to the psychological study of academic achievement, the major theme of this dissertation. Greg and Carol together provided a welcoming yet challenging intellectual home, one that exemplifies

the key combination of high standards and assurance that helps people thrive (see Chapter 1). In the course of my graduate career, they have provided me with countless opportunities for intellectual growth and potent advice about research and life.

I must also thank the other three members of my dissertation committee: Zak Tormala, Geoff Cohen, and Lee Ross. I have learned a great deal from all three of them by taking their classes, reading their writings, and discussing research.

The Department of Psychology at Stanford is a vibrant community that brings together an amazing, and eclectic, group of people. I have grown smarter—and probably also stranger—as a result of my many interactions with them. It has been a pleasure to be a sounding board for their ideas, academic and otherwise. In particular, I must thank the members of my graduate school cohort, especially Carissa Romero, Lauren Szczurek, Sarah Gripshover, Anna Merritt, Alan Gordon, Ricardo Bion, Yula Paluy, Reno Bowen, and Alex Russell — also honorary members Michael Slepian, Kody Manke, Jason Okonofua, and Lauren Howe. These people made graduate school much more fun and stimulating than I could have imagined. Thanks are particularly due to Alan Gordon for his eagerness to spend hours talking about statistics on Saturday mornings and for his ability to have fun in the process.

Last, but by no means least, I must thank the people who most directly made this dissertation possible, my colleagues at the Project for Education Research That Scales (PERTS). The work in this dissertation is the result of our combined efforts,

and I could not imagine working on a better team. The PERTS core staff: Carissa Romero, Ben Haley, Chris Macrander, Jacquie Beaubien, Eric Smith, and Rachel Herter have all worked tirelessly and selflessly on all aspects of this project. To name a few of their contributions, they have written software, applied for grants, conducted analyses, designed work-flows, recruited and managed schools, co-written manuscripts (several included here as chapters), and helped to edit this dissertation. I am proud of the work we have done together and excited about the work we have yet to do. I am also enormously thankful for the support, advice, and camaraderie of the our team's advisors, collaborators, contributors, and advocates, including Greg Walton, Carol Dweck, David Yeager, Claude Steele, Geoff Cohen, James Gross, Angela Duckworth, and Cyndi Geerdes.

I would also like to acknowledge individuals who made significant contributions to each study presented here but are not named above. Brian Spitzer helped create some materials for the high school intervention described in Chapter 3. Jascha Sohl-Dickstein and Joseph Williams helped design and set up the Khan Academy study described in Chapter 5.

The research described in this dissertation has been funded by the William and Flora Hewlett Foundation, the National Science Foundation, the Institute of Education Sciences, and the Stanford University Office of Community Engagement.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Education as a Context for Psychological Research

Education is important. It is a key pathway to social advancement and economic stability for individuals, and it is the means by which nations develop highly skilled workforces that are necessary for economic growth and competitiveness (Augustine, 2005; Baum, Ma, & Payea, 2010; Carnevale, Rose, & Cheah, 2011; Day & Newburger, 2002; Schwab, 2012). To understand and affect this important context, psychologists have been conducting research in schools for well over a hundred years (Berliner, 2006; Thorndike, 1903). In this chapter, I argue that school has been a locus for psychological research both because education is socially important and also because it is in many ways an ideal context for research.

The importance and difficulty of academic success, coupled with schools' complex social structures, make school a context full of powerful, conflicting psychological

forces. Tensions between students' goals and worries about the academic and social environment at school are myriad, and this gives researchers the opportunity to study high-pressure systems in which small changes can have big effects (Gladwell, 2000; Ross & Nisbett, 1991; Yeager & Walton, 2011). The ability to produce large effects through seemingly small but precise changes makes this type of research interesting both theoretically and practically: theoretically because it makes for clear and dramatic demonstrations of the underlying processes and practically because the ability to have a large impact with a small investment conserves resources and spurs progress.

The education context is not only socially important and psychologically rich, it is replete with high quality data. Academic records are regularly collected by schools; these records track students for long periods of time; and they do not merely index academic success — they largely define it. By working with schools, researchers effectively "outsource" the collection of a dataset that would otherwise be virtually impossible to generate. This constellation of factors makes educational contexts a good place to refine theory while enriching society.

In this chapter, I conceptualize schools as Lewinian tension systems that field researchers can influence with targeted interventions (Lewin, 1947; Ross & Nisbett, 1991). I draw on the psychological forces that have been described in the educational context, and I describe a family of social psychological field interventions that have

successfully manipulated the forces in this tension system to effect outsized gains in student achievement. I discuss the promise of these interventions for improving academic achievement broadly and for revitalizing the social psychological tradition of theory building through precise, context-wise field research. Finally, I close by discussing the obstacles to the fulfillment of these promises, a topic that I explore more fully in the following chapter.

## Tension Systems in Education

According to a recent Gallup poll, 95% of Americans say that earning a postsecondary degree is either "very important" (70%) or "somewhat important" (25%) (Gallup and Lumina Foundation, 2013). This is for good reason. Economic research suggests that Americans who graduate from high school earn more than $260,000 in additional income over their lifetime than those who do not (Rouse, 2005). Those who earn a college degree earn more than twice as much (Carnevale et al., 2011; Day & Newburger, 2002). The more highly educated also benefit from greater job security (Carnevale, Smith, & Strohl, 2010) and report enjoying their jobs more (Baum et al., 2010).

Education outcomes are at least as important for society at large as they are for the individual. According to one economic analysis, for example, the individuals who dropped out of high school in 2007 will contribute $300 billion less to the economy

than they would have had they graduated (Alliance for Excellent Education, 2007). High school dropouts between the ages of 20-67 will also contribute $50 billion less in taxes every year (Rouse, 2005), all the while requiring greater public expenditures. For example, those who drop out of high school are eight times more likely to be incarcerated (Bridgeland, DiIulio, & Morison, 2006; Harlow, 2003), at an average cost of $22,650 per state inmate per year (Stephan, 2004). More generally, better educated nations outcompete less educated nations because they generate more technologies and have more high-skilled workers to staff the most economically productive jobs (Augustine, 2005; Schwab, 2012).

The importance of educational credentials for economic success is understood by most students, who see education as a key pathway to achieving their career aspirations (Markow & Pieters, 2011). For example, in 2004, a full 87% of graduating seniors expected to attend college at some point[1] (Chen, Wu, & Tasoff, 2010).

Education makes an interesting context for research in large part because it has such a large impact on people's lives — and because the resulting struggle to succeed academically and socially makes school a whirlwind of competing psychological forces. Over the last several decades, research has identified many such forces. To give a sense for the breadth of this research and some of the well-established forces in this system, I summarize some high-level findings below.

[1]Only 67% actually did so the following year (Bureau of Labor Statistics, 2005).

- Students may adopt academic identities (Finn, 1989) or come to identify with particular subjects, e.g., math (Gainor & Lent, 1998; Lent, Lopez, & Bieschke, 1991). Those who do generally outperform those who do not. However, these academic identities may support or conflict with other identities, such as students' gender or racial group membership. In some cases, these conflicts lead students to divest their effort from school or distance themselves from the conflicting identity (Altschul, Oyserman, & Bybee, 2006; Nasir, McLaughlin, & Jones, 2008; Pronin, Steele, & Ross, 2004).

- Students work harder and are more satisfied when they are made to feel agentic and autonomous in school — when they are made to feel that they are choosing to learn (Vallerand, Fortier, & Guay, 1997; Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004, for a review, see Vansteenkiste, Lens, & Deci, 2006).

- Students form influential relationships with others. Those who perceive their teachers to be more relatable and emotionally supportive feel a stronger sense of belonging in school, pursue more adaptive classroom goals, hold higher achievement expectations, and earn better course grades (Furrer & Skinner, 2003; Goodenow, 1993; Osborne, 1997; Patrick, Ryan, & Kaplan, 2007; Wentzel, 1997, 2002).

- Students feel anxious for myriad distinct reasons in school, and these anxieties

distract them or undermine their motivation in a variety of specific ways (Erdley, Cain, Loomis, Dumas-Hines, & Dweck, 1997; Good, Aronson, & Harder, 2008; Hembee, 1990; Juvonen, Wang, & Espinoza, 2011; Steele & Aronson, 1995) — though they can learn to interpret these anxieties in more adaptive ways (Alter, Aronson, Darley, Rodriguez, & Ruble, 2010; Jamieson, Harkins, & Williams, 2010).

- Students divest their attention and efforts from school when they encounter challenges unless they see themselves as capable of improving their abilities (Blackwell et al., 2007; Dweck, 1999; Grant & Dweck, 2003; Molden & Dweck, 2006; Mueller & Dweck, 1998).

- Students exert more effort when they think of their schoolwork as relevant to their lives (Durik & Harackiewicz, 2007; Harackiewicz, Rozek, Hulleman, & Hyde, 2012; Hulleman, Godes, Hendricks, & Harackiewicz, 2010) or when they believe that learning a lot in school will help them be a force for positive change in the world (Yeager & Bundick, 2009).

The psychological aspects of the academic experience are clearly too many to review here. What the preceding list makes clear is that countless interrelated, often competing psychological factors influence students' success. These factors also appear in other contexts in which people care about achieving difficult goals and face adversity in attaining them. For example, stereotype threat was discovered in an

academic context, by observing the disparities between White and Black students' achievement after they arrived at college (Steele, 2010, p. 20). It was then experimentally investigated in a lab setting using an academic outcome (Steele & Aronson, 1995), and it has since been documented in golf-putting, car-parking, and childcare (Bosson, Haymovitz, & Pinel, 2004; Derks, Scheepers, Van Laar, & Ellemers, 2011; Stone, Lynch, Sjomeling, & Darley, 1999).

This generalizability of insights from education to other domains in which people struggle to succeed, coupled with the domain's importance and data-density, make it an outstanding context for theoretically rich, high-impact research. This potential has arguably been best realized in field interventions that wisely deflect or rebalance psychological forces in precise ways that interact with students' environments to affect long-term outcomes.

## 1.1 Intervening Wisely in School

Over the last several decades, a handful of seemingly small psychological field interventions, lasting hours or even minutes, have affected students' achievement over periods of months or years (e.g. J. Aronson et al., 2002; Blackwell et al., 2007; Cohen et al., 2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Good et al., 2003; Harackiewicz et al., 2012; Hulleman, Godes, Hendricks, & Harackiewicz, 2009; Jamieson, Mendes, Blackstock, & Schmader, 2010; Miyake et

al., 2010; Oyserman, Bybee, & Terry, 2006; Walton & Cohen, 2007, 2011; Wilson & Linville, 1985). Although they have been likened to magic (for a refutation, see Yeager & Walton, 2011), the impressive effects of these interventions were the result of a precisely contextualized application of research-derived principles. These interventions follow in the footsteps of a long tradition of intervention research in social psychology (e.g., E. Aronson, Blaney, Stephin, Sikes, & Snapp, 1978; Lewin, 1947, 1958; McCord, 1978). They import the precise methodology and theoretical insight developed through basic research (e.g., Festinger, 1954; Lord, Ross, & Lepper, 1979; Ross, Greene, & House, 1977; Steele & Liu, 1983) and apply it to field contexts with careful attention to the psychological forces already at play.

These interventions are different from traditional educational reforms. They do not convey curricular content or restructure pedagogy. Instead, they try to open students to existing learning opportunities by removing psychological barriers that may otherwise inhibit them. In doing so, social psychological interventions rely on the pre-existence of positive forces, like good teaching and students' desire to do well, to propel achievement when specific psychological barriers are removed.

Since these types of interventions rely on the displacement of existing forces, researchers designing these types of interventions must attend to the forces at play in the target environment and find deft ways to remove them. For example, Wilson and

Linville's (1982) classic attribution intervention identified a common fear among college freshmen: That they are incapable of doing well in college. Their work focused on how this fear could become a source of downward pressure for some students.

> Most freshmen enter college with trepidation about whether they can handle the work, and many are dismayed to discover that the amount of studying required far exceeds that needed to do well in high school. When academic problems first occur, as they do with many freshmen, students may see this as confirming their worst fears about their inability to succeed at college. Such attributions may cause additional worrying and anxiety, making it even more difficult to study (p. 368).

Wilson and Linville suspected that students' anxieties about an initial poor performance could lead to an exacerbation cycle — one in which worries lead to distraction that reduces performance that fuels further worries. To disrupt this cycle, Wilson and Linville drew on attribution therapy research (cf. Nisbett & Valins, 1972; Storms & McCaul, 1976; Storms & Nisbett, 1970). They reasoned that they could forestall this exacerbation cycle by helping students see their underperformance as a normal and temporary part of the transition to college. They were right. The students assigned to the reattribution treatment subsequently earned higher grades and were less likely to drop out of college over the next year. Furthermore, this field experiment provided two scientific insights: 1) Anxieties about an initial poor performance can

significantly depress subsequent performance, and 2) it is possible to alleviate these anxieties by changing individuals' interpretation for the cause of their poor performance (in this case from one that is permanent and intrinsic to one that is temporary and extrinsic).

Subsequent field interventions have to a large extent repeated this pattern while focusing on different psychological factors. Downward forces on achievement in a particular context were identified; then a way to displace them was devised and tested. In this section, I briefly summarize several social psychological intervention types that have faithfully executed this pattern to obtain long-term, positive effects on students' academic achievement. I describe how each type of intervention plays on existing forces to affect long-term achievement and discuss the circumstances under which each intervention, based on underlying theory, should be effective.

### 1.1.1 Belonging

Belonging interventions work to change students' interpretation of ambiguous cues about their social fit in a given context (Walton & Cohen, 2007, 2011). In doing so, they can make students more secure about their relationships with others and address the important need to belong (Baumeister & Leary, 1995).

As described by Walton and Cohen (2007), students' doubts about belonging may take the form of a hypothesis. Students who have reason to believe that others hold

negative expectations for them, for example, because they are members of a negatively stigmatized group, may scan their environment for evidence that they do not belong. This vigilance can make them more sensitive to any existing evidence that supports this hypothesis; it can also make them more likely to interpret ambiguous cues about their belonging as potential evidence that they do not belong. Walton and Cohen (2007) report that, when students encounter such "evidence," they feel a lower sense of fit in a given academic environment, and they may counsel students like themselves to avoid the environment. Such students are also less likely to seek contact with their instructors and often earn lower grades.

Successful social belonging interventions have attended carefully to who may feel a lack of fit in a given environment, and they have presented those individuals with information intended to help them interpret ambiguous cues about belonging in benign rather than threatening ways (Walton & Cohen, 2007; Walton, Cohen, et al., 2013; Walton, Logel, Peach, Spencer, & Zanna, in prep.). For example, Walton and Cohen (2007) created an intervention that was intended to reduce Black students' worries about belonging at an elite university in which they were numerically underrepresented. To do so, they were careful to communicate explicitly and implicitly that concerns about belonging are universal, rather than confined to members of specific groups, and that they naturally pass with time. In this way, the intervention normalized individuals' concerns and helped them to reinterpret their feelings in a

more adaptive way.

Because these interventions mitigate individuals' uncertainty about their social belonging, their effects should be confined to specific groups in specific contexts. That is, they should only work for people for whom a sense of belonging is likely to be in doubt. Belonging interventions that worked for underrepresented Black students were not expected to, and did not, influence the achievement of White students because these students were well-represented and faced no negative stereotypes in that environment.

### 1.1.2 Affirmation

Value-affirmation interventions are grounded in Self-Affirmation Theory (Steele, 1988) and subsequent laboratory studies (for a review, see Sherman & Cohen, 2006). These studies have explored how affirming one's self-worth in one domain can affect behavior in the face of threats to self-worth in other domains.

Self-Affirmation Theory contends that individuals try to preserve their self-worth and, further, that self-worth is a cross-domain psychological resource. The proposed cross-domain nature of self-worth suggests that, when a person's self-worth is affirmed in one domain, it is buffered from a variety of threats in other domains. For example, the theory holds that helping a person feel secure about her relationships with friends and family can make her less anxious about potential failures in other domains, like

performing poorly on a test.

A number of studies support this perspective. People who affirm their personal values are less anxious in evaluative situations (Creswell et al., 2005), and they are more willing to accept threatening information (Cohen et al., 2007; Sherman, Nelson, & Steele, 2000). They are also more resistant to normative social pressure (Binning, Sherman, Cohen, & Heitland, 2010) and are more likely to accept information that reflects poorly on their in-group (Adams, Tormala, & Obrien, 2006).

Because they mitigate stress and threat, value-affirmation interventions have had positive effects among students for whom stress or threat act as a barrier to success. For example, Black students may be concerned about confirming negative stereotypes about their group's intellectual abilities, and this concern may undermine their performance (Schmader & Johns, 2003; Steele, 1997; Steele & Aronson, 1995). In one value-affirmation intervention, Cohen and colleagues (2006; 2009) randomly assigned two successive cohorts of 7th grade students at an ethnically diverse school to affirm an important personal value. Black students who had completed the affirmation earned significantly higher course grades over the next two years. Additionally, they showed lower cognitive activation of words pertaining to negative stereotypes about Black Americans — suggesting that the affirmation may have worked at least in part by diminishing concerns over these worries.

Other studies have focused on value-affirmations' stress-buffering properties in

academic situations. One study by Sherman, Bunyan, Creswell, and Jaremka (2009) found that college students assigned to write about an important personal value experienced significantly lower sympathetic nervous system activation over the course of the exam period. The effect was stronger among students who were especially vulnerable to stress because of high evaluative concerns.

Convergent evidence suggests that value affirmation exercises mitigate individuals' experience of psychological threat, such as the threat of being evaluated poorly or of confirming a negative stereotype about one's group. The affirmation field interventions have extended this research by showing the generality and importance of this psychological experience in authentic learning environments. They have confirmed that members of stigmatized groups often do feel greater threat and that this threat inhibits their ability to perform at their best.

Value-affirmation may be especially sensitive to local tensions because it mitigates threat and stress. So its effects largely depend on the role of threat and stress in the environment. One may expect value-affirmation to enhance performance when such performance is encumbered by threat. In contrast, one may expect it to reduce performance in situations in which performance is enhanced by this threat (e.g., Chajut & Algom, 2003; Jamieson & Harkins, 2009; Jamieson, Harkins, & Williams, 2010; Woolf, McManus, Gill, & Dacre, 2009).

### 1.1.3 Wise Feedback

Wise feedback interventions change students' interpretations of critical feedback (Cohen, Steele, & Ross, 1999; Yeager et al., in press). Critical feedback helps students grow and learn (Hattie & Timperley, 2007); it is thought to be one of the most important contributors to student achievement (Hattie, 1999). However, if delivered or interpreted in maladaptive ways, it can dampen students' self-worth and motivation. Wise feedback interventions help students interpret critical feedback adaptively.

A critical component in wise feedback interventions is the development of trust between the student and the teacher. Unsurprisingly, students who do not believe their teacher is acting in their best interest are less likely to accept feedback and learn from it (Bryk & Schneider, 2002; Cohen et al., 1999). So trust is an important precondition for students to benefit from feedback. Creating trust is especially important for negatively stereotyped students, who may worry that feedback is a sign of bias or prejudice (Cohen et al., 1999; Yeager et al., in press). Instead of trying to fix their mistakes, students who do not trust their teachers may instead disengage and miss out on important opportunities to learn.

Wise feedback interventions indirectly assure students that criticism of their work results not from bias but from a firm belief in their potential to reach a higher standard (Cohen et al., 1999). To effect this interpretation of critical feedback, wise

interventions must communicate teachers' high expectations and the belief that students can reach those expectations (Cohen et al., 1999; Yeager et al., in press). For example, Yeager and colleagues (in press) randomly attached different post-it notes to critical feedback about an essay that students completed for class. In the wise feedback condition, this note read, "I'm giving you these comments because I have very high expectations and I know that you can reach them." This note helped students interpret the teacher's criticism as a sign of their high standards, rather than as a sign of bias.

While these interventions help students develop trust with their teachers, their effects have been limited to stigmatized groups. For example, these interventions have been shown to work for minority students in academic contexts (Cohen et al., 1999; Yeager et al., in press). Majority group members generally do not benefit from this type of feedback in the same way, presumably because they have less reason to interpret it as a sign of prejudice or bias..

### 1.1.4 Theories of Intelligence

Research on people's implicit theories has revealed that the extent to which people view characteristics as malleable or fixed can influence their responses to setbacks and their interest in seeking challenges (Dweck & Leggett, 1988). Intuitively, the extent to which an individual views a particular quality, like intelligence, as malleable should

influence their goals concerning developing that quality. Why would someone work hard to change something that they believe cannot be changed? Similarly, negative information about a personal quality should be less worrisome if the quality is viewed as malleable rather than fixed. In the former case, the negative information identifies an area for improvement; in the latter, it directs attention to an indelible personal flaw.

A considerable body of empirical evidence supports this intuitive account. When individuals view qualities as malleable rather than fixed, they tend to respond to setbacks more adaptively across a variety of contexts from test-taking and studying to school bullying and international conflict (Bar-Tal, Halperin, & Oren, 2010; Mangels, Butterfield, Lamb, Good, & Dweck, 2006; Molden & Dweck, 2006; Mueller & Dweck, 1998; Yeager, Trzesniewski, Tirri, Nokelainen, & Dweck, 2011). Students' beliefs about the malleability of intelligence seem to be particularly relevant to their feelings about school.

Students who believe intelligence to be malleable are more likely to adopt mastery goals and to earn higher grades over the course of middle school (Blackwell et al., 2007). For these students, school is an opportunity to improve themselves — to become smarter — and effort is a route to growth. For students who believe that intelligence is fixed, on the other hand, effort is sign of insufficient intrinsic ability, and school is a place that can either validate or impugn one's intelligence — not a place

that can grow it. These divergent perspectives on effort and evaluation can lead to entirely different academic goals, priorities, and anxieties (Dweck, 1999; Molden & Dweck, 2006). For example, students who are led to adopt a fixed view of intelligence are more likely to overstate (lie about) their scores on an achievement test (Mueller & Dweck, 1998), presumably because they do not want others to view their fixed abilities in a negative light. Students who endorse fixed views of intelligence are also less likely to pay attention to and benefit from critical feedback that could help them improve (Mangels et al., 2006).

Several interventions targeting students' beliefs about the malleability of intelligence have shown that these beliefs are themselves malleable and that they have a causal influence on achievement (J. Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003). For example, one intervention taught middle school students that intelligence is malleable in the course of two sessions that were part of an eight-session course about the brain. Students who were taught about the malleability of intelligence were more likely to be named by their teachers as students who experienced significant improvements in classroom motivation; they also earned significantly higher grades in math over the following academic term. More recently, computer-administered versions of this intervention also succeeded at increasing scores. In one study, middle school students who were in the bottom half of their class by GPA experienced a significant increase in GPA in the trimester after they were taught about

the malleability of intelligence (Romero, Paunesku, & Dweck, 2011).

### 1.1.5 Self-Relevance

Research has shown that when students find school more relevant to their own lives, they are more motivated and engaged in academics (Harackiewicz et al., 2012; Hulleman & Harackiewicz, 2009; Jang, 2008).

School can seem pointless to students who have trouble seeing how the things they learn in class are relevant to their lives. Students who fail to see the relevance of schoolwork tend to lose interest in coursework, especially in math and science classes for which self-relevance is less immediately apparent (Brophy, Biswas, Katzlberger, Bransford, & Schwartz, 1999; Hidi & Harackiewicz, 2000). When students are disinterested, they are more likely to opt out of more advanced classes (Harackiewicz et al., 2012), and in extreme cases, even drop out of school (Bridgeland et al., 2006). On the other hand, when students see how school is relevant to their lives and career aspirations, they invest more effort in school and perform better (Harackiewicz et al., 2012; Hulleman et al., 2009; Hulleman & Harackiewicz, 2009).

Research has shown that it is possible to increase academic relevance through "self-relevance" interventions that ask students to write about how particular concepts or lessons learned in class are relevant to their lives (Hulleman & Harackiewicz, 2009). The success of self-relevance interventions at raising interest and achievement

in school, particularly in STEM subjects (Harackiewicz et al., 2012; Hulleman & Harackiewicz, 2009), provides direct causal evidence that students underperform and under enroll in these courses because not enough is done to link these courses to their personal lives.

These interventions also reveal that helping students see the relevance of their coursework need not be difficult or expensive. Some schools and academic enrichment programs try to make schoolwork personally relevant through internships and research experience (Barnard College, 2013; University of California Santa Cruz, 2012), and many have had success doing so (Packard & Nguyen, 2003; Stake & Mares, 2001; Yamauchi, 2003). Such programs can no doubt be a positive force for student engagement; however, the results of the psychologically wise self-relevance interventions show that the benefits of making school more relevant can be realized far more easily and cost-effectively, through well-timed conversations between students and their parents or by setting aside time for students to reflect for themselves on how their schoolwork is relevant (Harackiewicz et al., 2012; Hulleman & Harackiewicz, 2009). Schools and parents can take such steps immediately at virtually no cost and without suffering through the often long, expensive process needed for structural reform.

### 1.1.6 Sense-of-Purpose

Sense-of-purpose interventions are related to but distinct from the self-relevance interventions described above. They draw on a long line of research which shows that students exhibit greater tenacity when they connect their schoolwork to a sense of purpose that encompasses a commitment to causes that transcend the self (Damon, Menon, & Bronk, 2003; Deci & Ryan, 2000; McKnight & Kashdan, 2009; Yeager & Bundick, 2009). For instance, high school students who describe their motivation as a desire to contribute to society focus more on learning and less on merely avoiding failure (Lee, McInerney, Liem, & Ortiga, 2010). A number of experimental studies also demonstrate that linking one's current learning to a meaningful purpose can lead to greater motivation and deeper learning (Jang, 2008; Vansteenkiste et al., 2004).

This research suggests that helping students see the beyond-the-self purposes of their schoolwork can lend to an inspiring — and motivating — interpretation of otherwise boring activities. For example, a purpose for doing well in science class may be to become an engineer who can make devices that make people's lives better (a beyond-the-self goal). This type of purpose is quite different from extrinsic motives, like earning a good grade or getting a high-paying job. References to these latter, extrinsic, motives are ubiquitous in popular culture, and they are used frequently by adults to exhort students to do well (Ames, 1992; Boggiano & Katz, 1991; Deci & Ryan, 2004). So it is no surprise that many students cite such motives when asked

about the reasons they do schoolwork (Yeager & Bundick, 2009).

Notably, extrinsic motives are not well aligned with learning deeply. They focus students on demonstrating ability to others in exchange for credentials and financial rewards. From this perspective, schoolwork — especially boring or difficult schoolwork — is an obstacle to achieving one's ultimate goals, and avoiding work or doing the least amount needed to meet the external criteria is reasonable, even advisable. In contrast, students who are motivated by their desire to use what they learn to contribute to a beyond-the-self purpose can see tedious and difficult schoolwork differently. For these students, schoolwork provides a chance to build skills that will help them make a positive impact on others. Interpreting schoolwork in this way — as in line with and enabling the pursuit of a self-relevant long-term goal — may help these students bring to bear additional effort and self-control (Fishbach & Trope, 2005; Fishbach, Zhang, & Trope, 2010; Loewenstein, 1996; Mischel, Cantor, & Feldman, 1996; Rachlin, 2000; Thaler & Shefrin, 1981; Trope & Fishbach, 2000; Zhang & Fishbach, 2010).

Students who interpret their work through this prosocial lens may also circumvent certain cultural pressures against high-achievement that is viewed as individualistic or competitive. Researchers have long observed that high-achieving students may get stigmatized as a "teacher's pet" or "curve raiser" and face social costs, e.g., have fewer friendships (Coleman, 1961; Fryer & Torelli, 2010; Tyson, Darity, & Castellino,

2005). However, the perceived motivations of high-performing students affect how peers respond to them. Students who adopt prosocial, beyond-the-self motivations may sidestep certain social pressures against high achievement.

Consistent with this assertion, research finds that high achievers who signal communal as opposed to competitive or individualistic values are better respected by their peers (Boykin et al., 2005; Marryshow, Hurley, Allen, Tyler, & Boykin, 2005), especially in ethnic minority and low-SES cultural contexts that emphasize communal and cooperative goals over individualistic or competitive goals (American Psychological Association, 2002; Tyler et al., 2008). For example, a number of studies report that Black elementary school students perform better on a variety of academic activities, including math, social studies, and language arts exercises, when those activities are presented as communal as compared to competitive or individualistic (Boykin, Lilja, & Tyler, 2004; Dill & Boykin, 2000; Hurley, Boykin, & Allen, 2005). This research suggests that sense-of-purpose interventions may help students achieve in part by licensing them to work harder in school. Unlike the more typically invoked individualistic motives around professional prestige and financial success, the motives invoked by sense-of-purpose interventions are communal in nature. In that respect, they may be more consistent with the cultural values of students from traditionally underperforming groups, and they enable students to pursue academic mastery without experiencing cultural dissonance or incurring social costs from their peers.

To review, purpose interventions may work through two distinct paths: By changing students' interpretation of schoolwork to be more meaningful and motivating and also by licensing students to try their best without defying communal cultural practices. No sense-of-purpose interventions have yet been published. However, several manuscripts in preparation, including two reported in this dissertation, describe how these interventions have helped students view mundane schoolwork in more meaningful ways (Yeager, Henderson, et al., 2013) and how they have elevated academic achievement among low-performing students (see Chapters 3 and 4). The success of these field interventions suggests that students do want to make a positive impact on the world and are more motivated in their schoolwork when they see it in this light. It also suggests that students' desire to do good in the world is a relatively untapped motivational resource in normal school environments.

## 1.2 Fulfilling the Promise

"If you want truly to understand something, try to change it!"

- Kurt Lewin (as cited in Stam, 2006)

The interventions described above in many ways represent a return to the original values of a discipline that made its mark by shedding light on the powerful, unseen mechanisms underlying many of society's biggest problems (Allport, 1954; E. Aronson et al., 1978; Darley & Latané, 1968; Lewin, 1947; Milgram, 1965; Zimbardo,

1969). Prominent figures have recently lamented that social psychology has to some extent retreated from the real world to the laboratory and, in the process, lost its focus on the questions that matter to society (Cialdini, 2009; Reis & Gosling, 2010; Ross, Lepper, & Ward, 2010). The methodologically precise and psychologically insightful interventions reviewed above may be an antidote to this unwelcome trend: They are simultaneously practically useful (satisfying the public) and scientifically informative (satisfying academic psychologists).

Social psychological interventions that affect academic achievement are compelling to the public and policy makers because they reveal unique, often low-cost ways to affect important social outcomes. In principle, many of these interventions could be administered at a price ranging from a few pennies to several dollars per student. For example, the seminal value-affirmation interventions were only a few sheets of photocopied paper handed out to students by their teachers during class; they took only 15 minutes to complete (Cohen et al., 2006). This is far less costly than structural reforms like decreasing class size (Ilon & Normore, 2006), creating tailored learning communities (Weiss, Visher, Wathington, Teres, & Schneider, 2010), or setting up conditional cash transfers (Goldrick-Rab, Harris, Benson, & Kelchen, 2011).

Educational field interventions also present a unique context for research. To paraphrase Kurt Lewin, they let researchers test if they truly understand processes by trying to change them. That is, they give researchers deep insight into the tensions

that govern people's worries, motivations, and behaviors by providing the opportunity to see what happens when those tensions are displaced.

Although such interventions hold significant promise for education reform and for fascinating research, questions about their robustness and generalizability must be answered before this promise can be fulfilled. Can these interventions work consistently in the hands of non-experts (are they foolproof)? Can they work in a variety of different educational contexts? To date, field experiments testing these interventions have utilized relatively small samples and were administered with considerable direct involvement from researchers. For example, in a field intervention conducted by Blackwell and colleagues (2007), 16 research assistants were personally trained by researchers to conduct 8 weeks of workshops with a total of 99 students. The high costs of a procedure like this in terms of researcher time, class time, and transportation severely limit the scalability of this type of protocol.

Therefore, before policy change can take place, we must determine if these interventions retain their effectiveness even when they are delivered in a scalable format — one that is cost-effective and does not rely on fixed resources (e.g., the time of preeminent psychological researchers). We must also determine whether they work in many different contexts.

In order for research to investigate these topics, the financial and logistical barriers to large-scale experiments in schools must be greatly reduced. Traditional education

experiments are far too costly to answer the myriad questions relevant to the scale-up of these interventions. For example, the Institute of Education Sciences awards an average of $2.6 million to evaluate the efficacy of a single intervention, usually with a sample of several hundred students (Institute of Education Sciences, 2012). Given that the effect sizes reported for social psychological interventions have been in the $d = .1$ to $d = .5$ range, a grant supporting an experiment with 500 students could only support 1-4 experiments with adequate (80%) statistical power.

In addition to being financially costly, field interventions are also far more time consuming than lab experiments, and they are less likely to provide the "clean" mediational data that has come to be expected in top disciplinary journals within social psychology (Cialdini, 2009; Wilson & Linville, 1982). This gives academic psychologists a strong disincentive to pursue such work.

In summary, for these interventions to realize their potential — to facilitate a better educated society and to catalyze a return to a more field-centered social psychology — a new research paradigm is needed. This paradigm must enable researchers to answer questions about the practical utility of these interventions and do so in an economically efficient way that stops short of monopolizing researchers' time. The next chapter explores how this type of research might look.

# Chapter 2

# Scaling Up Social Psychology

## 2.1 Introduction

Over the last several decades, social psychologists have learned a great deal about changing human behavior in productive ways. By paying close attention to the psychological pressures influencing people's behaviors and by drawing on insights from psychological theory to redirect those pressures in targeted ways, they have crafted brief interventions to: increase voter turnout (Bryan, Walton, Rogers, & Dweck, 2011); improve relationship stability (Finkel, Slotter, Luchies, Walton, & Gross, in press); reduce power consumption (Allcott, 2011; Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007); and raise academic achievement (see Chapter 1 for a review). The creation and evaluation of these cost-effective interventions draws on certain psychological technologies — theoretically grounded and locally contextualized techniques, approaches, and practices — that hold the potential to alleviate

important social problems while advancing knowledge.

In exercising and advancing these technologies, researchers can simultaneously improve society, elevate the status of psychological science, and learn how psychological factors influence people over time, in interaction with the social environments they inhabit. This is an exciting opportunity. It is also a great challenge because it pushes psychological research to consider new questions and work at larger levels of analysis. This challenge deserves our field's attention because questions about the robustness, cost-effectiveness, and generalizability of these interventions must be answered before they can be responsibly disseminated.

Below, I describe why these questions are especially important in the context of intervention research. I also describe why large, diverse samples are necessary, though not sufficient, for answering these questions while maintaining the public trust that is critical for field research. In the subsequent section, I use the work of the Project for Education Research That Scales (PERTS) as a case study in scaling up social psychological research.

## Robustness & Generalizability

In describing when education innovations should be broadly disseminated, Bryk (2009) argued that it is critical "to know how to make [the innovation] work reliably over many diverse contexts and situations." Indeed, this is true of any innovation.

To be socially useful, an innovation must work consistently over time and in different environments. Often, they do not (Dusenbury, Brannigan, Falco, & Hansen, 2003; Fullan, 2007; Graczvk & Weissberg, 2003; Labaree, 1998; Tyack & Cuban, 1995). Carefully crafted practices that show promising effects in initial, small-scale evaluations often lose their effectiveness as they get implemented more broadly by individuals with less expertise and lower incentives to maintain fidelity (Berman & McLaughlin, 1976). For example, in initial efficacy trials, researchers may personally visit a school to increase faculty support for an intervention study or to learn how to tailor the intervention to the local context. In doing so, however, they almost guarantee that the intervention will be administered more faithfully and with better targeting in their study than it will in subsequent practice. That is, in subsequent practice, a given intervention would probably be less effective because it would be administered by a less motivated and less expert team and in a different context from the one for which it was tailored.

Of course, it is perfectly reasonable for researchers to first test whether a novel intervention can work under ideal circumstances. That is an advisable first step in any research program. For example, trials for clinical interventions separately consider their efficacy (performance under ideal circumstances) and effectiveness (performance under realistic circumstances) (Flay et al., 2005; Rush, 2009; U.S. Food and Drug

Administration, 1998). However, the later steps, i.e., tests of effectiveness under realistic circumstances, often never take place in social psychological research. In part, this is because publishing practices disincentivize academic researchers from conducting replication studies (Giner-Sorolla, 2012; Koole & Lakens, 2012) and because larger scale work is unavoidably more logistically difficult. In practice, this means that many promising innovations never get broadly implemented or get implemented despite questionable effectiveness at scale.

Therefore, researchers who want their interventions to be broadly applied should strive to test them under realistic circumstances. That means that they should minimize the degree to which they personally supervise or tailor an intervention in ways that would be impractical in subsequent dissemination. They should also test interventions in multiple, diverse contexts because doing so decreases the probability that the eccentricities of a particular context will drive (or inhibit) an effect.

Both of these practices — more diverse samples and less enforcement of fidelity — are likely to increase the variability of outcomes and decrease observed effect sizes. That means that researchers conducting these more realistic evaluations must increase their sample size to compensate for the loss of statistical power.

## Cost & Effectiveness

Researchers who intend for their interventions to get broadly disseminated should also consider cost-effectiveness. At their best, intervention studies can affect the real world by providing decision-makers with the information they need concerning the likely benefits of policy changes. For such decisions to be wise, they must weigh the likely costs and benefits of implementation, and an intervention's effect-size is central to the assessment of its benefits.

Theorists have discussed at length when and how effect-size measurements are important for research and theory development, and they have put forward arguments for why and when conventionally "small" effect-sizes can be "impressive" (e.g., Prentice & Miller, 1992). In the context of basic research, this can make for interesting and nuanced intellectual debate. In the context of applied research, however, the arguments are rather more clear-cut.

In applied research, it would be patently vacuous to base a judgement of "importance" solely on statistical effect size because effect size is just one term in a cost-benefit formula. It would be like trying to determine the volume of a box using only its width. To determine whether an intervention should be implemented, one must know how big its effect is statistically, how important this effect is practically, and how much it costs to obtain this effect. To give an extreme example, if an intervention extended humans' lifespan with an effect size of $d = .05$ (which is statistically

"small"), it would equate to approximately 1 extra year of life and would clearly be practically important. At that point, this benefit would have to be considered in reference to its cost and compared to other policies competing for the same resources.

Although statistical effect size is only one of several factors to consider when evaluating the promise of an intervention, it is important to know it with some degree of certainty. If there is a great variability in an expected effect size, there will also be great variability in the expected social benefit, making it hard to determine whether, where, and at what cost an intervention is worth scaling up.

Critically, the variability of observed effect-sizes varies considerably with statistical power. As power decreases, the variability of observed effect-sizes increases. For example, assume that a treatment has a true effect-size of $d = 0.2$, which is the average reported effect size across all psychological studies in the 20th Century (Richard, Bond, & Stokes-Zoota, 2003). Simulation studies demonstrate that, even if this intervention actually produced an effect of $d = 0.2$, random variability would cause the observed effect of this treatment to be negative (below zero) approximately 1% of the time with a sample of 1000 (500 per cell), 9% of the time with a sample of 200, and 35% of the time with a sample of 30. In other words, an intervention that actually exerts an effect of $d = 0.2$ would, by random chance, appear to have a negative effect almost 1 in 10 times with a sample of 200 and more than a third of the time with a sample of 30! Furthermore, with these sample sizes, it would only reach conventional

levels of statistical significance ($\alpha = .05$) if the observed effect were actually larger than the true effect! If the true effect-size were empirically observed, it would only reach a conventional level of statistical significance with samples of 386 (193/cell) or more. That is to say, when evaluated in an underpowered study, an actually effective intervention would appear to be harmful a sizable percentage of the time, and its positive effects would only be detectable at a statistically significant level a small percentage of the time.

Therefore, researchers who hope to influence policy should strive to make their effect-size estimates as accurate as possible by harnessing sufficient statistical power. Of course, sufficient power alone is not enough to guarantee that an intervention will be as effective in implementation as it was during a field study. As discussed above, an intervention's effect also varies with the fidelity of implementation and the similarity between the population in which it is tested and subsequently implemented. However, when statistical power is insufficient, the observed effect-size would provide an inaccurate frame of reference for the subsequent effect-size, even if the population were perfectly matched and the treatment fidelity were extremely high.

The other key component to assessing cost effectiveness is cost. Most psychological interventions hypothetically have low monetary costs; for example, they may involve students completing easy-to-photocopy reading and writing exercises (e.g., Cohen et al., 2006; Hulleman & Harackiewicz, 2009) or showing students survey data and

then video recording their responses to it (e.g., Walton & Cohen, 2007). These activities are therefore seemingly no more expensive in terms of dollars, teacher time, or student time than a single math worksheet or a high school video project. All of these are plentiful, low-cost resources.

However, a focus on the tiny material costs of these interventions ignores the potentially enormous costs of the expertise required to administer them with sufficient fidelity. Theorists have explained that the impressive effects of these interventions are made possible by the carefully choreographed psychological experiences they elicit (see Yeager & Walton, 2011). To the extent this is so, a part of the true cost of these interventions must account for the great pains that eminent researchers go through to ensure a school site executes their protocol faithfully. If such on-site supervision is actually necessary, then these interventions are far more expensive than advertised. Not only that, they may also be fundamentally limited in scale by the small worldwide supply of expert researchers.

I do not mean to argue that these interventions are in fact expensive or that they truly require on-site supervision from a full professor in a top-tier psychology department. The results presented later in this dissertation provide direct evidence to the contrary. However, I do mean to argue that traditional efficacy studies, which provide extensive on-site supervision, training, and encouragement to intervention facilitators, leave open the possibility that these "extras" are required for effectiveness.

To the extent that field experiments can avoid such hard-to quantify (and potentially impossible to scale-up) extras, their results provide a better estimate of their robust, scalable benefits.

## Public Perceptions

Another reason to use large samples in field research relates to public perceptions. Compared to laboratory studies, field studies are more likely to involve partnerships with external institutions that provide funding or research subjects. These institutions, unsurprisingly, often want to know the outcomes of studies that they fund or participate in.

When an underpowered intervention study fails to detect the effect of a particular treatment (as is likely to happen), policy makers and the public, who lack a nuanced understanding of statistical power, are likely to perceive such failures as disconfirming evidence of a treatment's effectiveness, even though it should be interpreted instead as an absence of evidence in either direction. In this way, underpowered studies substantially increase the risk that an effective treatment will come to be perceived as ineffective by the individuals in a position to fund and implement it or by those in a position to materially or politically support subsequent research.

## 2.2 A Case Study: The Project for Education Research That Scales

In this section, I use the Project for Education Research That Scales (PERTS) as a case study for the creation of social psychological interventions with the goal of broad-scale evaluation and dissemination. I describe the circumstances that led to the project's formation and the principles that guide its work; the logistical restrictions we placed on ourselves to ensure that our field interventions were tested in realistic ways and that enough students participated to assess their generalizability; how we chose interventions that would be maximally effective given these restrictions; and the modifications we made to maximize the effectiveness of the interventions we selected. Although PERTS is focused on the academic context, many of the principles described here could apply to other domains as well.

### Recognizing the Opportunity

As detailed in Chapter 1, a number of small but rigorous field experiments demonstrated that brief social psychological interventions could have long-term effects on academic performance (J. Aronson et al., 2002; Blackwell et al., 2007; Cohen et al., 2006, 2009; Walton & Cohen, 2007, 2011). These interventions showed surprisingly large effects on academic outcomes considering that they took from only 15-minutes

(Cohen et al., 2006) to 3.3-hours (Blackwell et al., 2007) of class time to administer. Additionally, they seemed to have the potential for extremely low implementation costs. For example, the value-affirmations evaluated by Cohen and colleagues (2006) required only a photocopied piece of paper and the growth mindset activities administered by Blackwell and colleagues' (2007) research assistants appeared, at least in principle, to be translatable into a text-based format. If these interventions could be administered broadly, efficaciously, and cost-effectively, there would be many direct and indirect benefits.

Most directly, millions of students could potentially experience academic gains that could translate into more enjoyable, higher paid jobs for them and greater economic productivity for society. To the extent that certain interventions continued to disproportionately affect students from stigmatized groups, the broadly disseminated interventions could also reduce group-based achievement gaps by as much as 40% (see Cohen et al., 2006). Additionally, because the interventions were brief and delivered directly to students, these benefits would not require politically difficult structural changes or commitments to extensive teacher professional development programs.

Indirectly, the successful scale-up of these interventions could have even farther-reaching impacts on education and psychological research. If these interventions achieved broad-based success, they would encourage more attention to students' psychology among educators and education policy makers. Whereas educational reforms

usually focus on curriculum and pedagogy, the broad success of psychological interventions would signal the importance — and productivity — of considering students' psychological state as well. This would spur more research about effective psychological interventions and perhaps also about psychologically wise ways to reform education practice more holistically.

In short, existing evidence suggested that the potential benefits of the successful scale-up of these interventions could be enormous. However, it was equally clear that broad-scale dissemination would be premature without further study. It was not known whether these interventions would work broadly or without the close supervision of researchers who could not feasibly supervise programs in tens, to say nothing of thousands, of schools every year.

## Building the Infrastructure for Scalable Research

To learn whether social psychological interventions could broadly and robustly affect achievement, we needed to build a new type of infrastructure for educational research. This infrastructure would need to let us deliver and evaluate these interventions with many students in many geographically disparate schools. It would also need to enable us to do so in a way that would maintain a relatively high degree of fidelity without demanding too much time and effort from participating schools (otherwise it would be too challenging to recruit them). Furthermore, we wanted to ensure that

successful interventions could be administered to many more students without major modifications because, if modifications were required, so too would new rounds of evaluation.

To accomplish this, we created an online platform to administer the interventions and simple logistical processes that could be applied consistently across schools. After a school expressed interest in participating in an intervention, a brief phone call (15-30 minutes) was arranged to explain the participation procedure to the person who would act as the site coordinator; usually this was a teacher or counselor at the school who would then recruit and orient their colleagues to the study. In addition to this brief phone call, there were also 5-10 emails between research staff and the site coordinator until after the intervention, at which point this individual was asked to send academic records to the research team. Following the phone call, the site coordinator then handled the small amount of "training" that their colleagues required before they could administer the intervention. In some schools this meant holding a faculty meeting to explain the registration process to participating teachers; in other cases, coordinators simply emailed instructions to participating teachers.

Several things about the PERTS process described above should be noted. First, the amount of time needed was far lower than in traditional interventions in which there are usually several long, face-to-face meetings between multiple research-team members and multiple individuals at each site (often all of the participating teachers).

Besides requiring far less time of fewer people, this approach cut travel time and obviated complex scheduling between researchers and school staff, who are likely to have different, busy schedules. Because the process could be asynchronous, it was far more flexible. Whereas multi-person face-to-face meetings may have to be scheduled weeks or months in advance, asynchronous coordination has fewer dependencies and can take place over the course of a few days. The use of a site coordinator provided much of the benefit of having staff visit the site, but it did so in a scalable way. The ability to coordinate remotely also removed geographic barriers that would have otherwise made working with distant schools prohibitively difficult.

Once oriented, teachers registered at perts.net. This was a standard, 5-10 minute process during which they provided their email, name, and demographic characteristics. They would also agree to follow study procedures, e.g., download a parent information form and send it home to parents. Upon registration, teachers would enter a label for each participating class and schedule when it would participate. This scheduling also let researchers and site coordinators track whether a particular teacher was on schedule to complete study activities when they said they would (see Figure 1). This was done through a web panel showing the schedule and participation status across dozens or hundreds classes in a school or study (see Figure 2). It is important to note that the flexibility of the automated scheduling enabled teachers to plan the intervention activities around their own schedule without difficulty; this

made the intervention far less disruptive to course plans.

The benefits of the online platform were equally pronounced during the completion of the online materials, i.e., surveys and interventions. The use of online materials instead of paper and pencil materials meant that intervention materials did not need to be transported and that students' responses to survey questions did not need to be transcribed; this saved time and money. It also meant that progress could be tracked in real time, as students answered questions.

Another key aspect of the scalable infrastructure was shortening the amount of class time required for each study. We restricted intervention studies to 1-3 sessions of one class period each, i.e., 45-135 minutes of class time in total would be devoted to the studies. The small time footprint helped recruit schools that may have otherwise been hesitant to cede significant class time.

A final important aspect to this infrastructure was the iterative process we went through to refine it and make it easy for schools to use. Our staff visited school sites during pilot testing to check for anomalies and errors that would be disruptive to students' experiences. We also actively sought feedback from participating students and teachers through a brief online survey at the end of the final online activity. At the end of the study, we also scheduled a 15 minute debriefing phone call with each school site coordinator. This feedback improved the experience in future iterations.

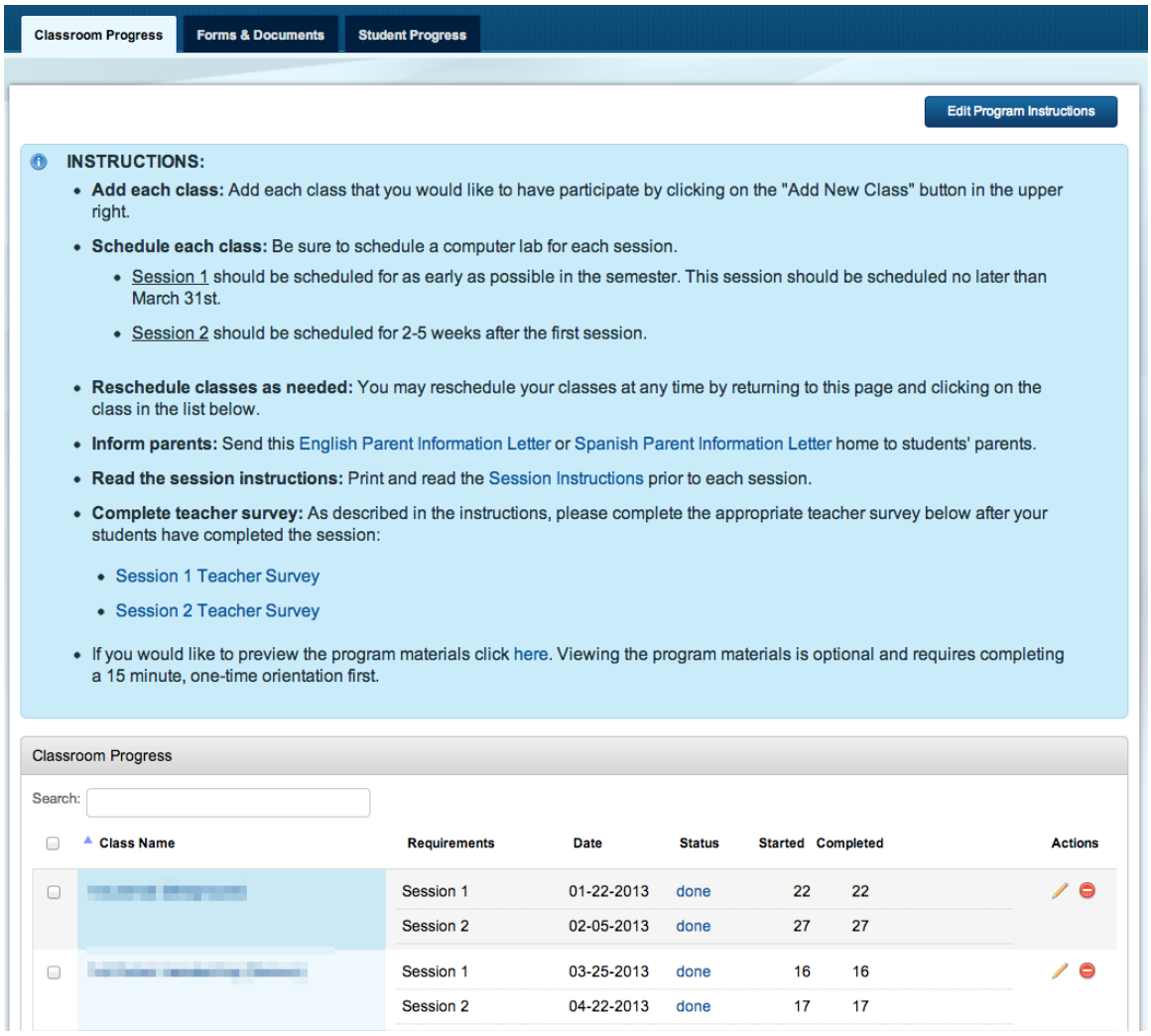



Figure 1: *Teacher Panel.* When teachers are logged in, they can add and schedule classes, download necessary materials, and take the teacher survey.

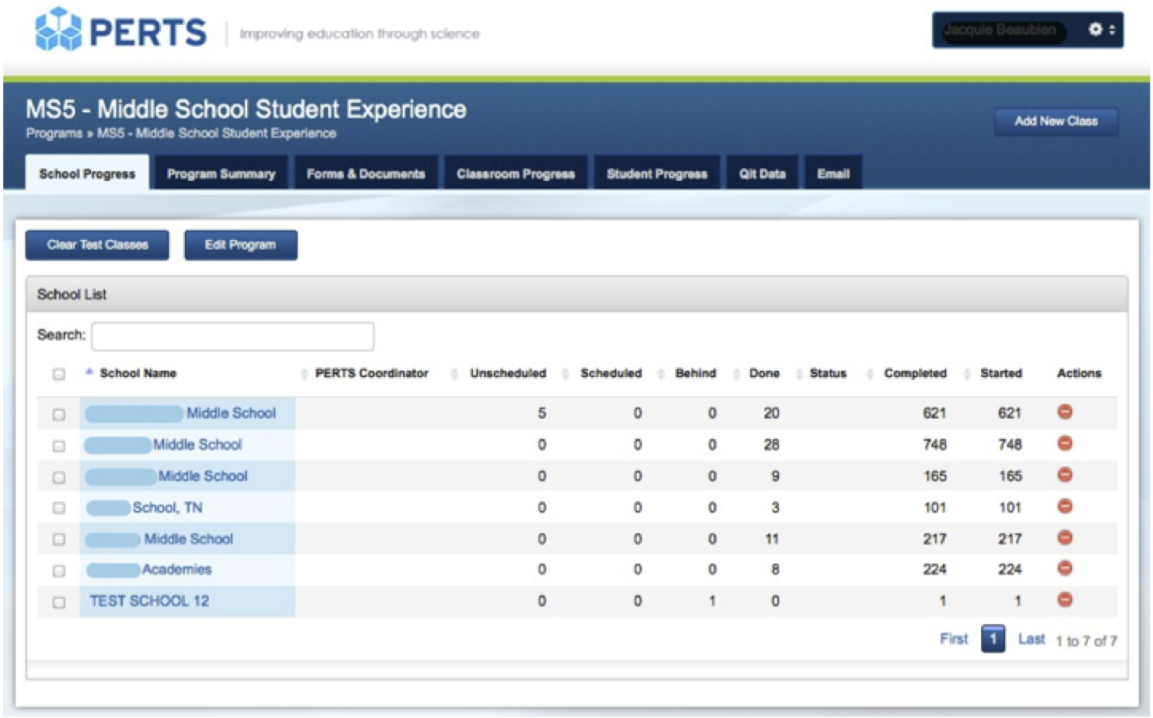



Figure 2: *Researcher panel.* A researcher panel showing study progress across multiple schools. These progress summaries can be viewed at multiple levels, by program (as above), by school, or by individual class.

## Scaling Up the Psychological Experience

Within the restrictions set by our infrastructure, our goal was to create materials that would consistently evoke powerful psychological experiences for many different groups of students. That meant we had to 1) choose interventions that would be most likely to affect achievement across a broad range of students and 2) distill the key messages so that they were short and simple but psychologically powerful.

A priori, we did not know what types of schools would get recruited to our sample. So it was important to choose interventions that affect common psychological barriers across many different types of schools. For example, value-affirmation or sense of belonging interventions were not ideal candidates given these restrictions because they would only be predicted to yield positive effects for certain groups and under specific circumstances. Both of these interventions selectively influence students from stigmatized groups (Cohen et al., 2006; Sherman et al., 2013; Walton & Cohen, 2007). Furthermore, these interventions alleviate group based-concerns that are more likely to be active when one's group is underrepresented in the context: Prior research suggests these interventions may not apply to members of stigmatized groups in contexts in which these groups are adequately represented (Inzlicht & Ben-Zeev, 2000; Logel, Walton, Peach, Spencer, & Zanna, 2010). When representation is adequate, the psychological pressures that belonging and affirmation interventions act

upon are less likely to depress students' achievement. Therefore, we generally focused scale-up efforts on growth mindset and sense-of-purpose interventions because these interventions target psychological barriers likely to exist in a wide variety of educational contexts (see Chapter 1).

In addition to choosing interventions that were likely to be broadly applicable, we worked to distill the key messages and create intervention activities that would help students effectively internalize them. First, the interventions were never presented as remedial so as to avoid reactance or defensiveness and to avoid stigmatizing students (cf. Sherman, Cohen, et al., 2009; Yeager & Walton, 2011). Instead we used a "stealthy" approach: Teachers were instructed to describe the activities as a study about learning rather than an intervention. Following in the footsteps of earlier successful interventions (e.g., J. Aronson et al., 2002; Walton & Cohen, 2007; Wilson & Linville, 1982), we explained to students that we needed their help to communicate the intervention content to other students in their own words. This tactic was employed to defuse defensiveness and to get students to internalize the ideas by publicly endorsing them (cf. J. Aronson et al., 1999; Bem, 1965; Cooper & Fazio, 1984). This writing exercise also enabled students to "self-customize" the intervention message and to author its implications for their own lives. This is important because prior research suggests that students are more responsive to intervention messages when given the opportunity to self-generate them instead of

simply being told to accept them (Godes, Hulleman, & Harackiewicz, 2007).

Finally, the materials were designed to communicate norms about the intervention messages in ways that would mitigate defensiveness and model acceptance and integration of key ideas. For example, in the sense-of-purpose intervention, we were concerned that students may think it counter-normative to frame their learning in terms of the prosocial, beyond-the-self goals that it would help them achieve. To dispel these concerns, we presented survey results suggesting that many students think about school in these terms (cf. Prentice & Miller, 1993). All the while we tried to increase acceptance of this information by explaining that it is normal for students to not initially think this way but to come around to it over time; this tactic was intended to diffuse the possibility that students would think themselves foolish for not having already thought about school in this way. We also modeled the process of conceptual transformation that we wanted students to have using vignettes that described students' thinking changing over time through self-reflection, e.g.:

> For me, getting an education is all about learning things that will help me do something I can feel good about, something that matters for the world. I used to do my homework just to earn a better grade, but now I realize it's much more important than that. Doing well in school is all about preparing myself to do something that matters, something that I care about.

## Summary

By paying close attention to the psychological forces at play in real-world environments and by designing interventions to deflect those forces, social psychologists can effect positive changes in people's lives and causally test the strength of the hypothesized forces. Such work presents unique opportunities to help mitigate important social problems while advancing knowledge. However, before these types of interventions can fulfill their promise through broad, effective dissemination, their robustness must be tested. As I argue above, this type of testing requires an explicit focus on scalable intervention approaches: Approaches that can be tested with large, diverse samples and without significant "hidden inputs," like the supervision of expert researchers. Adherence to these approaches ensures that field interventions can, once shown to be effective, be broadly disseminated without significant alteration and with reasonably estimated benefits and costs.

In the following four chapters, I present data from three separate studies investigating the robustness of several scalable psychological interventions. The next three chapters (3-5) describe the methods and primary results of these studies. Chapter 6 focuses on local norms about growth mindset and the way they influence students' responses to interventions.

# Chapter 3

# Scalably helping at-risk high school students: Brief social-psychological interventions as a systemic solution to academic underperformance

## 3.1 Introduction

In recent years, rigorous field-experiments have demonstrated that brief social-psychological interventions can change how students think about learning and effect lasting gains in academic performance (see Chapter 1). The success of these experiments suggests the potential for social-psychological interventions to contribute to efforts to remedy persistent problems in education (Yeager & Walton, 2011). However, these initial demonstrations were relatively small in scale, and they provided expert researchers with more direct control over implementation than would be feasible on a large scale (see Chapter 2). Because of this, it is uncertain whether the effects of these initial

field experiments are robust and generalizable: Would these interventions work under realistic administration circumstances? Would they work across many education contexts?

These questions are important for practice and for research. Practically, interventions are far more useful if they work broadly and consistently (Bryk, 2009). Scientifically, findings are more helpful for building theory if they are robust and replicable (Pashler & Wagenmakers, 2012; Simmons, Nelson, & Simonsohn, 2011; Spies et al., 2013). They are also more useful for understanding people broadly if they are generalizable (see Yeager, Krosnick, et al., 2011).

In the present research, we sought to determine whether two distinct social psychological interventions can robustly affect academic achievement in a variety of educational contexts. To do this, we created internet-based versions of two distinct social psychological interventions and tested them in a randomized field-experiment targeting a diverse sample of high school students ($N = 1594$ attending 13 different schools). One intervention was designed to convey a growth-mindset, to teach students that intelligence is malleable and grows with challenge and effort (J. Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003). The other intervention was a sense-of-purpose intervention that aims to help students find meaningful, beyond-the-self reasons to learn and work hard in school (Damon et al., 2003; Yeager & Bundick, 2009; Yeager, Henderson, et al., 2013). The hypothesized mechanisms

behind both of these interventions are described in detail in Chapter 1.

Our use of internet-based administration was intended to reduce the logistical complexity of this field research and to lessen the time commitment required of schools and researchers. In this way, the use of internet administration permits significantly larger and more heterogeneous samples. It also ensures that effective treatments can be scaled up to more students at high fidelity and with little additional cost. This novel approach enabled us to shine light on previously unstudied questions.

First, how commonplace are the psychological processes targeted by social psychological interventions? Are they common in and harmful of students' outcomes in diverse school settings? Much laboratory research in psychology with small, select samples assumes that the results speak to people's lives more broadly (see Yeager, Krosnick, et al., 2011). But are interventions that address key psychological processes in education relevant in diverse school settings? If they are, it would suggest that social-psychological interventions may constitute an important element of education reform.

Second, can we change complex social psychological processes like students' beliefs about schooling in broad samples? Psychologists have perfected tools to change psychology with relatively homogenous samples in controlled laboratory settings. Do our tools translate to large-scale field settings where it would be prohibitive to "hand-deliver" or tailor a stimulus to large numbers of people?

We hypothesized that these interventions would affect students' academic mindsets (specifically, their beliefs about the malleability of intelligence and their construal of the purpose of school work) and that they would raise achievement among struggling students. Of course, students may struggle for a variety of reasons. For example, they may simply be academically underprepared, or they may face psychological threats that make it more difficult for them to study or perform effectively (e.g., Steele & Aronson, 1995; Taylor & Walton, 2011). But whatever the specific source of their difficulties, it is especially important for students in challenging circumstances to keep themselves motivated in order to succeed, and research shows that students with adaptive academic mindsets are more motivated and resilient (Farrington et al., 2012; Lesgold & Welch-Ross, 2012).

Thus, we anticipated that these mindset interventions would be most immediately beneficial to students with a history of poor performance—those who are more likely to experience school as challenging. Such effects would be consistent with past research, in which social psychological interventions have proven most effective at raising achievement among students at-risk of poor performance, including students with a history of poor performance and negatively stereotyped ethnic minority students (Cohen et al., 2009; Walton & Cohen, 2011). To examine this in the present research, we test whether intervention effects are most evident at helping students in challenging academic situations. So we examine intervention effects in challenging

core academic courses, among previously low-performing students, and on markers of poor performance, like course failure rates.

## 3.2 Methods

### Participating Schools and Students

Thirteen schools signed up to participate and agreed to provide participating students' academic records. These schools were located in the Southwestern and Eastern United States. Eight were public schools, four were charter schools, and one was a private school. They varied widely in socioeconomic characteristics; in five schools, 0-10% of students received free or reduced lunch because of their low household income; in two schools the figure was 11-50%; and in the remaining six schools, more than half of students were from low-income households.

Across the 13 schools, 1650 students participated in the study. Analyses focus on the 1594 students for whom both pre and post-intervention semester grades were available. Analyses of academic outcomes included all students who were assigned to condition, regardless of whether they completed session 2 (28% of students did not). That is, we used a conservative, "intent-to-treat" analysis. This included 525 Latino, 277 Asian, 371 White, 174 Black, and 247 other/mixed ethnicity students in grades 9-12. Individual students' socioeconomic characteristics (SES) were not collected;

however, school-level SES was available via the free and reduced lunch rate for all public and charter schools (12 out of 13 in the sample).

## Procedure

The study consisted of two 45-minute sessions spaced approximately two weeks apart ($M = 13$ days). Both sessions were administered in each school's computer lab during the spring semester, between January and May 2012. Teachers were instructed to introduce the activities as a part of an ongoing Stanford University study about why and how students learn. Teachers were explicitly instructed not to present the activities as an intervention. Upon signing into the study website, each student was randomly assigned to a control condition or to one of three intervention conditions — a growth-mindset intervention, a sense-of-purpose intervention, or a combined intervention condition. The first session started with the pre-study survey and moved on to the growth-mindset intervention or its control condition. The second session started with the sense-of-purpose intervention or its control condition; it ended with the post-study survey. The content of each intervention is described below.

*Growth-mindset intervention.* In the growth-mindset intervention, students read an article describing the brain's ability to restructure itself as a consequence of effortful practice. The article focused on the implications of these neuroscientific findings for students' potential to become more intelligent through study and practice. Figure

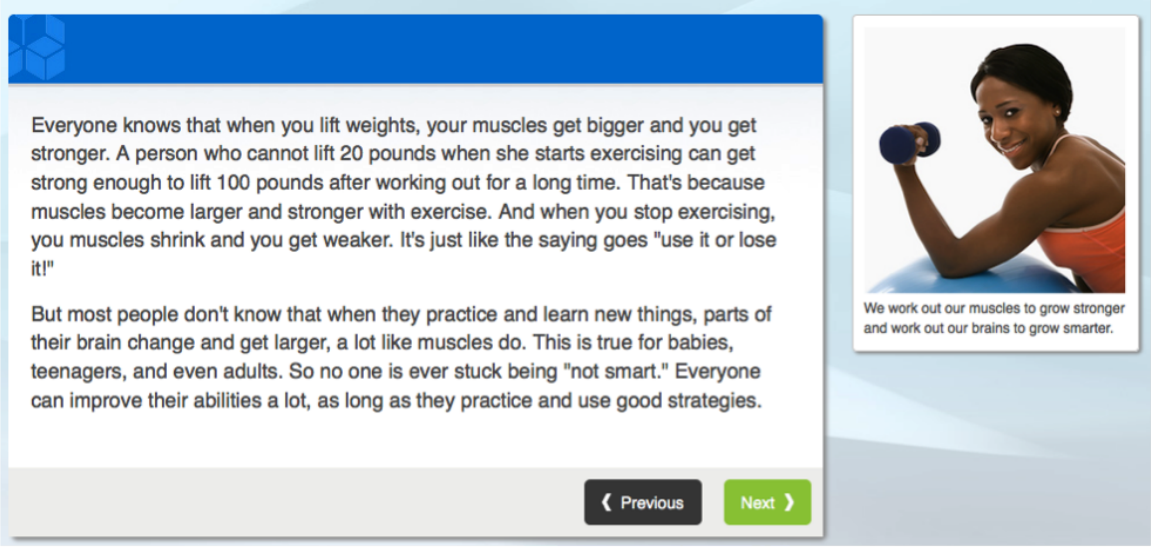



Figure 3: *Intervention Screenshot.* The growth mindset intervention uses the metaphor that the brain is like a muscle.

3 shows a screenshot of this intervention. This message was reinforced through several writing exercises. In one, students summarized the scientific findings in their own words. In the second, they read about a hypothetical student who was becoming discouraged and starting to think of himself as "not smart enough" to do well in school. The writing exercise asked participant students to advise this student based on what they had read. In the control condition, students read and completed similar materials. However, these materials focused on functional localization, not neural plasticity. They thus lacked the key psychological message that intelligence is malleable.

*Sense-of-purpose intervention.* The sense-of-purpose intervention was designed to

help students understand their schoolwork in the context of meaningful, beyond-the-self goals (see Chapter 1). To start students thinking about meaningful, beyond-the-self goals, the intervention first asked students how they wish the world could be a better place. It then went on to explain that many students work hard in school so that they can grow up to be empowered individuals who can make the world a better place, e.g., to "make a positive impact on the world," or "make their families proud or be a good example for other people." Students were then asked to think about their own goals and how learning and working hard in school could help them achieve these goals. In the control condition, students completed either of two similarly formatted web modules. One asked students to describe how their lives were different in high school than before high school. The other focused on economic self-interest as a reason to work hard in school.

## Psychological Measures

Growth Mindset. We measured students' beliefs about the malleability of intelligence using two questions from the theories of intelligence scale used by Blackwell and colleagues (2007). The two items were, "You can learn new things, but you can't really change your basic intelligence." and "You have a certain amount of intelligence and you really can't do much to change it." Only two items were used to save time during questionnaire administration. These two items formed a reliable composite,

$\alpha = .84$.

Construal of Schoolwork. We assessed students' interpretation of a variety of mundane academic tasks, e.g., taking notes in math class, revising an essay, solving math equations, using the Construal of Schoolwork task (Yeager, Henderson, et al., 2013). This 8-item measure is designed to assess whether students view schoolwork at a mechanical, low level-of-construal or at a more abstract and meaningful, high level-of-construal. For example, the measure identifies a particular academic task, e.g., "Doing your math homework," and then asks students to, "Choose the description that more naturally comes to your mind." The two options include, "Typing numbers into a calculator and writing formulas" (low level-of-construal) or, "Building your problem-solving skills" (high level-of-construal). It formed a reliable composite, $\alpha = .72$.

## Measures of Academic Performance

Schools provided participating students' transcripts. We calculated each student's fall- and spring-semester grade point average (GPA) in core academic courses (i.e., math, English, science, and social studies; respectively, pre- and post-intervention GPA). We focused on core academic courses because these courses are generally considered the most crucial to students' success and because they are the most challenging

(the average grade across core subjects was 2.47 vs. 3.15 in other subjects). The difficulty of these courses made them most relevant to academic mindset interventions intended to buttress students' motivation and resilience.

## 3.3 Results

### Manipulation Checks

*Growth Mindset.* A linear regression controlling for pre-study mindset showed that students in the growth mindset group endorsed an incremental view of intelligence to a greater extent than students in the control group, $b = .17$, $t(1009) = 2.647$, $p = .01$.

*Construal of Schoolwork.* A linear regression controlling for pre-treatment construal of schoolwork showed that students in the sense-of-purpose group endorsed a higher level of construal of mundane academic tasks than students in the control group, $b = .05$, $t(1002) = 2.451$, $p = .01$.

### Post-Intervention Grade Point Average

To determine whether each treatment influenced post-study GPA and whether it did so to a greater extent for previously poor-performing students, we subjected post-intervention semester GPA to a linear regression including pre-intervention GPA

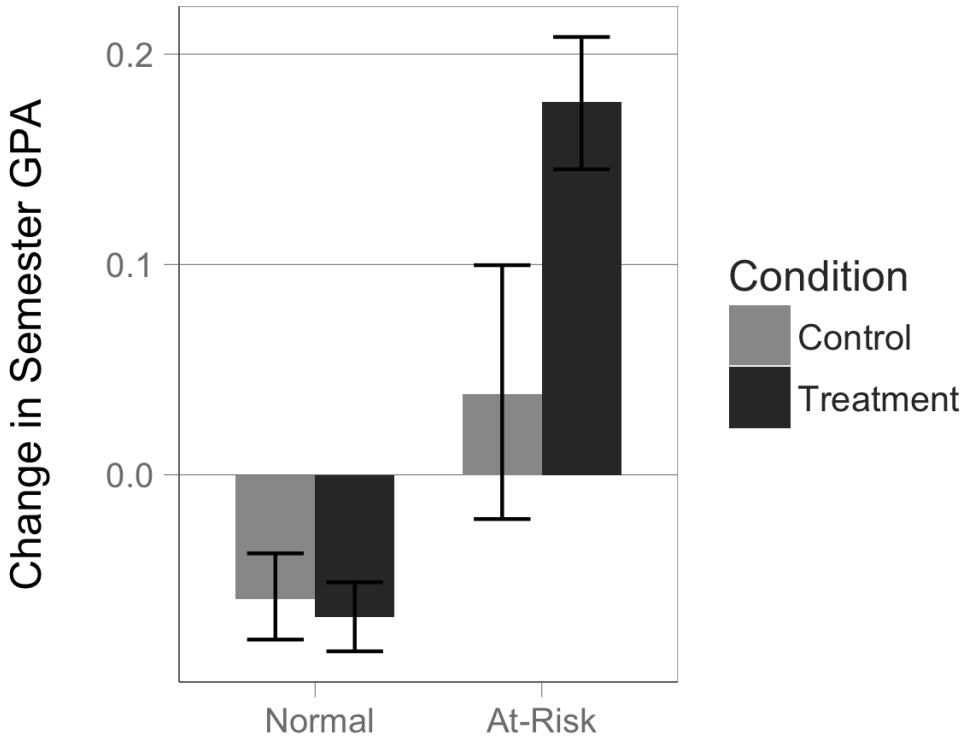



Figure 4: *GPA effects.* Gains in semester GPA were observed for at-risk students.

(standardized), a dummy code for each treatment, and their multiplicative interaction. Because all of the interactions terms were of a similar magnitude and significance ($bs$ -.05 to -.06, $ts$=-1.582 to -1.806), we collapsed the interventions conditions together into a single treatment dummy code (0=control, 1=treatment). We then conducted the same regression with the new treatment dummy code. The predicted interaction was significant, $b = -0.06$, $t(1590) = -2.163$, $p = 0.03$. The simple effect of treatment was significant one standard deviation below mean GPA, $b = .09$, $t(1590) = 2.475$, $p = .01$, but not one above it, $t < 1$. That is, the intervention raised the grades of initially lower-performing students.

We also sought to test the effects of the intervention on the grades of students who would be considered academically at-risk based on widely accepted criteria. To identify such students, we used the indicators developed by the Consortium from Chicago School Research and the National High School Center (see Allensworth & Easton, 2005, 2007; Heppen & Therriault, 2008). Using these criteria, we marked as at-risk all students who earned an overall GPA of 2.0 and below and those who earned an "F" in any core academic course; these students are at a significantly higher risk of dropping out than others. There were 467 such students, 29.3% of the sample.

We conducted another linear regression to determine how much larger the effect was for at-risk students. The interaction included at-risk status (0=normal, 1=at-risk), treatment (0=control, 1=treatment), their interaction, and a covariate

| | |
|---|---|
| Intercept | $-1.03^{***}$ |
| | (0.33) |
| Time | $-0.06$ |
| | (0.16) |
| Treatment | $-0.18$ |
| | (0.19) |
| Time X Treatment | $-0.52^{**}$ |
| | (0.19) |
| AIC | 3730.74 |
| Log Likelihood | $-1858.37$ |
| Num. Obs | 3366 |
| Num. Students | 467 |
| Num. Courses | 199 |
| Num. Schools | 12 |
| Variance Student | 1.61 |
| Variance Course | 1.84 |
| Variance School | 0.64 |
| Variance Residual | |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 1: Regression parameters for the course failure mixed-effects model.

for pre-study GPA. The interaction was significant, $b = .15$, $t(1589) = 2.529$, $p = .01$, demonstrating that the treatments had a significantly greater impact on at-risk students (mean change in GPA control=.04, growth mindset=.18, purpose=.19, combined=.16). See Figure 4.

## Course Failure

We also examined the effect of the intervention on course failures among at-risk students (these students accounted for 96% of failed courses). After the intervention, the rate of course failure was lower in each treatment condition than the control group (control=50%, growth mindset=40%, purpose=40%, combined=41%). To assess statistical significance, we used a logistic mixed-effect model (Bates, Maechler, & Bolker, 2012). As the outcome, we specified failure in each core academic course before and after treatment; as fixed effects, we specified treatment (dummy-coded), time (0 pre-treatment, 1=post-treatment), and their interaction; as random intercepts we specified each student and course. See Table 1 for regression parameters.

The rate at which at-risk students failed courses was significantly reduced as a function of the interaction between time and treatment group, $OR = .59$, $z = -2.743$, $p = .006$. See Figure 5. In the semester before the intervention, the control group and treatment group had very similar rates of course failure (49% versus 48%). In the semester after the intervention, the rate stayed flat in the control group ($< 1\%$ change) but dropped in the treatment group (by 8%). The treatment led the 329 at-risk students in the treatment group to pass an estimated 94 courses more than would be expected based on control-group passing rates[1].

[1]Change in passing in treatment relative to control: 8% * 1177 courses taken in the treatment group = 94 extra courses passed

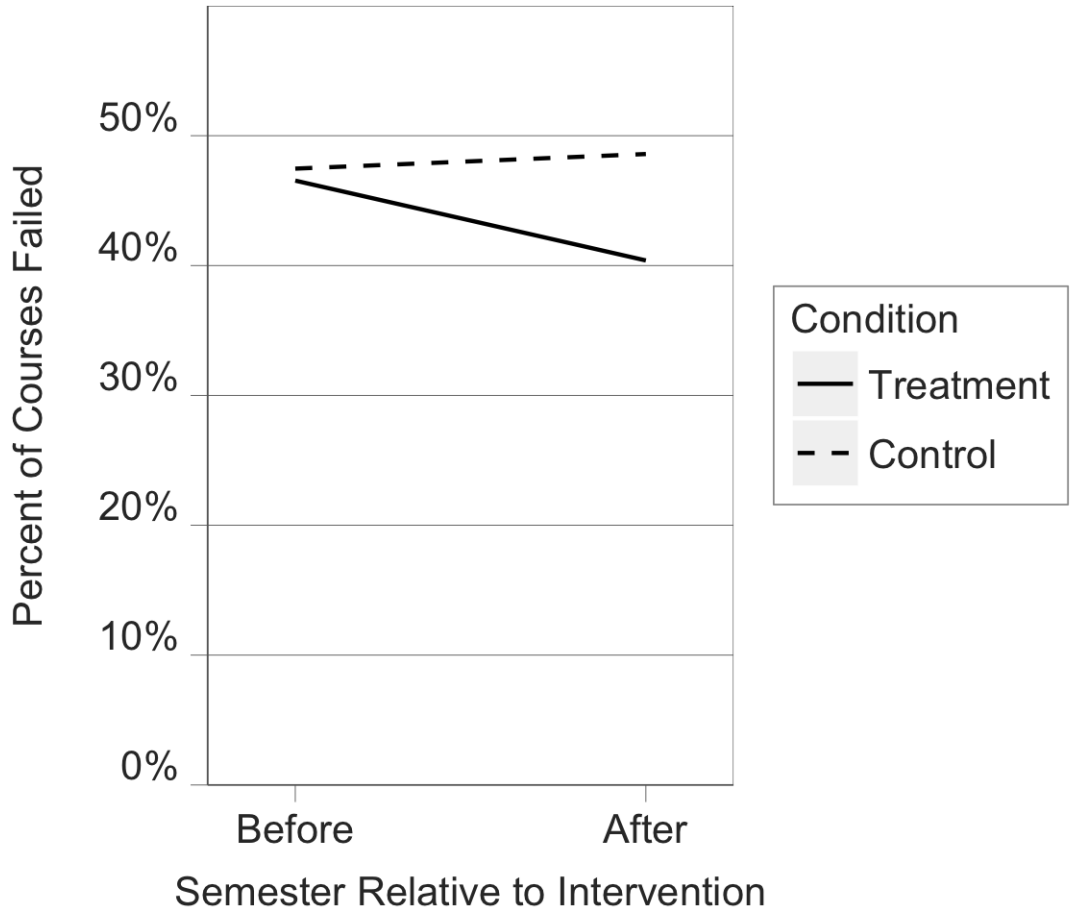



Figure 5: *Course failure effects.* Treatment reduced the rate at which at-risk students failed courses.

## 3.4 Discussion

The two social psychological interventions tested in the reported experiment raised academic achievement among the students for whom school is most challenging: Those considered at-risk by established indicators (Allensworth & Easton, 2005, 2007; Heppen & Therriault, 2008). These students, who had failed at least one core class or had a GPA of 2.0 and below in the semester prior to the study, earned higher grades and passed more courses if they were assigned to the treatment group. Furthermore, these effects were obtained in a diverse sample of students from heterogeneous schools, through a scalable format, in less than 90 minutes of class time, and with minimal supervision from researchers. The methods and sample size of this study represent a significant departure from prior social psychological intervention research, and they speak to the robustness of these effects under realistic administration circumstances.

These interventions prevented an estimated 94 course failures among 329 at-risk high school students. In a sample of 329,090 students, would they prevent 94,000 course failures? In a sample of 4.93 million students—the worst performing third of high school students nationwide—would they prevent a proportional 1.4 million course failures? The present results suggest this possibility, and the methods provide a feasible way to test it. A large, nationally representative study of internet-based psychological interventions could provide key information concerning where these interventions work and reliable estimates for their effect sizes in different populations.

Such data could act as a springboard for wide-scale dissemination in the populations where they would do the most good. They would also provide countless new insights for research concerning which psychological factors inhibit different groups of students from learning—insights that could be used in the formation of new intervention methods.

Although we only observed effects on academic achievement among at-risk students in the present study, it is important to note that "at-risk" is not a static individual difference but a function of an individual's level of challenge in a given context. Based on prior theory and research, we would expect the two tested mindset interventions to raise achievement in a variety of settings and for a variety of students, provided that the students are trying to learn under circumstances that they find challenging. In the case of high school, that group can be operationalized as students who have low prior GPAs. In other environments, that operationalization may be quite different. For example, the higher performing students who adopted a growth mindset as a result of the present intervention may be served well by that mindset when they move onto college and start to take more challenging classes.

Despite their many strengths, internet-based interventions are not magic bullets for school reform. Decreasing course failures by 8% is an extraordinary feat for a 90-minute exercise, but surely far more good could come of finding effective ways to foster adaptive student psychology during the other 74,310 minutes of the average

high school students' academic year (Kolbe, Partridge, & O'Reilly, 2011). How can the lessons of this kind of research help inform what parents and teachers do and how peer interactions are structured? No such techniques are currently established (Farrington et al., 2012), but this is an important question for future research.

Future research should also investigate how these interventions interact. In the present study, the combined treatment condition was no more powerful than either of the single treatments. There are multiple potential explanations for this finding, none of which can be ruled out with the collected data. For example, it could be that struggling students can only focus on a single, motivating reason to try hard in school at a given time. If that were the case, then the second intervention message may provide no marginal benefit because it competes for attention with the first message. This possibility suggests that integrating the two interventions together, as part of a coherent framework, may improve treatment effectiveness by giving students one powerful reason to try hard instead of two separate, less powerful reasons to try hard. Another possibility is that there was not enough time for the combined treatment to exert a larger effect on grades. Perhaps stronger effects will emerge in the combined treatment condition over time, as students encounter more situations in which a sense-of-purpose or a growth mindset is particularly helpful.

Finally, future research should explore how to scale up interventions that address achievement gaps. Value-affirmation and belonging interventions target psychological

processes that affect some students more than others, e.g., stereotype threat (Steele & Aronson, 1995), and they can reduce group-based inequalities in school achievement (e.g., Cohen & Prinstein, 2006; Sherman et al., 2013; Walton & Cohen, 2011). If scaled up, perhaps they could even reduce achievement disparities systemically. However, these interventions are less generalizable by design in so far as they only target certain groups of students. They are also less robust in so far as they are highly sensitive to timing (Cook, Purdie-Vaughns, Garcia, & Cohen, 2011). Future research should consider the proper ways to scale up these interventions so that they reduce achievement disparities without wasting school resources and the time of those students who would be unlikely to benefit from them.

# Chapter 4

# Increasing Resilience in Community College Math

## 4.1 Introduction

This chapter describes our efforts to improve academic outcomes in community college mathematics by adapting and widely testing two scalable, psychological interventions. As described below, community college is an important avenue to social mobility, but students encounter significant difficulties in completing required courses, particularly mathematics. These courses are challenging for community college students, who are often unprepared for difficult coursework and who often come to college with histories of academic failure. To succeed under these difficult circumstances, students must stay resilient. In the research described below, we sought to test whether two interventions designed to target mindsets relevant to academic resilience could improve community college students' academic outcomes.

## The Community College Context

Community colleges are widely recognized as an essential avenue to social advancement (Baum et al., 2010). These colleges serve as entry-points to higher education and higher-paying jobs, especially for low income, first-generation, and minority students, who are more likely to attend them than four-year schools (Bailey & Alfonso, 2005; Horn & Nevill, 2006; Kolesnikova, 2009; Mullin, 2012). Students attending community colleges are the fastest growing population in higher education (Silva & White, 2013), accounting for 56% of the entire U.S. student population in 2012 (American Association of Community College, 2013). However, community colleges have also been criticized for high attrition and failure rates (Edgecombe, 2011; Jenkins & Cho, 2011), especially in mathematics courses (Bailey & Alfonso, 2005).

Many factors contribute to community college students' low graduation rates. They are more likely to have families, to work, or to come from economically disadvantaged backgrounds (Bailey & Alfonso, 2005; Fike & Fike, 2008). However, one of the greatest obstacles is that so many arrive academically unprepared for college-level coursework, especially math (Bailey, Jeong, & Cho, 2010; Grubb, 2010; Levin & Calcagno, 2008; Silva & White, 2013). Many are also psychologically unprepared to stay resilient through the requisite hours of frustration-filled study required to catch up and succeed (Yeager, Bundick, & Johnson, 2012).

Community college students, particularly those needing developmental classes, are

more likely to have a history of academic failure (Adelman, 2006; Lesgold & Welch, 2012; Silva & White, 2013). This may make them especially likely to arrive with low academic expectations and maladaptive beliefs about school that, in turn, can make them less likely to pursue strategies that could help them succeed (Fike & Fike, 2008; Lesgold & Welch, 2012). In two recent surveys of thousands of community college students by the Center for Community College Student Engagement (CCCSE) only 23% said that they often worked with others outside of class (Center for Community College Student Engagement, 2011), and 76% said that they never sought face-to-face tutoring (Center for Community College Student Engagement, 2012). In addition, there is evidence that these students struggle to see the purpose and relevance of their coursework (Yeager et al., 2012).

This problem is especially pronounced in math. Math is often referred to as the gateway subject for degree completion because so many students fail to complete their math requirements and are prevented from earning degrees as a result (Adelman, 2006). At least one college-level math course is required to graduate, and for many majors several college-level math courses are required (Grubb, 2010). Yet, approximately 60% of all incoming community college students are required to complete "developmental" (pre-college) math classes before they even start college-level math (Levin & Calcagno, 2008). Of these, only 31% complete the required classes within three years (Bailey et al., 2010; Grubb, 2010). For those referred to the lowest level

developmental math classes, this number drops to a dismal 16% (Edgecombe, 2011). The prospect of several semesters of non-credit classes — often revisiting material they have failed in the past — causes many students to never enroll at all. Of those who do, over half quit within the first few weeks (Silva & White, 2013). Clearly, many community college students experience math as particularly challenging.

To summarize, math requirements serve as a critical "gateway" to degree completion, and there is evidence that some students fail to complete these requirements in part because they do think about school in maladaptive ways that make it difficult to stay motivated. To the extent that an underlying cause of this educational problem is psychological, it provides an opportunity to create and test psychological solutions.

## Social Psychological Interventions

Academic mindsets have started to receive considerable attention because of their important role in facilitating student motivation and success (Farrington et al., 2012; Yeager & Walton, 2011). A recent report from the National Academy of Sciences drew attention to psychological factors as key determinants of adult learners' persistence (Lesgold & Welch, 2012).

Psychological factors are a particularly attractive target for investigation because they can be changed with brief, targeted interventions (J. Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003; Hulleman & Harackiewicz, 2009;

Jamieson, Mendes, et al., 2010; Walton & Cohen, 2007, 2011; Yeager & Walton, 2011; for a review, see Yeager & Walton, 2011). In the present work, we sought to impart community college students with two academic mindsets that we believed would help them remain resilient in challenging math classes. In doing so, we hoped to 1) identify a partial solution to the social problem of underperformance in community college math and 2) model an example of how the principles of social psychology can be applied at a socially meaningful scale within a particular type of context.

Theorists have observed that successful psychological interventions must be wise to the contexts in which they unfold (Yeager & Walton, 2011). Guided by this principle, we focused on two distinct psychological factors that we believed, based on the evidence discussed below, would be pertinent to community college students' experiences in math and also malleable through brief, scalable interventions.

These two psychological factors were students' growth mindset, i.e., their beliefs about the malleability of intelligence, and their sense-of-purpose about schoolwork, i.e., whether they understand their schoolwork in the context of meaningful beyond-the-self goals. Roughly, these factors map onto the way students answer the questions "Can I do my schoolwork?" and "Should I do my schoolwork?"

### Growth Mindset

One social psychological factor that may affect community college math achievement is students' mindsets about intelligence — students' beliefs about whether intelligence is fixed or malleable. Research demonstrates that mindsets about intelligence have important consequences for how students' respond to challenges (Blackwell et al., 2007; Good et al., 2003; Yeager & Walton, 2011). Students who believe intelligence is fixed — who have a fixed mindset — value appearing smart over learning (Dweck & Leggett, 1988). These students are less likely to ask for help when faced with a challenge out of fear of exposing their inability (Blackwell et al., 2007; Nussbaum & Dweck, 2008). Because they believe that ability in a subject is inherent, they often withdraw effort or even cheat when faced with a challenge, seeing no other remedy for failure (Hong, Chiu, Dweck, Lin, & Wan, 1999). Conversely, students who believe intelligence is malleable — who have a growth mindset — value learning in the classroom and respond to challenges more adaptively (Blackwell et al., 2007; Robins & Pals, 2002). They seek help, try new strategies, and increase their effort when faced with setbacks (Blackwell et al., 2007; Hong et al., 1999). Not surprisingly, students with a growth mindset perform better academically than students with a fixed mindset (Blackwell et al., 2007). These differences are pronounced in challenging subjects, like math, where setbacks are common and perseverance is necessary for success (Blackwell et al., 2007). In several rigorous, small-scale studies, teaching a growth

mindset has led to greater math achievement in junior high students (Blackwell et al., 2007; Good et al., 2003) and higher grade point averages in African American students at an elite university (J. Aronson et al., 2002).

A growth mindset intervention may be particularly beneficial to community college students, who are disproportionately likely to have had histories of academic failure (Adelman, 2006; Lesgold & Welch, 2012). Their personal histories and systematic evidence that they do not seek out help when needed (Center for Community College Student Engagement, 2011, 2012) suggests community college students may be especially likely to internalize doubts about their academic abilities and to worry about whether they are capable of doing well or growing intellectually. For instance, students can come to believe that if they struggle or require help, it means they are too dumb to pass. When this happens, it can make positive engagement in school—like seeking help from peers or faculty — seem like a tacit admission of inability.

### Sense-of-Purpose

Community college students' achievement may also be influenced by their interpretation of why they should learn. A number of studies by Hulleman, Harackiewz, and colleagues (2012; 2008; 2009; 2010) suggests that students are more motivated and perform better when they can understand how their school work is relevant to their

own lives. Other research suggests that particular types of self-relevance may be especially motivating: Those that enable students to interpret schoolwork in reference to a sense of purpose, that is meaningful goal that transcends the self.

Students exhibit greater tenacity when they connect their schoolwork not only to their life but to a sense of purpose that encompasses a commitment to causes that transcend the self (Damon et al., 2003; Deci & Ryan, 2000; Yeager & Bundick, 2009). For instance, when high school students reported that they are motivated by a desire to contribute to society, they adopted more learning goals and showed less of a focus on simply avoiding failure (Lee et al., 2010). Experimental studies have also causally demonstrated that linking one's current learning to a meaningful purpose can lead to greater motivation and deeper learning (Jang, 2008; Vansteenkiste et al., 2004). When students come to understand otherwise menial tasks, like algebra problem sets, in the broader context of their meaningful life goals, they may come to view these actions as more interesting and relevant.

Few community college students appear to spontaneously create the links between their schoolwork and such beyond-the-self goals. In a national survey of community college students conducted by the Carnegie Foundation for the Advancement of Teaching, a full 68% of students said that they did not understand the purpose of the content of their coursework (Yeager, Paunesku, et al., 2013). This large group of students may wonder whether putting effort into abstract academic tasks — like

solving math problems — will ultimately help them fulfill life goals that they find meaningful. This lack of understanding of the purpose of their schoolwork, combined with doubts about their own abilities, can prevent students from taking advantage of even well conceived supports from their college.

## Present Research

In the present study, we examined the effects of a growth mindset intervention and a sense-of-purpose intervention on community college students' math achievement. The interventions were both customized to target community college students by referencing specific concerns that students like them were likely to have. For example, the growth mindset intervention made references to the malleability of even adult brains and the implicit fallacy of statements like, "I am not a math person." Similarly, the sense of purpose intervention normalized and modeled the process of realization about the purpose of schoolwork — the process we hoped students would undergo over the course of the intervention. For example, students read vignettes attributed to community college students which described that they did not always think about their schoolwork in terms of a beyond-the-self purpose but that they came to do so through self-reflection over the course of college.

These interventions also presented certain practical benefits. They required no

customization to particular course content because they focused on elements of learning that transcend the material of a particular course. That is, they did not focus on highly-specific messages, like how the materials in a particular week's algebra lesson are relevant to students' lives. Instead, they tried to change how students made sense of learning as a more general process, e.g., they conveyed the idea that effort and effective strategies beget improvement and that becoming better educated can help one achieve meaningful goals. Those higher-level messages are powerful because they generalize beyond a particular week's or even course's assignments, and they are also more scalable because they do not require extensive customization. This differentiates them from other interventions that involved customizing intervention messages to each week's lesson plan (e.g., Hulleman & Harackiewicz, 2009).

To deliver these interventions, we took advantage of an online delivery mechanism that affords a number of benefits. First, online delivery removed geographic constraints that would have otherwise made it challenging to work with study sites 314 miles and 1,976 miles away from the primary research site. Second, it vastly eased the logistical and class time burden for the participating colleges: Participation required virtually no training for participating colleges and enabled us to deliver these interventions as online homework assignments, significantly decreasing requisite class time. These factors made it far easier to recruit the large sample of students necessary to test these interventions with sufficient statistical power and to provide

meaningful estimates of effect size. Finally, the lower variability associated with a computer-administered intervention meant that these interventions, if effective, could be scaled up to reach many more community college students with high fidelity and at little additional cost.

## 4.2 Methods

### Participants

Two community colleges agreed to participate in this study and provide academic outcomes for participating students. One of these community colleges was located in Southern California, and the other was located in the Midwest. Between these two colleges, 884 students were enrolled in a math course during the semester of the intervention (463 Southern California, 421 Midwest; 531 female, 353 male; mean age = 23). These students were enrolled in 29 different math courses for a total of 908 math-person combinations[1]. Our sample was racially and ethnically diverse, including 413 White, 176 Latino, 78 Asian, 76 Black, and 143 other ethnicity or unknown ethnicity students. Of the 884 participants, 58% (517) were new students and 391 were returning students, for whom prior grades were available.

[1]This number is larger than the total sample because some students were enrolled in multiple math courses.

## Procedure

The study consisted of two 45-minute sessions spaced approximately two and a half weeks apart (mean 18 days). Both sessions were administered either in each school's computer lab ($n = 239$) or as homework ($n = 645$) during the fall semester, between September and November of 2012. When students first signed into the intervention, they were individually randomly assigned to a control condition or to one of three intervention conditions — a growth-mindset intervention, a sense-of-purpose intervention, or a combined intervention condition. Students who were assigned to receive only one treatment always completed that treatment during the first session. In the second session, they completed the control condition associated with the other treatment, e.g., a student assigned to the sense-of-purpose only condition completed the sense-of-purpose treatment during session 1 and the growth mindset control condition during session 2. Students assigned to the combined treatment completed the growth mindset intervention during session 1 and the purpose intervention during session 2, and students assigned to the control condition completed the growth mindset control activity during session 1 and the sense of purpose control activity during session 2.

In the growth mindset intervention, students read an article describing the brain's ability to restructure itself as a consequence of effortful practice. The article focused on the implications of these neuroscientific findings for students' potential to become more intelligent through study and practice. This message was reinforced through

several writing exercises. In one, students summarized the scientific findings in their own words. In the second, they read about a hypothetical student who was becoming discouraged and starting to think of himself as not smart enough to do well in school. The writing exercise asked participant students to advise this target student based on what they had read. In the growth mindset control condition, students read a similarly formatted web module about the brain. However, it focused on functional localization instead of neural plasticity. It was thus devoid of the key psychological message that intelligence is malleable.

The sense-of-purpose intervention was designed to motivate students by helping them to see the value of trying hard in school for their ability to have a personally meaningful life as an adult. Specifically, students were led to focus on personally meaningful, prosocial reasons to try their best in school. The intervention started by asking students how they wish the world could be a better place; it then went on to describe some of the reasons other students report trying hard in school, e.g., to "make a positive impact on the world," "have a career that they enjoy," or "make their families proud or be a good example for other people." Students were then asked to think about their own goals and how learning and trying hard in school could help them achieve those goals. Students not assigned to the sense-of-purpose intervention were assigned to a similarly formatted web module that asked them to describe how their lives are different now that they are in college.

## Measures

Participating students' transcripts were collected from each college, and their fall math grades served as the primary outcome. Students were enrolled in 29 different math courses, and 22 students were simultaneously enrolled in 2 math courses. Most grades (86%) were convertible to a numeric point score, (F = 0, D = 1, C = 2, B = 3, A = 4). However, three P ("passed") grades and 120 W ("withdrew") grades could not be converted to numeric values because they do not get entered into the calculation of students' grade point average.

## Analyses

Analyses were performed using the R environment for statistical computing (R Development Core Team, 2006). Because the data were cross-nested, i.e., some students were in multiple classes and most classes had multiple students, we used the lme4 package to create multi-membership mixed-effect models (Bates et al., 2012). In all analyses, random intercepts were set for each student, course, and the intervention administration format (in-class or homework). To control for past performance, each model included a dummy code for each student's academic history (new student or GPA mean of F, D, C, B, A)[2]. In each analysis we specified random intercepts for

[2]We used a dummy code to control for academic history instead of a numeric GPA score because over half of participants were first semester students for whom grades were unavailable. Multiple imputation would be inappropriate in this case because the prior GPA data of new students are deterministically different from those of returning students.

each student, course, and the program administration format (in-class or homework). We conducted two separate analyses because not all data were readily convertible to numeric grade points.

The grade point analysis used a linear mixed-effects model to predict the numeric grade points earned by the 770 students who received A, B, C, D, or F. The satisfactory completion analysis also included students who withdrew from the math class or took the course on a pass/fail grading basis ($n = 886$). This analysis combined together the grades A, B, C, and P as "satisfactory" and D, F, and W as "unsatisfactory." The distinction between "satisfactory" and unsatisfactory grades is important because, at participating colleges, only satisfactory grades permit a student to receive transfer credit, to count the course as a prerequisite for more advanced courses, or to count the course towards their general education requirements.

## 4.3 Results

*Grade Point Analyses.* To determine whether each treatment influenced post-study math grades, we subjected post-intervention math grades to a linear mixed effects model with each treatment condition and their multiplicative interaction as predictors. The growth mindset and sense-of-purpose terms did not differ, $t < 1$, and the interaction was not significant, $t < 1$. So we collapsed the intervention conditions together into a single treatment dummy code (0=control, 1=treatment). A mixed

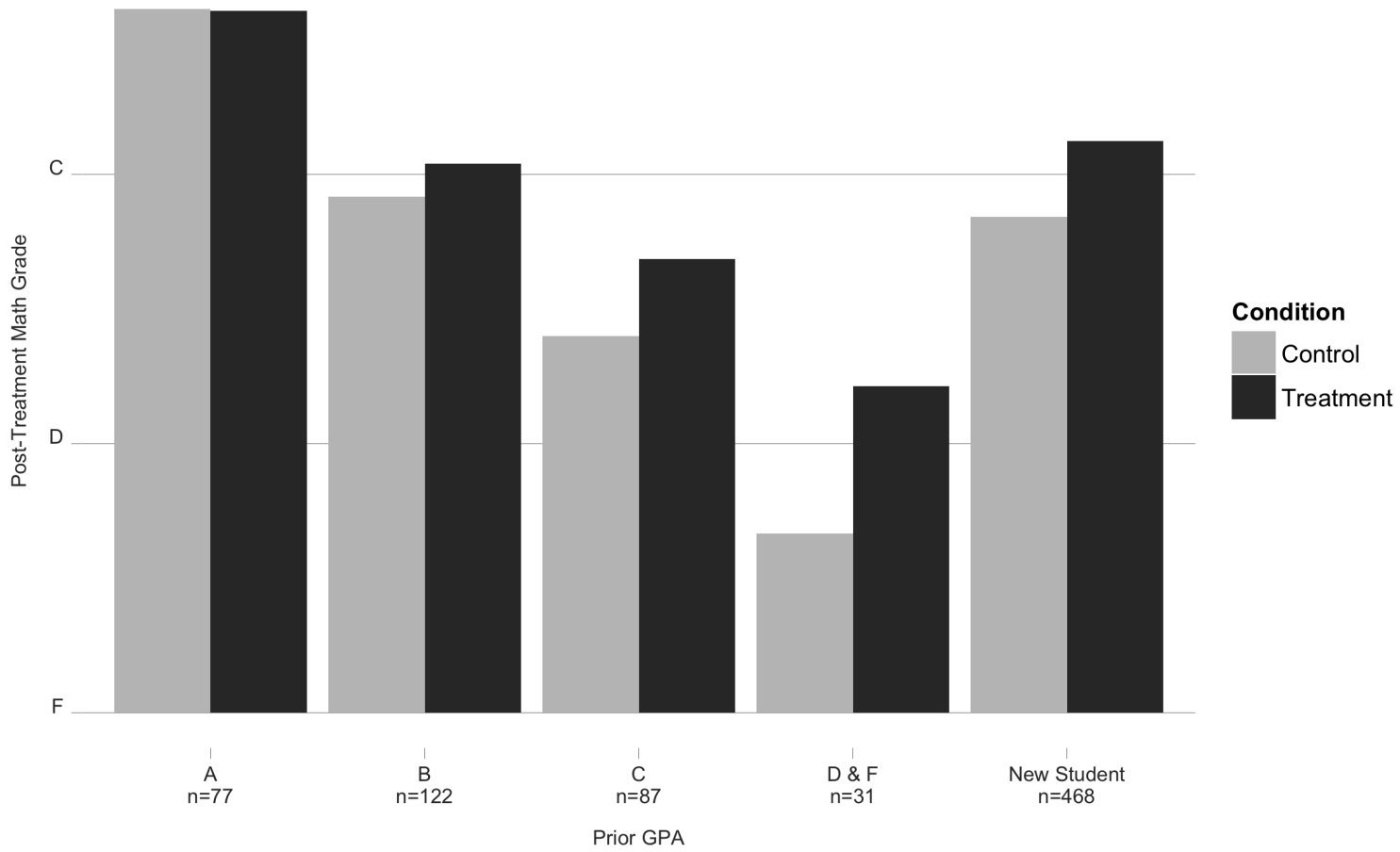



Figure 6: *Effect of treatment by prior GPA.* A mixed effect model revealed that students in the treatment group earned higher grades than control group students.

effect model revealed that students in the treatment group earned higher grades than control group students, $b = .25$, $t = 2.236$, $se = .11$, $p = .03$[3], see Table 2. See Figure 6.

*Satisfactory Completion.* Students in the treatment groups (mindset=56.0% , purpose=60.3%, combined=59.6%) were more likely than students in the control group (52.2%) to earn a satisfactory grade in math. A mixed effect logistic regression model revealed that the effect of treatment on satisfactory grade completion was statistically significant, $OR = 1.46$, logit $= .38$, $z = 2.365$, $p = .02$, see Table 2. See Figure 7. Since students in the treatment group took 680 math courses and their completion rate was 6.4% higher, this intervention cumulatively led to 44 additional students earning satisfactory grades in math.

## 4.4 Discussion

In response to the crisis in community college math programs, we created two contextualized psychological interventions intended to help students see themselves as capable of succeeding academically and to see their coursework as personally meaningful. Relative to the control condition, the treatment conditions increased by 6.4% the rate at which students earned the As, Bs, Cs, and Ps needed to make degree progress. As a result of this treatment at the participating colleges, an estimated

[3]To assess statistical significance, we used a likelihood ratio test, as recommended by Pinheiro and Bates (2000)

| | |
|---|---|
| Intercept | $0.91^{\cdot}$ |
| | (0.55) |
| Treatment | $0.25^{*}$ |
| | (0.11) |
| D Prior-GPA | $-0.30$ |
| | (0.48) |
| C Prior-GPA | 0.24 |
| | (0.37) |
| B Prior-GPA | $0.59^{\cdot}$ |
| | (0.37) |
| A Prior-GPA | $1.24^{*}$ |
| | (0.38) |
| New Student | $0.76^{*}$ |
| | (0.35) |
| AIC | 2702.42 |
| Log Likelihood | $-1340.21$ |
| Num. Observations | 785 |
| Num. Students | 770 |
| Num. Courses | 25 |
| Num. Programs | 2 |
| Variance: Student | 0.86 |
| Variance: Course | 0.07 |
| Variance: Program | 0.33 |
| Variance: Residual | 0.88 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 2: Regression parameters for the course failure mixed-effects model.

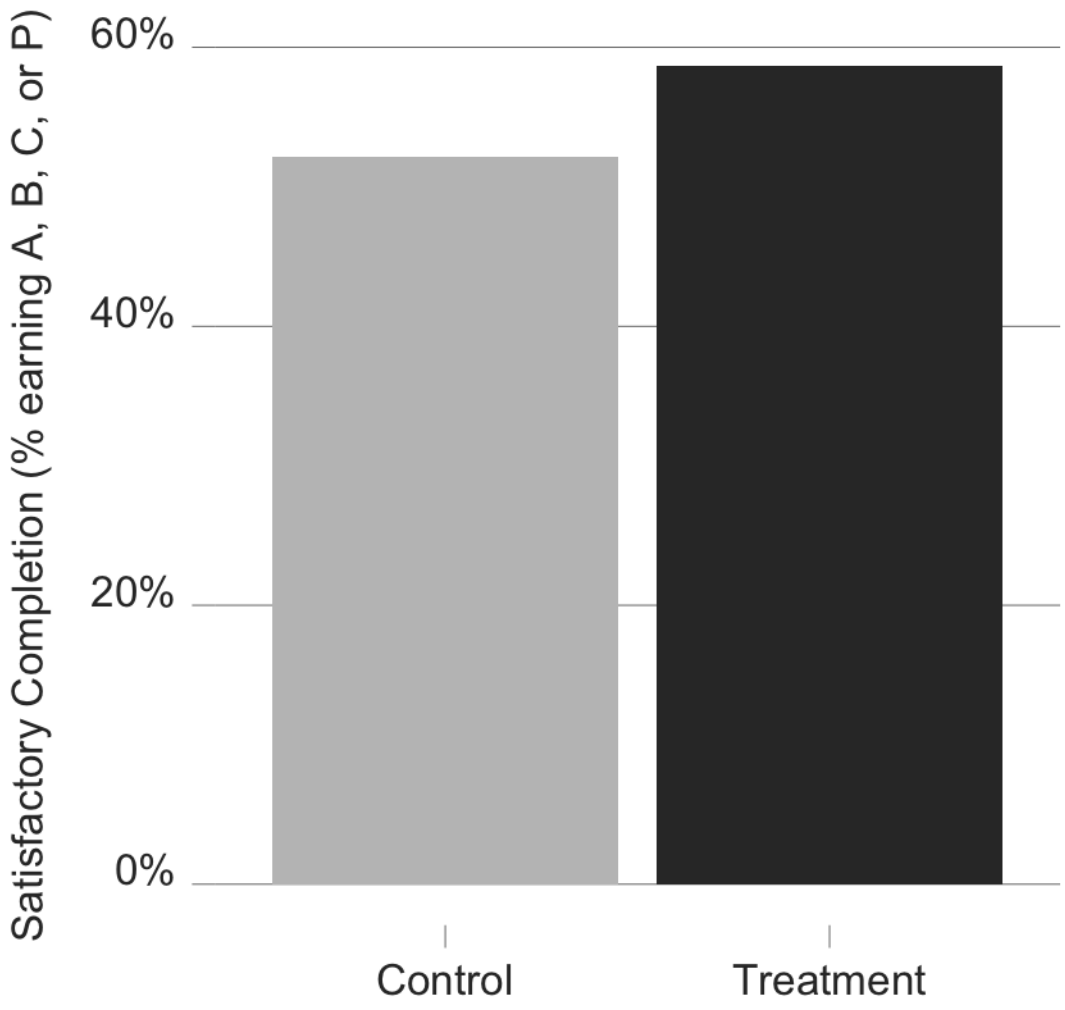



Figure 7: *Class completion by treatment.* A mixed effect logistic regression model revealed that students in the treatment group were more likely to earn a satisfactory grade.

44 additional students earned the grades they needed to progress to more advanced courses or to fulfill their general education requirements. These results are significant for several reasons.

First, the interventions tested in this study made a substantial dent in an important problem in education, and they did so in a low-cost, highly scalable way. The treatment used in the present study reduced the rate of unsatisfactory math course completion by 13.3%. Furthermore, it did so using web-based materials that could be used by virtually unlimited numbers of students at little additional cost; it took less than 90 minutes of student time to complete; it took almost no time away from coursework because it could be completed from home; and it required almost no training of college faculty — faculty merely provided students with a ready-made page of instructions for participation and credited students once participation was complete. This stands in stark contrast to other, much more expensive reform efforts that made no discernible impact on math achievement (Rutschow et al., 2011).

Second, this study is notable because it provides a generalizable model for applying psychology to efficiently solve real-world problems. Although the materials used in this study were contextualized for community college, the approach we utilized could be applied in a variety of contexts. We first identified a problem (community college math course completion) that was likely to — at least in part — have a psychological root. We identified specific psychological factors that appeared to contribute to the

problem. We drew on previous successful interventions to learn how to affect these types of factors, e.g., by using "saying is believing" exercises (J. Aronson et al., 2002; Walton & Cohen, 2007; Wilson & Linville, 1982), and we leveraged information technology to make the process scalable and easy for researchers and participants.

This study opens the door to several important lines of inquiry for future research. One concerns the processes by which these two interventions affect students' achievement. It is unclear, for example, why the combined intervention condition did not lead to larger effects than either intervention alone. The same subadditive effect was observed with a large sample of high school students, and several possible explanations for it are discussed in Chapter 3. In order to investigate why these effects may be subadditive, future work could collect psychological process measures, like daily diaries, and try to discern whether students exposed to multiple treatments only assimilate one intervention message or whether they assimilate both and recruit them to deal with different types of challenges. We avoided such process measures in the present investigation because they would have made it more challenging to recruit a sufficiently large and diverse sample — colleges are hesitant to surrender even 90 minutes of course time, and asking for a larger commitment for the purposes of process outcomes would have dissuaded many faculty from participating. However, colleges may become more willing to surrender class time to research as its utility for their students becomes more apparent.

A related line of inquiry could focus on identifying those students or groups of students who respond especially well or especially poorly to each of these interventions. At the simplest level, students may benefit less from mindset interventions that are redundant with their existing mindsets. For example, a student who believes intelligence is malleable should presumably benefit less from a growth mindset intervention than a student who believes intelligence is fixed.

More complicated treatment response patterns should be investigated as well. For example, consider two students with fixed mindsets: one generally disidentifies with school; the other wants to do well but becomes frustrated by the difficulty he has in math class. Because of his fixed mindset, the second student interprets his difficulties as a sign that he lacks the natural ability to do well in math and starts to give up. If both students were exposed to a growth mindset intervention, we may expect the motivated student to benefit from it more. For him, the intervention would clear a primary obstacle to success; once this obstacle were removed, his existing high motivation could drive his success. For the academically disidentified student, on the other hand, many other psychological obstacles would presumably remain even if his fixed mindset were overturned.

As researchers' understanding of such factors improves, so too will their ability to predict who will benefit from which message and customize messages to students for greatest effect. In the process, these interventions will improve in effectiveness and,

with them, our understanding of what constellations of psychological factors spur students to succeed or hold them back.

# Chapter 5

# Wise Encouragement: Promoting learning one sentence at a time

## 5.1 Introduction

Students who view intelligence as malleable exhibit a variety of adaptive academic attitudes and behaviors. Compared to students who view intelligence as fixed, students who have a growth mindset — those who believe they can grow their intelligence through effort and by learning new strategies — are more likely endorse mastery goals (Blackwell et al., 2007), select difficult rather than easy tasks (Mueller & Dweck, 1998), try to learn from their mistakes (Mangels et al., 2006), and perform better in challenging courses (Blackwell et al., 2007; Grant & Dweck, 2003, see also Chapter 4).

A number of studies have investigated how students come to have a growth mindset. Research on naturalistic praise and lab experiments both suggest that process

praise encourages a motivational framework consistent with a growth mindset (Gunderson et al., 2013; Mueller & Dweck, 1998). Researchers have also conducted targeted interventions designed to instill a growth mindset. These were usually focused, interactive activities lasting from 30 minutes to several hours (e.g., Blackwell et al., 2007; Good et al., 2003). For example, they have included 8 weeks of lectures, 30 minutes of reading and writing (Chapters 3-4), and integration into the content of a multi-week tutoring program (J. Aronson et al., 2002).

In the present study, we wanted to test whether a much more minimal approach, encouragement presented one sentence at a time, could also bring about adaptive learning behaviors. Brief words of encouragement map readily onto the myriad interactions teachers, tutors, and parents have with students in efforts to raise their motivation (Gunderson et al., 2013; Henderlong & Lepper, 2002). Although encouragement is ubiquitous, research has not yet explored whether and how it can instill a growth mindset.

In the present experiment, we tested whether students try harder and learn more (math) when exposed to messages of encouragement that convey a growth mindset, e.g., "When you learn a new kind of math problem, you grow your math brain!" and "Mistakes help you learn. Think hard to learn from them". Importantly, we contrasted these growth mindset messages against face-valid exhortations to "try harder" that students often hear from parents and teachers, e.g., "Set your goals

high, and don't stop until you get there!", "This might be a tough problem, but we know you can do it". These messages served as a key control condition in the present study because, like the growth mindset messages, they encourage students to invest more effort into learning math. However, they lack the key ingredient present in growth mindset messages: The idea that trying harder makes one smarter. In this way, only growth mindset messages transform the meaning of effort from something that helps one accomplish a particular task into something that fundamentally builds one's capabilities.

We performed this experiment in a large online learning environment because it provided an easy way to test the effects of different types of encouragement in a large, diverse population naturalistically engaged in learning. The use of an online learning environment was also timely considering that online learning is growing quickly and that students who learn online may be especially likely to benefit from additional motivational support.

## Online Learning

Each year online enrollment in degree-granting postsecondary institutions has increased by 17.5% — a pattern that has held strong for 10 years with only minimal signs of slowing (Allen & Seaman, 2013). By comparison, general enrollment has only seen average increases of 2.5%. In 2002, 1.6 million students were enrolled in

at least one online class, but by 2011 this number jumped to over 6.7 millions students. Today, 31% of all students attending college or university are taking at least one online class, and 69% of higher education institutions now say online-learning is a critical part of their long-term strategy (Allen & Seaman, 2013). The rise of the internet as a delivery tool for educational content provides unprecedented access to high quality educational materials to countless students. It also presents educators with new challenges for motivating students and researchers with new opportunities to study learning on a larger scale, with more granular outcomes (Romero & Ventura, 2007).

Students who learn online are more likely to become disengaged and drop out (Hiltz, 1997; Moore & Kearsley, 2005). One reason for this may be an inverse selection pressure. Online learning environments generally have fewer barriers to entry than face-to-face learning environments (Dutton, Dutton, & Perry, 2002), and these lower barriers may attract students who are less committed or less capable than those who are willing to attend school in person. In some cases, online educational content may specifically attract or even be intended for those who are struggling to learn a particular concept and require additional assistance (Khan, 2011).

Online learners also tend to receive less social and structural support than students learning in traditional settings (Aragon & Johnson, 2008; Park & Choi, 2009). They are more likely to be older, to work full time, and to have family or personal situations

competing for their time (Aragon & Johnson, 2008; Park & Choi, 2009). These factors combine to put online learners at greater risk of disengaging. Therefore, it is critical that online educational content be designed to support these students in becoming more committed and empowered to persist.

The shift towards online classes within higher education is not the only way online learning is changing the education landscape. Within the k-12 education system, blended learning — the incorporation of internet-based educational content into face-to-face curricula — is on the rise as well (Watson, Murin, Vashaw, Gemin, & Rapp, 2010). In a report issued by the Institute of Educational Science, a 2009 survey revealed that 97% of teachers reported having at least one computer in the classroom and 93% of these had internet access daily. The average ratio of students to computers in 2009 was 3:1, down from 7:1 just eight years earlier (Snyder & Dillow, 2011). Nationally, approximately 50% of all districts either have a blended learning program or are planning to implement one in the near future (Watson et al., 2010).

## Present Research

In the present research, we partnered with Khan Academy (khanacademy.org) to test the effects of growth mindset encouragement in an online environment. Khan Academy is one of the largest, free web-based educational content providers (Roush,

2012). Over 4200 k-12 instructional videos, primarily in math and science, are available on their website along with exercises and tools for students to learn and teachers to monitor students' skill mastery. Every month, it serves over 6 million unique users from 216 countries. Educators understand their value and Khan Academy videos are now used to enhance curriculums in over 30,000 classrooms worldwide. The number of users engaging in Khan Academy activities enabled us to test the robustness of these effects with an unprecedented sample size.

## 5.2 Methods

All of the math videos on Khan Academy are accompanied by practice problems intended to help students learn and practice the associated concepts. These practice problems are divided by topic (e.g., multiplication and division, fractions, decimals, percent) and further subdivided into specific exercises (e.g., "Fraction Word Problems," "Ordering Fractions," "Multiplying Fractions"). By default, students stay within a particular exercise until they are deemed proficient, that is until they have a greater than 94% probability of answering the next problem accurately. Notably, users can skip around to any exercise they choose regardless of their progress within a given exercise.

## Procedures

We modified all fractions exercises on Khan Academy to randomly present users with one of five sets of header messages immediately above each math problem (see Figure 8). Once assigned to a header condition, users were randomly exposed to one of the within-condition header messages on all subsequent fractions exercises.

The conditions included a *no-header* control group in which users did not see a header; this is the default on Khan Academy. There were also two control statements groups: *standard encouragement*, e.g., "Some of these problems are hard. Just do your best," and *science statements*, e.g., "Did you know: An elephant brains weighs 7/2 as much as a human brain." Finally, there were two versions of the *growth mindset* headers. One version was simply a growth mindset encouragement message, e.g., "Remember, the more you practice the smarter you become!" and "If you make a mistake, it's an opportunity to get smarter!" The other version of growth mindset headers also included a hyperlink to another page on which students could read more about the malleability of the brain.

The present sample includes all data collected from 265,082 participants who were randomized to condition within the first 37 days of the experiment. Their demographic characteristics are not known because Khan Academy does not collect such data.

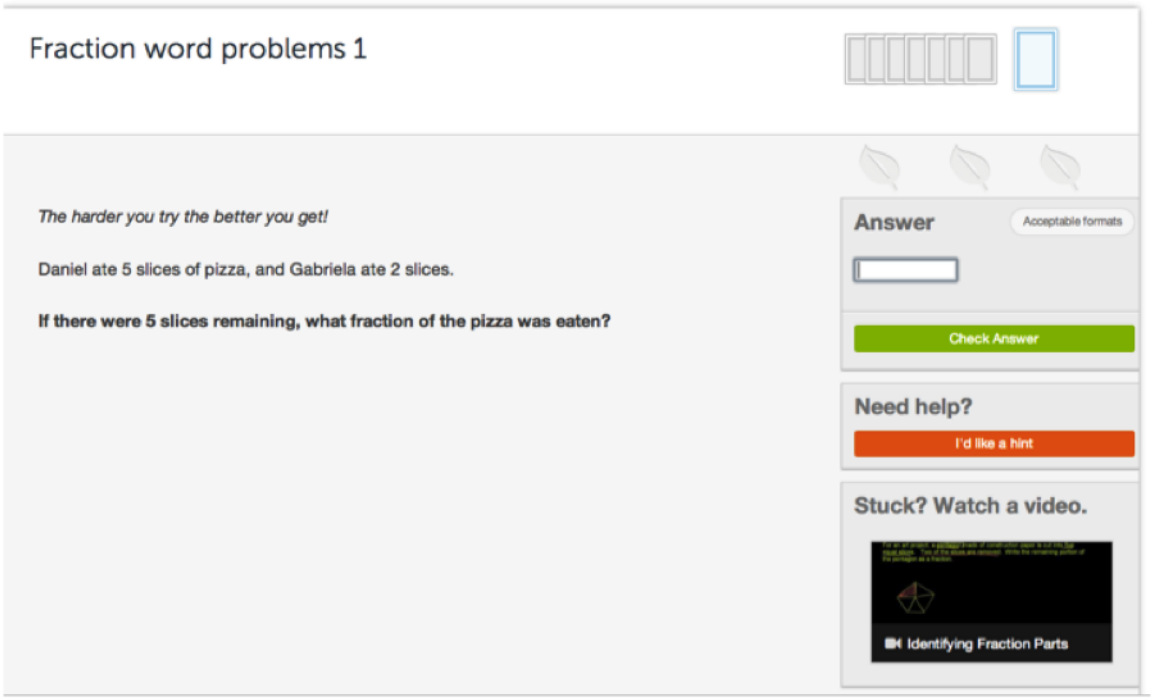



Figure 8: *Khan Academy Treatment Manipulation.* The manipulation consisted of condition-specific headers (it italics) above each math problem in the fractions exercises.

## Measure

Khan Academy deems students to be proficient at a particular concept when they correctly answer enough problems that there is an above 94% probability that they will answer the next problem correctly. We calculated the total number of proficiencies students earned after assignment to an experimental condition.

### Analysis

Analyses focused on the number of proficiencies participants earned after assignment to condition. We used a negative binomial regression model to assess treatment effects on proficiencies. A negative binomial model was used because proficiencies are a count outcome but overdispersed relative to a Poisson distribution, standard deviation/mean = 2.57 (Cameron & Trivedi, 1998). The two growth mindset conditions did not significantly differ from one another; therefore, they were collapsed together for reporting simplicity.

## 5.3 Results

The results of the negative binomial regression model are displayed in Table 3. Relative to the no headers group, the growth mindset encouragement group earned proficiencies at rate that was 2.9% higher, $z = 3.212$, $p < .01$. Neither of the control statement types influenced the rate at which students earned proficiencies, $z < 1$.

## 5.4 Discussion

Students exposed to growth mindset encouragement while learning fractions mastered math concepts at a higher rate. Importantly, the standard encouragement control group showed no gains in learning even though this is precisely the type of

Table 3: Negative binomial regression model predicting proficiencies

| | Negative binomial | |
|---|---|---|
| Parameter | Estimate | $z$-Value |
| Intercept | 1.708 | $359.259^{***}$ |
| Standard Encouragement | 0.000 | 0.001 |
| Science Statements | 0.013 | 1.074 |
| Mindset Statements | 0.029 | $3.212^{**}$ |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

encouragement students often hear from teachers and parents or see on motivational posters. This standard encouragement condition communicated that students should try hard and that they could master the fractions concepts with enough work. However, this condition did nothing to change the meaning of effort for students. By contrast, students in the growth mindset condition were exposed to an interpretation of effort as a way to growing one's intelligence. For these students, trying harder became not merely a way to answer more math problems correctly. It became a way to make themselves smarter.

Consistent with other experiments, we found that a growth mindset intervention could improve learning outcomes (e.g., J. Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003). However, the format, setting, and practical implications of this study are quite different. First, this was the first study to focus on brief words of encouragement, as opposed to targeted praise (e.g., Mueller & Dweck, 1998) or more extensive interventions (e.g., J. Aronson et al., 2002; Blackwell et al., 2007). Second,

this was the first study to investigate the effect of a growth mindset intervention in a massive online educational context — a context in which more and more students are learning every year (Allen & Seaman, 2013). Third, the sample was over 1000 times larger than the largest published growth mindset intervention to date (n = 138 in Good et al., 2003).

That this brief intervention was effective even with a massive, heterogenous sample suggests that its effect is extremely robust, and it carries practical implications. Although the absolute size of the gain (3%) may appear small, it is practically significant considering that it could be easily implemented across a range of learning contexts. Parents, teachers, and tutors frequently provide words of encouragement to spur their students to learn (Gunderson et al., 2013; Henderlong & Lepper, 2002). This study experimentally identified messages of encouragement that affect learning and compared them directly against other messages that have face-validity but show no effects on learning. Furthermore, this line of research is only beginning.

In future research, we will expand on this work in several ways. We will create and test new types of encouragement to determine the specific characteristics that predict positive effects. For example, perhaps messages centering on the value of mistakes or those that evoke the idea of the brain growing are especially motivating.

Future work should also experimentally manipulate timing. In the present study,

students were exposed to one of several encouragement messages above every problem. However, students may "tune out" these messages after repeated exposure. So a natural extension of this work would be to test whether encouragement is more effective on an intermittent schedule or when timed to coincide with challenging problems. For example, we could expose students to an intervention when they hit a streak of problems on which they are struggling.

Finally, future research should also assess disaggregated outcome data. In the present study, we summed the effects of the intervention across all proficiencies completed after exposure. However, we also collected problem-level data that could provide additional insight about the process. For example, we could examine whether the growth mindset intervention is, as theory would predict, especially effective at improving outcomes on challenging problems. We could also examine whether the effects of exposure to the mindset messages decrease, increase, or remain steady over time.

# Chapter 6

# School Norms and the Growth Mindset: Influences on achievement and response to intervention

## 6.1 Introduction

Research has shown that individuals' implicit theories about intelligence — whether they view intelligence as static or malleable — influence their academic goals and achievement (Blackwell et al., 2007; Good et al., 2003; Yeager & Walton, 2011). To date, this research has focused only on individuals' *own* beliefs about intelligence. However, a rich body of psychological research suggests that the perceived beliefs and behaviors of *others* can also influence individuals' attitudes and behaviors (Cialdini & Goldstein, 2004; Miller, Brickman, & Bolen, 1975; Newcomb, 1943; Prentice &

Miller, 1993). Based on this this research, we hypothesized that local growth mindset norms, i.e., students' *peers'* beliefs about the malleability of intelligence, would (1) influence students' achievement and (2) affect students' acceptance of messages promoting a growth mindset.

## Theories of Intelligence

Individuals think about intelligence in different ways. Some view intelligence as a fixed, unchangeable quality — they have a "fixed mindset" — while others view it as a malleable quality that can be grown through effort — a "growth mindset" (Dweck, 1999). These mindsets carry important implications for the way individuals conceive of their intellectual potential and for the role they ascribe to effort in realizing that potential. Students with a growth mindset tend to view effort at challenging academic tasks as a way to grow one's abilities and intelligence. In contrast, those with a fixed mindset view effort as a sign that one lacks natural ability, e.g., "if you were smart, you wouldn't have to try hard." These different perspectives lead to divergent academic behaviors and outcomes.

Compared to those who have a fixed mindset, students with a growth mindset respond more adaptively in the face of various challenges (Molden & Dweck, 2006): They pay more attention when their mistakes are corrected (Mangels et al., 2006); they seek out information that could help them improve their performance (Nussbaum

& Dweck, 2005); and they attribute poor test performances to factors within their control — like a lack of preparation as opposed to a lack of innate ability (Blackwell et al., 2007). Over time, students who have a growth mindset, or endorse the types of learning goals associated with a growth mindset, tend to earn higher grades (Blackwell et al., 2007; Romero, Master, Paunesku, Dweck, & Gross, 2013).

Importantly, research has shown that these mindsets are themselves malleable, and it has identified some the ways they take shape. For example, one longitudinal study showed that the way children are praised early in life affects their motivational frameworks several years later (Gunderson et al., 2013). Whereas praising a child for her intelligence suggests that some innate trait is the key to success, praising her for the process she used to accomplish a task suggests that effort and strategies are the keys to success. Consistent with this interpretation, children who were praised for process rather than intelligence at the ages of 1-3 were more likely to have a motivational framework aligned with a growth mindset 5 years later—they were more interested in engaging in challenging tasks than easy tasks (Gunderson et al., 2013). The same pattern of findings has been observed in short-term, experimental studies with older children (Mueller & Dweck, 1998).

These findings suggest that individuals' mindsets and associated approaches to challenge are affected by the statements and behaviors of those around them. In this way, certain school environments may be more conducive to the adoption of a

growth mindset than others. For example, a school's culture may celebrate effort and learning as ways to reach one's potential; or it may derogate those who have to work hard and celebrate those who seem to achieve effortlessly. Classic social psychological research suggests that norms like these are likely to influence students' attitudes and behaviors.

Social psychological research has extensively studied the influence of social norms on students' attitudes and behaviors. It has found that students — like people in general — tend to adopt the real or perceived norms of those they interact with. Students are more likely to drink alcohol heavily if they believe their schoolmates enjoy doing so (Prentice & Miller, 1993); they are more likely to become politically liberal if they attend a predominantly liberal college (Newcomb, 1943); they are more likely to conform to the eating behaviors of their sorority sisters (Crandall, 1988); and they are more likely to clean up after themselves or to study math if they think their classmates do so (Miller et al., 1975). In this way, real or perceived school norms can influence individual students' behaviors and become self-reinforcing.

Given social norms' ability to influence students' attitudes and behaviors, it is important to consider how normative attitudes about intelligence at a school may influence individual students' mindsets and associated behaviors. Although students and teachers may not explicitly discuss their beliefs about intelligence, they may implicitly communicate them through other means. For example, teachers may praise

children for putting forth significant effort or, alternately, for accomplishing tasks seemingly effortlessly. Students may congratulate each other for successfully struggling through difficult lessons, or they may call each other "dumb" whenever they have to work hard to understand a topic. In these ways and others, a variety of common classroom situations could instill or reinforce a particular mindset.

Based on existing research, one may expect schools with growth mindset norms to foster higher achievement than schools with fixed mindset norms. However, it is less clear whether such local norms would influence students' achievement over and above students' individual mindsets. On one hand, students may benefit from a school's growth mindset culture only to the extent that they personally come to view intelligence as malleable. On the other hand, even students who individually endorse a fixed view of intelligence may nonetheless benefit from the strong emphasis on learning and effort that would be expected in a growth mindset culture. For example, fixed mindset students at growth mindset schools may seek out challenging problems and work hard to solve them because that is what everyone else in their class does and because that is what their teachers expect. Consistent with this example, we anticipated that a growth mindset culture would benefit even students who individually endorse a fixed mindset, and we sought to test this hypothesis in the present study.

We also tested a related question: Are students in a growth mindset culture more accepting of an intervention designed to instill a growth mindset? A number of field

experiments have shown that targeted psychological interventions can lead students to adopt a growth mindset (Blackwell et al., 2007; Good et al., 2003; Paunesku, Yeager, Romero, & Walton, 2012; Romero et al., 2011). However, this research has not systematically investigated whether local norms influence the ability of growth mindset interventions to instill a growth mindset. Classic research suggests that local norms should have such an influence. For example, attitude change interventions were found to be more effective when their messages were perceived as consistent with social norms (Lewin, 1958). That is, individuals appear to be more likely to accept and act on an intervention message if the message is consistent with their peers' beliefs or behaviors. On that basis of this research, we hypothesized that a growth mindset intervention would be more effective at changing individual students' mindsets in environments in which such mindsets are already normative.

## Present Research

In the present study, we explored the relationship between theories of intelligence and achievement at the student and school levels. First, we tested whether school mindset norms influence students' achievement over and above individual students' mindsets. We predicted that growth mindset norms would produce a culture more supportive of learning and effort and that this culture would positively influence students' achievement even after controlling for students' individual mindsets. Second,

we tested whether school mindset norms would affect students' responses to a growth mindset intervention. Here, we predicted that students in schools with a growth mindset culture would be more accepting of an intervention designed to impart a growth mindset.

## 6.2 Methods

### Overview

The present investigation used data from a larger study run by the Project for Education Research That Scales (PERTS). The primary focus of the larger study was to test the efficacy of two social psychological interventions with a large, heterogeneous population: a growth mindset intervention and a sense-of-purpose intervention (see Chapter 3). Because the present investigation is focused specifically on growth mindsets in relation to school context, analyses focus on the data relevant to students' mindsets and to the effects of the growth mindset intervention (see Analytic Strategy section).

## Procedures

The study consisted of two internet modules that each took 30-45 minutes to complete. Students completed these modules in their own schools' computer labs between January and April of the spring academic semester (this semester covers January to May or June, depending on the school). Students were supervised by their own teachers, who were blind to condition assignment. The study was described as research about how students learn and what motivates them; it was not described as an intervention because of concerns that this could undermine the effectiveness of the intervention (cf. Sherman, Cohen, et al., 2009). Schools were instructed to space the two modules two weeks apart, mean spacing between the two modules was 13 days.

Module one started with a baseline survey that assessed students' pre-intervention mindsets (described in Measures section). After the survey, students were randomized to a growth mindset intervention or to a superficially similar control condition. In the growth mindset intervention, students were provided with scientific evidence pointing to the plasticity of the brain and of intelligence. The intervention also discussed the implications of neural plasticity for learning and effort. The control condition materials described functional localization in the brain, e.g., that visual information is processed in the occipital lobe at the back of the head.

The beginning of the second module contained a different psychological intervention that targeted students' sense-of-purpose (Damon et al., 2003; Yeager & Bundick,

2009). For simplicity, data from this second intervention were not used in the present investigation; the Analytic Strategy section below describes which data were used in each analysis.

## Measures

*Student Mindset.* Pre-study and post-study surveys were collected at the beginning and end of the first and second internet modules, respectively. These questionnaires assessed students' implicit theories of intelligence, and they were composed of two questions on 6-point Likert scales. These items formed a reliable composite ($\alpha = .75$, mean = 3.75, n = 2850).

*School Mindset.* To get an estimate of the local growth mindset norm of each school, we calculated the mean student mindset score within each school (mean = 3.75, n = 23 schools).

*Semester Academic Grades.* Official academic records were collected from a subset of schools (n=13). Academic records included one outcome per student per course per semester. Grades were rated on a $0 - 4.3$ scale, 0=F to 4.3=A+. An a priori decision was made to only include academic courses in analyses, i.e., math, science, social science, English, and foreign languages. Non-academic courses, e.g., student government, physical education, and art, were excluded because these courses were generally less challenging (mean GPA in academic courses = 2.16 and non-academic

courses 3.25) and therefore less relevant to students' mindsets about intelligence.

## Analytic Strategy

The data used in the present study were collected in the course of a larger study focused on maximizing the number of students participating in social psychological interventions. Great efforts were made to make participation easy for schools. Because of this open recruitment policy, the sample is considerably larger and more heterogeneous than those in previous social psychological education interventions. For the same reason, however, commitment from schools was variable and degrees of freedom vary between analyses.

Ten of twenty-three schools did not provide academic transcripts from participating students (this accounted for 33.7% of students who completed the initial survey and intervention). The schools that did provide transcripts had more students participate (mean n=167) than those that did not (mean n = 73). Post-study mindsets are unavailable for 42% of students who did not complete module 2; there was no meaningful difference in attrition between conditions, $\chi^2(1) = 1.8287$, $p = .18$.

We conducted all analyses using linear mixed-effects models with the lme4 package (Bates et al., 2012) in the R environment for statistical computing (R Development Core Team, 2006). To assess statistical significance, we used a likelihood ratio test, as recommended by Pinheiro and Bates (2000). In each analysis, we used the most

inclusive sample possible, but results did not meaningfully differ if more restrictive samples were used.

## 6.3 Results

### School Mindset Norms and Academic Achievement

To assess the association between school mindset norms and students' academic achievement, we specified a linear mixed-effect model predicting students' average course grade in the semester prior to the intervention. Random intercepts were specified for each school (n = 13), and fixed effects were specified for students' baseline mindsets (standardized), school mindset norms (standardized), and their multiplicative interaction. The analysis included 1580 students who completed the baseline survey and whose schools provided academic transcripts.

Individual and, critically, school mindsets both predicted higher GPAs. Individual students' pre-intervention mindsets were positively associated with their GPAs, $b = .056$, $se = .020$, $t = 2.878$, and school mindsets were also positively associated with students' GPAs, $b = .54$, $se = .17$, $t = 3.141$. Removing individual student or school mindset from the model significantly decreased the goodness of fit, as indicated by likelihood ratio tests — effect of individual mindset $\chi^2(1) = 7.9984$, $p = .005$; effect of school mindset $\chi^2(1) = 8.0412$, $p = .005$. The interaction was not significant,

$b = .015$, $se = .019$, $t = .766$. See Figure 9.

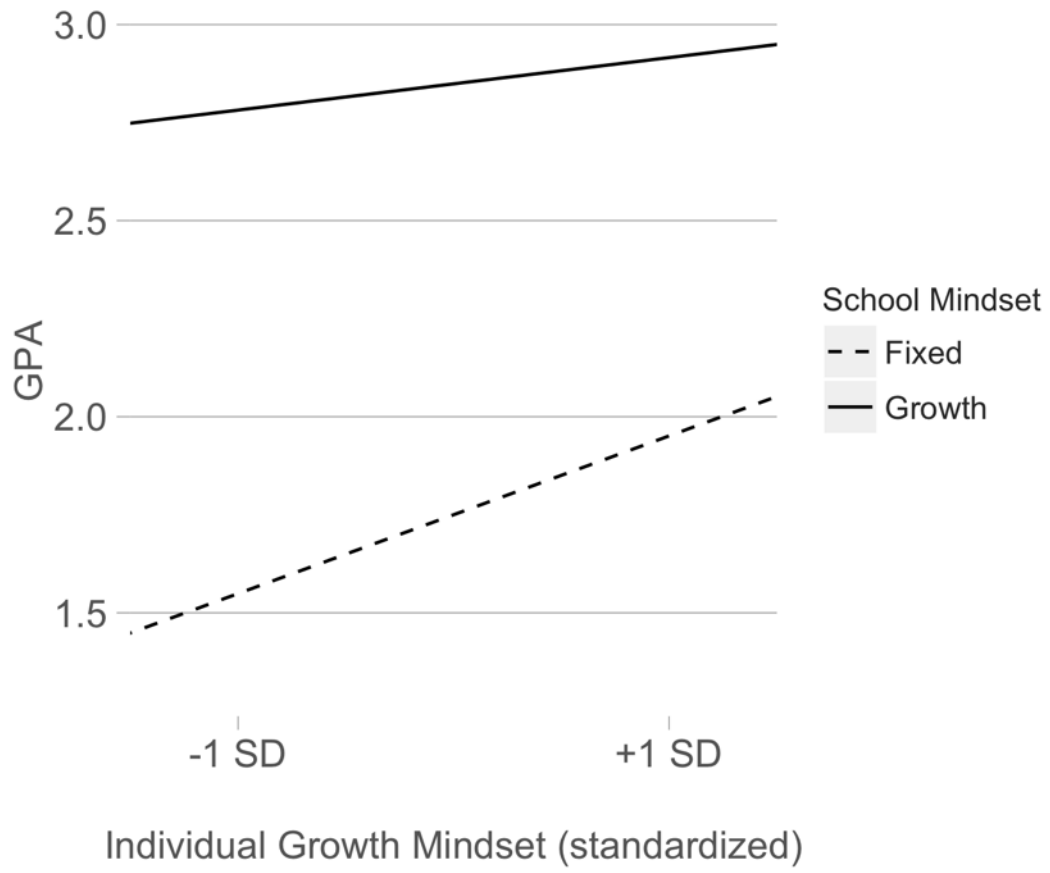
## Intervention-Driven Changes in Mindset

First, we sought to confirm that the growth mindset intervention affected students' post-intervention mindsets. To do so, we specified a linear mixed effect model predicting each student's post-study mindset (standardized for ease of interpretation). We specified school as a random intercept ($n = 23$). As fixed effects, we specified an intervention dummy code and each student's baseline mindset. The growth mindset intervention made students significantly more likely to endorse a growth mindset, $\beta = .17$, $se = .037$, $t = 4.45$, $\chi^2(1) = 19.759$, $p < .001$. The analysis included 1655 students who completed the pre- and post-intervention surveys.

We then tested the effects of individual- and school-level mindsets on the impact of the intervention. We subjected post-intervention mindset to a linear mixed-effect model with student-level mindset (standardized), school-level mindset (standardized), growth mindset treatment, and all 2-way and 3-way interactions. School was included as a random intercept. The three-way interaction was dropped because it was not significant $\beta = -.03$, $se = .038$, $t = -0.840$, but all two-way interactions were significant (see Table 4). On the individual level, students with relatively fixed mindsets experienced more change in mindset as the result of the intervention, and, most crucially, students in schools with growth mindset norms experienced greater changes in

Figure 9: GPA by individual and school mindset

Baseline GPA was predicted by individual students' mindsets and also by the average school mindset. School mindsets are split into lines at the median in this chart for visualization purposes.



| | |
|---|---|
| Intercept | −0.13 |
| | (0.03) |
| Treatment | 0.18 |
| | (0.04) |
| School Mindset | 0.01 |
| | (0.03) |
| Individual Mindset | 0.69 |
| | (0.03) |
| Treatment x School Mindset | 0.11 |
| | (0.04) |
| Treatment x Individual Mindset | −0.10 |
| | (0.04) |
| School Mindset X Individual Mindset | 0.05 |
| | (0.02) |
| AIC | 3788.42 |
| Log Likelihood | −1885.21 |
| Num. Students | 1655 |
| Num. Schools | 23 |
| Variance School | 0.01 |
| Variance Residual | 0.56 |

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 4: Regression parameters for the change in mindset mixed-effects model.

mindset as a result of the intervention. See Figure 10.

## 6.4 Discussion

This was the first study to investigate the role of school mindset norms, and it did so with an unusually large and diverse sample. Its findings show that school mindset norms influence students' academic achievement and students' responses to growth

Figure 10: Intervention effect by initial mindsets

Students with relatively fixed mindsets (left side of each cell) experienced more change in mindset as a result of the growth mindset intervention. Students in schools with stronger growth mindset norms (right cell) experienced greater changes in mindset as the result of the intervention compared to students in schools with relatively fixed mindset norms (left cell).

mindset interventions.

The first set of analyses revealed that students tend to earn higher grades both when they personally view intelligence as malleable and also, independently, when their peers view intelligence as malleable. The first of these relationships replicates, at a much larger scale ($n = 2850$ vs. $n = 373$), the association between individual students' mindsets and achievement (Blackwell et al., 2007). The latter effect is novel and invites more research about precisely how school norms around growth mindset influence individual students' achievement over and above their own mindsets. One possible mediator for this relationship is the degree to which mastery goals and associated behaviors are normative in these environments (Dweck & Leggett, 1988; Elliot & McGregor, 2001).

It is easy to imagine how a growth mindset culture that celebrates hard work could lead students to adopt mastery goals (Dweck & Elliott, 1983; Dweck & Leggett, 1988; Elliot & Harackiewicz, 1996), which have been associated with a growth mindset about intelligence (Blackwell et al., 2007), while a fixed mindset culture that disparages effort and mistakes could lead students to avoid mistakes and endorse performance-avoidance goals, which have been linked to higher academic anxiety, more self-handicapping, and lower achievement (Kazemi & Stipek, 2001; Linnenbrink, 2005; Midgley & Urdan, 2001).

In fact, research on "structural achievement goals" has already demonstrated that

classroom norms around the goals of schoolwork affect individual students (Turner et al., 2002; Wolters, 2004). In one study, Turner and colleagues (2002) assessed classroom structural goals by asking students to agree or disagree with statements like: "My teacher thinks mistakes are okay in math as long as we are learning." These perceptions were then aggregated to the classroom level. Students mastery-oriented classrooms were less likely to self-handicap or avoid novel techniques when doing schoolwork, and they were more likely to ask their teachers for help when they needed it. Put another way, students in mastery-oriented classrooms were more likely to exhibit a learning approach associated with a growth mindset.

The present study did not include achievement goal measures; so it could not directly test the prediction that growth mindset schools tend to be mastery-oriented while fixed mindset schools are performance-oriented. This is a direction for future research, and it could inform efforts to design classroom practices that foster a growth mindset. For example, perhaps teachers could by taught to reinforce the notion that mistakes are valuable for learning or to reassure students that they have the ability to improve and to reach a high standard (Cohen et al., 1999; Yeager et al., in press).

The second set of reported analyses focused on how school mindset norms influence students' responses to growth mindset interventions. Students at schools with growth mindset norms experienced a greater change in mindset as a result of the intervention than students at schools with fixed mindset norms. This result was consistent with

expectations from the social norms literature that it should be easier to affect an attitude towards rather than away from a group norm (Lewin, 1958). It implies that interventions aiming to instill a growth mindset—especially in fixed mindset environments—may do well to try to change the real or perceived norms around mindset. For example, following in the footsteps of classic field interventions, they may carefully guide groups of students to publicly endorse a growth mindset and, in this way, change the apparent group norm (cf. Lewin, 1958).

Importantly, that the intervention influenced *mindsets* to a greater degree in growth mindset schools should not be taken as evidence that the intervention also influenced students' *achievement* to a greater degree in these schools. A secondary analysis revealed that was not the case. Schools with fixed mindset norms had significantly more at-risk students than growth mindset schools, and, regardless of context, at-risk students experienced the greatest academic gains as a result of the intervention (see Chapter 3). Therefore, the absolute gain in achievement was actually greater at fixed mindset schools because these schools had more students in dire need of intervention. Does this mean that the growth mindset intervention was "redundant" with the norms of the growth mindset schools? Probably not.

Although the intervention had a smaller immediate impact on grades at growth mindset schools, it influenced students' mindsets to a greater degree in these schools, and this could have positive long-term effects. As discussed in Chapter 3, higher

performing students did not show any immediate gains in performance as a result of the intervention. Presumably, a growth mindset was less immediately relevant to these students because they experienced high school as unchallenging. However, what will happen as they start to take more challenging classes later in high school or college? At that point, their mindsets may become relevant and those who adopted a growth mindset early on — whether through local norms or through targeted intervention — may fare better.

# Chapter 7

# Conclusion

The theory and data presented in this dissertation introduce — or, perhaps more accurately, reintroduce — a specific methodological approach to social psychological research: Large-scale field experiments. This approach draws directly on a rich tradition of methodologically rigorous, theory-driven, socially important research in social psychology (e.g., Allport, 1954; E. Aronson et al., 1978; Darley & Latané, 1968; Lewin, 1947; Milgram, 1965; Zimbardo, 1969). Following the advice of Kurt Lewin, who famously said, "If you want to truly understand something, try to change it," this approach tests psychological theory by using its predictions to try to solve social problems. It does so by identifying and carefully modifying psychological tension forces in real-world environments using controlled experiments. This gives researchers the ability to make causal claims about what drives social problems; it also provides key insights concerning how to solve, or at least mitigate, these problems.

Although the field research methodology I describe and use in this dissertation

draws heavily on a long-standing tradition, it also builds on it in an important way. It focuses specifically on large-scale psychological field interventions. This is an important distinction because large-scale demonstrations are in many ways qualitatively different from small-scale demonstrations: They can speak directly to questions of generalizability, robustness, and cost-effectiveness. To borrow clinical terminology, they can answer questions about effectiveness in addition to efficacy (Flay et al., 2005; Rush, 2009; U.S. Food and Drug Administration, 1998). These questions matter for society because they determine whether these interventions — or policies modeled on them — constitute wise investments of public resources (for a discussion, see Chapter 2 and the Economic Impact section below). These questions should therefore also matter to researchers who want their work to be put into practice to broadly benefit society.

In Chapters 3-5, I presented data from three large-scale randomized controlled trials. These trials demonstrate that growth mindset and sense-of-purpose interventions can raise achievement among at-risk students in a variety of contexts, including a demographically diverse set of high schools and colleges and a sample of over a quarter million online learners. These studies provide robust evidence for the effectiveness of these interventions, and they provide an estimate for their effect-size when measured across diverse populations. Even more importantly, they model a highly scalable approach to randomized controlled trials in schools, an approach that could

be used to test these interventions (and enhanced variants thereof) even more broadly and cost-effectively (see Future Directions, below).

Large-scale field research is not only distinguished by its more direct implications for policy; it also fosters the pursuit of novel, otherwise untestable hypotheses. For example, Chapter 6 explores local-norms about growth mindset. It investigates the way these local norms interact with individual students' mindsets to influence their grades and their responses to a growth mindset intervention. This type of investigation shows how the tension systems operating in these schools differ from one another and how the same intervention impacts people differently depending on their social context (Lewin, 1947; Ross & Nisbett, 1991). Such a study would be impossible to conduct without a large sample of students from many different schools, and it suggests directions for future research — research that targets the social context as well as the individual.

## 7.1 Future Directions

The findings presented in this dissertation provide a proof-of-concept that social psychological interventions can effect meaningful change in schools on a large scale. This new methodology opens the door to many lines of research that may have otherwise been deemed unlikely to work or too costly and difficult to pursue.

One such line could focus on classroom practices. Reviews of research on students' mindsets note the promise of social psychological interventions but also the gap between research and actionable advice for teachers (Farrington et al., 2012). Farrington and colleagues (2012) lament that, "Unfortunately, the research does not directly translate into classroom strategies that teachers can use to support positive mindsets in their students." They conclude that, "a central tension arising from the research on academic mindsets revolves around how best to apply the research to improve student outcomes." Our own work and discussions with schools confirm that many educators want to learn how to boost their students' learning and engagement by leveraging psychologically wise practices. This unfilled need presents an opportunity for high-impact research.

Could the lessons learned through existing psychological interventions be translated into actionable advice for teachers? Teacher-mediated interventions could, if effective, be far more powerful than brief, online direct-to-student interventions because teachers spend a great deal more time interacting with students than an online intervention feasibly could. Teachers also, to a large extent, define the parameters of the environment in which students learn. For example, research has documented how teachers' framing of learning goals influences students' academic motivation and behavior (Hiebert & Grouws, 2007; National Council of Teachers of Mathematics, 2000; Turner et al., 2002; Wolters, 2004). It has found that students who agree

with statements like, "My teacher thinks mistakes are okay in math as long as we are learning," are less likely to self-handicap or to avoid novel techniques when doing schoolwork; they are also more likely to ask their teacher for help when they need it (Turner et al., 2002).

In light of this research, there is reason to believe that the way teachers praise students, discuss the meaning of mistakes, and frame the purposes of learning could foster different mindsets, learning goals, and behaviors. Future lines of research could therefore focus on changing the way teachers praise students or talk about mistakes. This work will present certain challenges. For example, there will inevitably be greater variability in the way teachers execute an intervention protocol than the way a computer executes it. However, the benefits could be enormous, and Elliot Aronson's work on the Jigsaw Classroom shows that classroom practices can be infused with psychological wisdom to great effect (E. Aronson et al., 1978).

Another future extension of this research could focus on developing predictive models for where and for whom specific interventions are effective. By leveraging large samples, investigators could test questions about contextual moderators, like those presented in Chapter 6. For example, investigators could test if, consistent with theory-based predictions, negatively stereotyped students benefit from belonging and value-affirmation interventions to a greater extent when they are in the numerical minority in their particular environment (e.g., Inzlicht & Ben-Zeev, 2000; Walton,

Spencer, & Erman, 2013).

They could also use machine learning tools to discover new patterns in large datasets. As the amount of psychometric and academic data collected grows, it may become possible to use data-mining techniques to find natural clusters of students who are psychometrically similar to one another across multiple dimensions (Bishop & Nasrabadi, 2006). These students could potentially be expected to share similar beliefs about school and to respond similarly to particular interventions. For example, maybe there are sizable groups of students who generally like school and have a growth mindset but suffer from debilitating math anxiety. Is it meaningful to consider the specific psychological experience of that subgroup of students, or are they such a small minority that their very existence is more likely to reflect measurement error? If such a subgroup did exist, could a well-timed value-affirmation reduce their math anxiety, boost their confidence, and start a virtuous cycle (Cohen et al., 2006; Martens, Johns, Greenberg, & Schimel, 2006)? Or would it be better to flesh out the implications of their growth mindset for mistakes — to help them embrace mistakes and seek out new strategies when they get stuck? The latter questions would have to be settled with experiments designed by clever researchers, but the clusters themselves could potentially be discovered by clever algorithms. In this way, the ability to automatically find meaningful patterns in data could be generative for theory. Such cluster classifications could also have practical effects if they could be used to tailor

interventions to students' particular concerns.

## 7.2 Economic Impact

What social and economic impacts can be expected from broadly integrating the insights and methods of social psychology into education practice? Such estimates depend on many assumptions. The interventions we tested may not work as well in more representative samples; on the other hand, they will likely work more effectively as we iteratively improve them and as we better learn to target them. For example, the effects of psychological interventions may also grow as we learn how to change classroom practices and norms in ways that reinforce the interventions' messages (see Chapter 6). Despite the many uncertainties, I would be remiss to not put forward some loose economic estimates.

In this analysis, I make a number of informed but untested assumptions to arrive at some "back of the envelope" impact estimates. Because of the many uncertainties involved, I try to err conservatively. I provide point estimates based on observed results and also estimates that assume that the interventions would only be *one third* as effective at scale. I focus on the results of the interventions tested with a diverse sample of high school students (see Chapter 3) because reliable data exist both for the economic impact of high school graduation and for the probability of graduating from high school given different grade point averages.

## The Economic Impact of High School Graduation

Numerous analyses have estimated the economic impact of high school graduation. The U.S. Census Bureau estimates that graduating from high school increases lifetime earnings by $250,000 (Day & Newburger, 2002). This number isolates the difference between dropping out and graduating from high school; i.e., it does not count as high school graduates those who go on to college (those individuals earn far more money). The Alliance for Excellent Education (2003) estimates that the increased spending by these richer graduates stimulates economic growth, increasing gross domestic product (GDP) by approximately $29,500 per student over the course of their career. In the following analysis, l assumed that the interventions reported in Chapter 3 only influenced students' likelihood of graduating from high school, not their likelihood of attending college.

## How Much Do Psychological Interventions Boost Graduation Rates?

To estimate the economic impacts of psychological interventions via their impact on high school graduation grades, I first needed to estimate their impact on high school graduation rates. The data collected by PERTS have not yet tracked students long enough to directly assess this. However, extensive research by the Consortium for Chicago School Research (CCSR) provides good estimates for the effects of high school

GPA on the probability of graduating. I used these data to estimate the increase in graduation rate associated with the implementation of these interventions.

The CCSR report, "What Matters for Staying On-Track and Graduating in Chicago Public High Schools", describes the graduation rate across all Chicago schools for students at each half-point in freshman year GPA (Allensworth & Easton, 2007). For example, a student who finishes their freshman year with a GPA of 1.0 has a 28% chance of graduating from high school, one with a GPA of 1.5 has a 53% chance, and one with a 2.0 GPA has a 72% chance.

In order to calculate the change in the probability of graduating for each freshman student, I used the CCSR half-point GPA cut-offs and created local linear interpolations between each cut-point (see Figure 11). For example, if a student earned a GPA of 1.5, then consistent with the CCSR numbers, I assigned her a 53% chance of graduating. If she earned a GPA of 1.6, I calculated her probability of graduating as one fifth of the way from 1.5 (53% chance of graduating) to 2.0 (72% chance of graduating), i.e., 56.8%.

Using this technique, I calculated each students' likelihood of graduating based on her pre-study GPA and post-study cumulative GPA. I then compared the changes in these rates between the treatment and control group. In the control group, the change in graduation probability was -.5% from fall to spring semester. In the treatment group, it was +.1%. This implies that the intervention increased the average student's

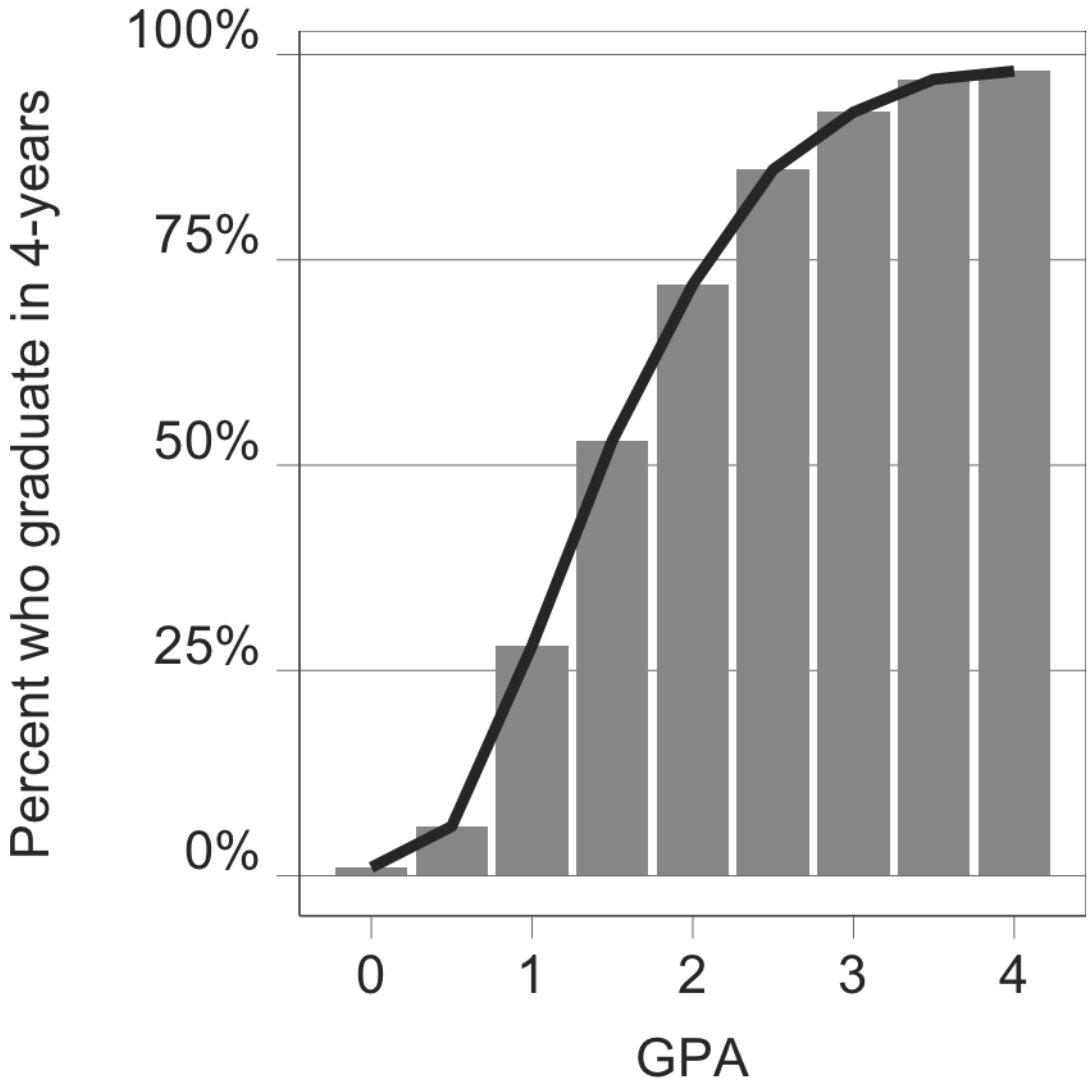



Figure 11: *Graduation Rate by GPA.*Each bar's value is taken from Allensworth and Easton (2007). The connecting lines are local linear interpolations used to convert GPA values into likelihood of graduating.

probability of graduating by .6%. Of course, most students are not in significant danger of dropping out. Those who earn GPAs below 2.0 are at a considerably higher risk, and they comprise a significant portion of the student population (40% in Chicago and 22.5% in the PERTS study population). In this group, the intervention raised predicted graduation rates by a much higher rate: 3.5%.

I should also note that the use of cumulative GPA to calculate the change in graduation probability may understate the size of this effect because students' cumulative grades after the study are "weighed down" by their pre-study GPA. This appears appropriate in this particular case because the intervention was administered during the second semester. However, there is evidence that social psychological interventions can have long-lasting effects (e.g. Cohen et al., 2009; Walton & Cohen, 2011; Wilson & Linville, 1985). So I also calculated the change in graduation rate if the intervention were delivered in the first semester and its effects maintained into the second semester. I did that by calculating the change in graduation likelihood using the post-study semester GPA instead of the cumulative GPA. This recalculation put the change in graduation rate at 1.8% across all students or at 5.9% among high-risk students (those with pre-GPAs of below 2.0).

In summary, I estimated that these psychological interventions, as tested, would lead to a gain of .6% in the graduation rate across a treated group of 854 freshman, i.e., that 5 additional people will graduate from high school as a result of having been

assigned to the treatment group in this study. This implies that U.S. GDP will grow by 5 x 29,500 = $147,000 and that these students will cumulatively earn an additional $250,000 x 5 = $1.25 million over their lifetimes.

These estimates assume a similar distribution of GPAs to that observed in out high school study, and they would depend considerably on the distribution of GPAs at each school. These interventions influenced students with low GPAs most strongly, and these students are also at greatest risk. Therefore, to the extent that the interventions are targeted at this high-risk population or implemented in school districts with many high-risk students, the associated increases in graduation rates would be higher. In the section below, I consider the costs of these interventions and finish by calculating their return on investment in several different scenarios.

## What are the costs of these interventions?

One of the benefits of the online delivery system we developed for this research is that it scales up cost-effectively. There are relatively high initial start-up costs associated with the IT infrastructure; however, these base costs remain virtually flat as more students are added. I assume a fixed base cost of $300,000/yr to run a center that creates and distributes online social psychological interventions; this assumes several managerial positions and IT costs.

There are also variable costs associated with technical support. School staff must

be trained to administer the interventions properly and tech-support is sometimes required. Based on current capacities, I assumed that one full-time technical support specialist (paid $50,000/yr) is required for every 50 new schools. This is a conservative estimate because that is approximately the number of schools PERTS support specialists currently handle, and they have many other research duties that require significant amounts of time. What cost does that imply per student? The average U.S. high school has 752 students and more than 25% are freshman because many students drop out (National Center for Education Statistics, 2001); so I assume an average freshman class of 200 per school. That means that, on average, one $50,000/yr coordinator is needed for 10,000 students — or that the incremental staff cost is $5/student.

I also factored in the cost of class time. According to the National Center for Education Statistics, in 2011, average U.S. spending per elementary and secondary public school student was $10,560 (Dixon, 2013). The average U.S. high school student spends 74,400 minutes in school (Kolbe et al., 2011). If we assume that all of that money is spent directly on educational activities, then each minute of a high school student's time costs $10560 / 74400 minutes = $.14. Assuming that the 90 minutes spent on these interventions directly displace 90 minutes of class time that is worth $.14/student, the cost of the intervention per student is $.14/minute * 90 minutes = $12.6.

To summarize, the formula for calculating total cost as a function of number of students ($n$) is:

Cost ($n$ = number of students) $= 300,000 + (5 + 12.6) * n$

## Return on Investment

Finally, what net gains could be realized with the broad-based implementation of these interventions? Drawing on the figures calculated above, I created several scenarios, each with the number of students ranging from ten thousand to one million. The return on investment is displayed in Figure 12. I calculate the return on investment as follows:

$$\frac{\text{Benefits}}{\text{Costs}} = \frac{\text{Change in Graduation Rate} * 29{,}500 * n}{(5+12.6) * n + 300000}$$

The following scenarios are graphed.

*Baseline*: 0.6% change in graduation rate. This estimate is based directly on observed PERTS data. It assumes interventions are delivered without any targeting to high-risk students.

*1/3rd Baseline*: 0.2% increase in graduation rate. This scenario assumes that the interventions would only be one third as effective when broadly instituted.

*Baseline Persist*: 1.8% increase in graduation rate. This variation on the Baseline Scenario assumes that the intervention is administered in the first semester of freshman year and that its effect persists through the entire year.

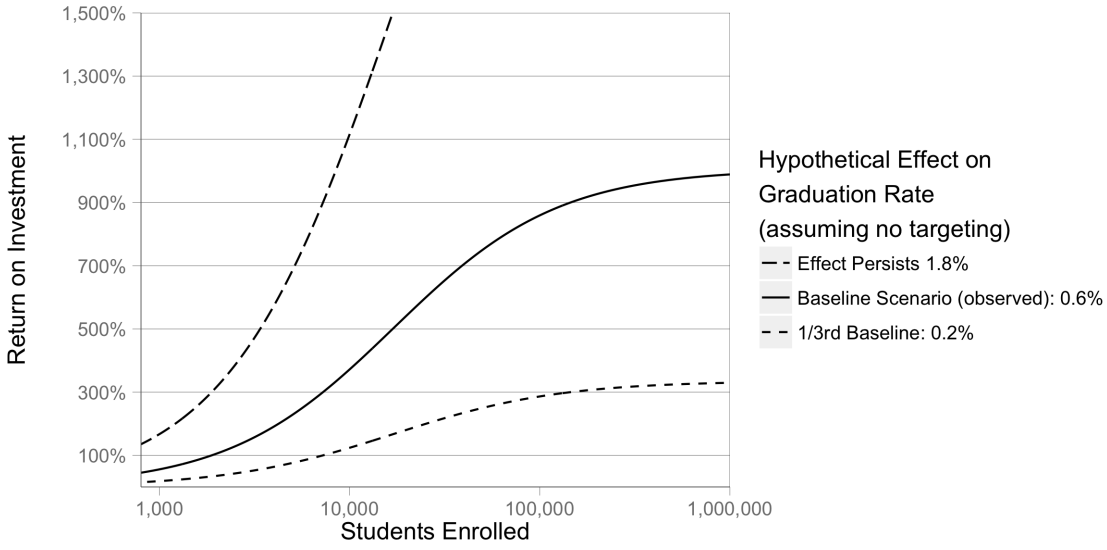



Figure 12: *Return on investment for untargeted interventions.* Return on Investment assuming that interventions are broadly implemented, i.e., that they are not targeted to high-risk students or school districts with higher concentrations of at-risk students.

As Figure 12 shows, the return on investment of these interventions is extremely high, even under a number of conservative assumptions. It also increases substantially as more students participate. For example, if 100,000 students participated, the Baseline Scenario projects that the benefits to the economy would be 9x greater than the costs of implementation. This is noteworthy because the Baseline Scenario makes a number of conservative assumptions.

First, it assumes that there is absolutely no improvement in the effectiveness of these interventions as a result of additional research and smarter targeting. For instance, it assumes no improvement from planned efforts to raise effect-size by enlisting teachers to reinforce messages and no improvement from increasing engagement by

adding audio, video, and greater interactivity. It also rests on a number of conservative assumptions about costs. It factors in distributed indirect costs, i.e., it treats as costs students' time that would not actually have to be remunerated. Furthermore, it uses an upper bound estimate for those costs: It assumes that all of the class time displaced by the intervention is completely wasted, even though students actually spend it engaged in academic activities like reading and writing. This scenario is also conservative because it assumes that no efforts will be made to selectively implement these interventions among higher risk students, despite their markedly higher impact on this population.

Even under this Baseline Scenario, these interventions would equal the return on investment of the widely touted Head Start program, 700-900% (Deming, 2009; Ludwig & Phillips, 2007), if 40,000-145,000 students participated. They would exceed Head Start's return on investment as more students participated.

What if we targeted at-risk students? If we were to target at-risk students on a track to earn GPAs below 2.0, the benefits could be substantially larger than the previous scenarios estimate. Even if we assumed that it would take 5 times as much staff time to coordinate for these students to participate (because they may need to get pulled out of other classes), the reduction in wasted time by other students would push down the costs substantially. Figure 13 shows the return on investment for interventions targeted in this way (incorporating a 5x higher staff cost). In this

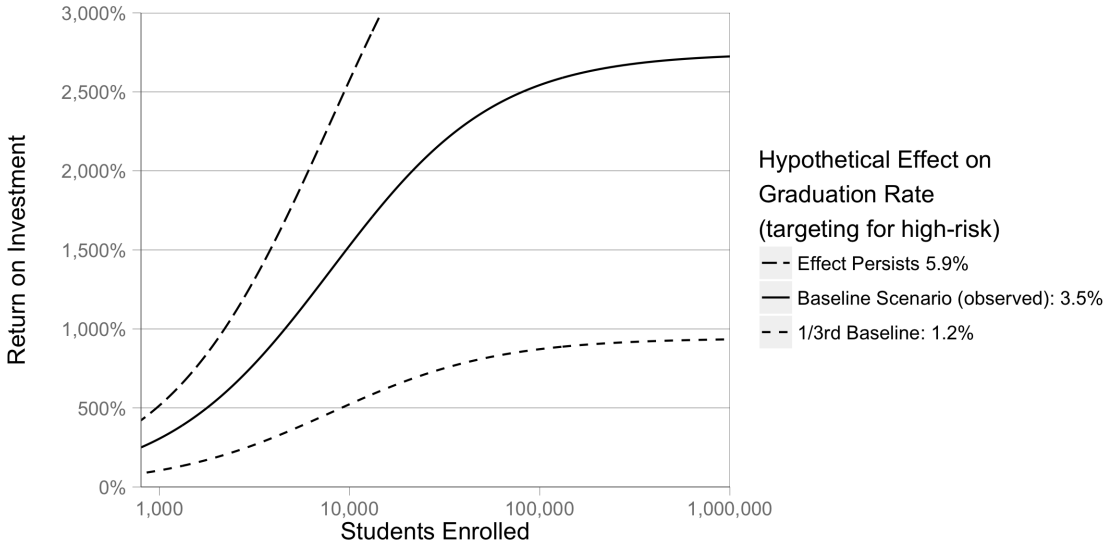



Figure 13: *Return on investment for targeted interventions.* Return on investment if interventions were targeted at high-risk students, those with GPAs below 2.0.

scenario, return on investment exceeds 15x when more than 10,000 high-risk students participate. Handled incorrectly, singling out high-risk students for the intervention could make it appear remedial and undermine its psychological impact. However, there may be ways to mitigate those concerns. Furthermore, some of the benefits of targeting high-risk students could be obtained simply by targeting schools with more high-risk students.

Further research in this area must continue. It is still too early to properly assess the impacts of these interventions or to gauge their full potential, but it is even more difficult to imagine that our team's nascent, bootstrapped effort to scale up social psychological interventions in schools — this "baseline estimate" — represents the

top boundary of this work's potential. More likely, it represents a low estimate. Even so, it is comforting that, should further efforts to scale prove much less effective than our initial demonstrations, the benefits will likely outweigh the costs if enough students participate. For example, an intervention with 1/6th the effect of the initial study would still have greater benefits than costs as long as 25,000 or more students participated.

## 7.3 Conclusions

This is an exciting time for social psychology. The careful field-research approaches pioneered by early social psychologists can now be combined with new technologies to conduct previously unimaginable experiments — experiments that definitively answer novel questions about human behavior and clear the path to important, widely realizable social changes. Today's new tools make social psychology's wisdoms more relevant and powerful than ever. The new era of the large-scale field experiment presents new opportunities to transform society and psychology's role in it. In the face of these opportunities, psychologists must not remain so tacit about their field's hard-won wisdoms (see Markus, 2005). Instead, they should find ways to use these tools wisely to generate new knowledge about human behavior, to learn how to make the world a better place, and to show the world the enormous value of a light, psychologically wise touch.

Of course, social psychology has long been relevant to important real-world domains. During World War II, governments called on social psychologists to devise ways to sell war bonds and conserve food (Lewin, 1947; Merton, 1946). For decades, marketers have drawn heavily on the theoretical insights from the attitude change branch of the psychology (e.g., Hovland, Janis, & Kelley, 1953; Petty & Cacioppo, 1986). Indeed, there are many excellent examples of psychology in action, but the current trend is different in an important way.

Now the cornerstone of psychological methodology — the randomized behavioral experiment itself — is gaining bottom-up popularity. Companies have started to run experiments to optimize user interfaces (e.g., Optimizely.com, Mixpanel.com) and to identify the most productive advertisements (e.g., Google AdWords). Experiments have gained tremendous popularity in these domains because computers have made it so easy to randomize features or ads and monitor their impacts on valued outcomes (e.g., sales). This trend will only grow as computers and their ability to collect data cheaply become more ubiquitous, as analytic techniques improve, and as the gains to be realized from the careful study of people's behaviors become more evident (Lohr, 2013).

Already, these trends are pushing medical and educational institutions to make data-informed decisions (Duncan, 2009; Easton, 2009; Shah & Tenenbaum, 2013;

Spencer Foundation, 2011) and generating new industries devoted to the easy interchange of such data (e.g., inBloom.org). Although the relevant technologies in education have barely started to mature, computer-based experimentation and data collection abilities are already making previously unimaginable research possible. The studies presented in this dissertation alone include data from over a quarter million students participating in randomized controlled trials — an experiment sample size that 20 years ago would have been difficult for all but the largest and most deep-pocketed departments of education. The era of the massive behavioral experiment has already dawned. Will modern Lewinians seize the opportunity to establish the value of their approach?

The challenge is to carefully examine the new frontiers that have recently been brought into the domain of the low-cost, mega-experiment. In these domains, we must investigate the interlocking tensile forces that — when tugged the right way — affect powerful changes, changes that shape our understanding of people and changes that, implemented broadly, can shape society itself.

The insights drawn from such work, when sufficiently robust and useful, become broadly-instituted social technologies that make people and organizations more productive. They help political candidates get out the vote (Bryan et al., 2011). They increase the rates of organ donation (Johnson & Goldstein, 2003). They increase sales (Freedman & Fraser, 1966). They reduce the waste of common goods (Allcott,

2011). They make workers more compliant in the face of change (Marrow, 1957). And they might soon help students do well in high school (Chapter 3), earn credits in college (Chapter 4), or learn over the internet (Chapter 5).

Certain psychological insights are simply too valuable to ignore for long. Over time, once-novel ideas about human behavior get integrated into the way people think about the world and the way successful organizations operate. They metamorphose from outlandish, counterintuitive psychobabble into common sense. In the process, they enrich society and send researchers looking for the next big hypothesis.

# References

Adams, G., Tormala, T., & Obrien, L. (2006). The effect of self-affirmation on perception of racism. *Journal of Experimental Social Psychology*, *42*(5), 616–626.

Adelman, C. (2006). The toolbox revisited: Paths to degree completion from high school through college. Washington, DC: U.S. Department of Education, Office of Vocational and Adult Education.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9), 1082–1095.

Allen, I., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Allensworth, E. M., & Easton, J. Q. (2005). *The On-Track Indicator as a predictor of high school graduation.* Chicago, IL: Consortium on Chicago School Research.

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on track and graduating in Chicago public high schools.* Chicago, IL: University of Chicago,

Consortium on Chicago School Research.

Alliance for Excellent Education. (2003). *Progress report on American high schools*. Washington, D.C.: Author.

Alliance for Excellent Education. (2007). The high cost of high school dropouts: What the nation pays for inadequate high schools (Report). Washington, DC: Author.

Allport. (1954). The Nature of Prejudice. Reading, MA: Addison-Wesley Pub. Co.

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C., & Ruble, D. N. (2010, January). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology*, *46*(1), 166–171.

Altschul, I., Oyserman, D., & Bybee, D. (2006). Racial-ethnic identity in mid-adolescence: Content and change as predictors of academic achievement. *Child Development*, *77*(5), 1155–69.

American Association of Community College. (2013). Community College Fact Sheet. Retrieved from `http://www.aacc.nche.edu/AboutCC`

American Psychological Association. (2002). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. Author. Retrieved from `http://www.apa.org/pi/oema/resources/policy/multicultural-guidelines.aspx`

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of educational psychology*, *84*(3), 261–271.

Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *The American Journal of Distance Education*, *22*(3), 146–158.

Aronson, E., Blaney, N., Stephin, C., Sikes, J., & Snapp, M. (1978). *The jigsaw class-room*. Beverly Hills, CA: Sage.

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.

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology*, *35*(1), 29–46.

Augustine, N. (2005). Rising above the gathering storm: Energizing and employing America for a brighter economic future. Washington, DC: National Academy of Science, National Academy of Engineering, Institute of Medicine, National Academy Press.

Bailey, T. R., & Alfonso, M. (2005). Paths to persistence: An analysis of research on program effectiveness at community colleges. In *New Agenda Series, 6*(1). Indianapolis, IN: Lumina Foundation for Education.

Bailey, T. R., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, *29*(2), 255–270.

Barnard College. (2013). Academic success and enrichment program. Retrieved from `http://barnard.edu/asep`

Bar-Tal, D., Halperin, E., & Oren, N. (2010). Socio-psychological barriers to peace making : The case of the Israeli Jewish society. *Social Issues and Policy*, *4*(1), 63–109.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes [R package Manual]. Retrieved from `http://CRAN.R-project.org/package=lme4`

Baum, S., Ma, J., & Payea, K. (2010). *Education pays 2010: The benefits of higher education for individuals and society.* Trends in Higher Education Series, New York, NY.

Baumeister, R., & Leary, M. (1995). The need to belong: Desire for interpersonal attachments. *Psychological Bulletin*, *117*(3), 497.

Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*, *1*(3), 199–218.

Berliner, D. C. (2006). Educational psychology: Searching for essence throughout a century of influence. In *Handbook of Educational Psychology* (2nd ed., pp.

3–27). Mahwah, NJ: Erlbaum.

Berman, P., & McLaughlin, M. W. (1976). Implementation of educational innovation. *The Educational Forum*, *40*(345–370).

Binning, K. R., Sherman, D. K., Cohen, G. L., & Heitland, K. (2010, December). Seeing the other side: Reducing political partisanship via self-affirmation in the 2008 presidential election. *Analyses of Social Issues and Public Policy*, *10*(1), 276–292.

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*. New York: Springer.

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.

Boggiano, A. K., & Katz, P. (1991). Maladaptive achievement patterns in students: The role of teachers' controlling strategies. *Journal of Social Issues*, *47*(4), 35–51.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology*, *40*.

Boykin, A. W., Albury, A., Tyler, K. M., Hurley, E. A., Bailey, C. T., & Miller, O. A. (2005). Culture-based perceptions of academic achievement among low-income

elementary students. *Cultural Diversity & Ethnic Minority Psychology*, *11*(4), 339–50.

Boykin, A. W., Lilja, A. J., & Tyler, K. M. (2004). The influence of communal vs. individual learning context on the academic performance in social studies of grade 4–5 African-Americans. *Learning Environments Research*, *7*(3), 227–244.

Bridgeland, J. M., DiIulio, J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts* (No. March). Washington, D.C.: Civic Enterprises.

Brophy, S., Biswas, G., Katzlberger, T., Bransford, J., & Schwartz, D. (1999). Teachable agents: Combining insights from learning theory and computer science. *Learning Technology*, *50*, 21–28.

Bryan, C. J., Walton, G. M., Rogers, T., & Dweck, C. S. (2011). Motivating voter turnout by invoking the self. *Proceedings of the National Academy of Sciences*, *108*(31), 12653–12656.

Bryk, A. S. (2009). Support a science of performance improvement. *Phi Delta Kappan*, *90*(8), 597–600.

Bryk, A. S., & Schneider, B. L. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation Publications.

Bureau of Labor Statistics. (2005). *College enrollment and work activity of 2004 high school graduates* (Report). Washington, DC. Retrieved from `http://`

www.bls.gov/news.release/archives/hsgec_03242006.pdf

Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. New York, NY: Cambridge University Press.

Carnevale, A. P., Rose, S. J., & Cheah, B. (2011). *The college payoff: Education, occupations*. Washington, DC: The Georgetown University Center on Education and the Workforce.

Carnevale, A. P., Smith, N., & Strohl, J. (2010). *Help wanted: Projections of jobs and education requirements through 2018*. Washington, DC: Georgetown University, Center on Education and the Workforce.

Center for Community College Student Engagement. (2011). Community college survey of student engagement (Report). Austin, TX: The University of Texas at Austin, Community College Leadership Program. Retrieved from http://www.ccsse.org/survey/survey.cfm

Center for Community College Student Engagement. (2012). A matter of degrees: Promising practices for community college student success (A first look). Austin, TX: The University of Texas at Austin, Community College Leadership Program.

Chajut, E., & Algom, D. (2003). Selective attention improves under stress: Implications for theories of social cognition. *Journal of Personality and Social Psychology*, *85*(2), 231–248.

Chen, X., Wu, J., & Tasoff, S. (2010). *Postsecondary expectations and plans for the high school senior class of 2003-2004 (NCES 2010-170rev)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U. S.

Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*(1), 5–6.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307–10.

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), 400–3.

Cohen, G. L., & Prinstein, M. J. (2006). Peer contagion of aggression and health risk behavior among adolescent males: An experimental investigation of effects on public conduct and private attitudes. *Child Development*, *77*(4), 967 – 983.

Cohen, G. L., Sherman, D. K., Bastardi, A., Hsu, L., McGoey, M., & Ross, L. (2007, September). Bridging the partisan divide: Self-affirmation reduces ideological closed-mindedness and inflexibility in negotiation. *Journal of Personality and Social Psychology*, *93*(3), 415–30.

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.

Coleman, J. S. (1961). *The adolescent society: The social life of the teenager and its impact on education*. New York, NY: Free Press of Glencoe.

Cook, J., Purdie-Vaughns, V., Garcia, J., & Cohen, G. (2011, April). *Longitudinal impact of a values affirmation and its timing on minority students' belonging and performance in middle school*. Presentation at the biennial meeting for the Society for Research in Child Development, Montreal, Canada.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. *Advances in Experimental Social Psychology*, *17*, 229–266.

Crandall, C. S. (1988). Social contagion of binge eating. *Journal of Personality and Social Psychology*, *55*(4), 588–598.

Creswell, J. D., Welch, W. T., Taylor, S. E., Sherman, D. K., Gruenewald, T. L., & Mann, T. (2005). Affirmation of personal values buffers neuroendocrine and psychological stress responses. *Society*, *16*(11), 846–851.

Damon, W., Menon, J., & Bronk, K. (2003). The development of purpose during adolescence. *Applied Developmental Science*, *7*(3), 119–128.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology*, *8*(4), 377–383.

Davidson, E. S., & Benjamin, L. T. (1987). A history of the child study movement

in America. In J. A. Glover & R. R. Ronning (Eds.), *Historical Foundations of Educational Psychology* (pp. 41–60). New York, NY: Plenum Press.

Day, J., & Newburger, E. (2002). *The big payoff: Educational attainment and synthetic estimates of work-life earnings*. Current population reports. Washington, DC: U.S. Census Bureau.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268.

Deci, E. L., & Ryan, R. M. (2004). Handbook of self-determination research. Rochester, NY: University Rochester Press.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics*, *1*(3), 111–134.

Derks, B., Scheepers, D., Van Laar, C., & Ellemers, N. (2011). The threat vs challenge of car parking for women: How self- and group affirmation affect cardiovascular responses. *Journal of Experimental Social Psychology*, *47*(1), 178–183.

Dill, E. M., & Boykin, A. W. (2000). The comparative influence of individual, peer tutoring, and communal learning contexts on the text recall of African American children. *Journal of Black Psychology*, *26*(1), 65–78.

Dixon, M. (2013). Public education finances: 2011 (G11-ASPEF). Washington,

D.C.: U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau. Retrieved from http://www2.census.gov/govs/school/11f33pub.pdf

Duncan, A. (2009, June 8). Robust data gives us the roadmap to reform. Speech made at the Fourth Annual IES Research Conference, Washington, DC. Retrieved from http://www.ed.gov/news/speeches/2009/06/06-82009.html

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology*, *99*(3), 597.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237–256.

Dutton, J., Dutton, M., & Perry, J. (2002). How do online students differ from lecture students? *Journal of asynchronous learning networks*, *6*(1).

Dweck, C. S. (1999). Self-theories: Their role in motivation, personality, and development. Philadelphia, PA: Taylor & Francis.

Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. In E. M. Hetherington (Ed.), Handbook of child psychology: Socialization, personality, and social development (Vol. 4, pp. 643–691). New York, NY: Wiley.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation

and personality. *Psychological Review*, *95*(2), 256–273.

Dweck, C. S., Walton, G. M., & Cohen, G. (2011). *Academic tenacity.* White paper prepared for the Gates Foundation, Seattle, WA.

Easton, J. Q. (2009, July). Using data systems to drive school improvement. Keynote address at the STATS-DC 2009 Conference, Bethesda, MD.

Edgecombe, N. (2011). Accelerating the academic achievement of students referred to developmental education (CCRC Working Paper No. 30, Assessment of Evidence Series). New York, NY: Columbia University, Teachers College, Community College Research Center.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology*, *70*(3), 461–475.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519.

Erdley, C., Cain, K., Loomis, C., Dumas-Hines, F., & Dweck, C. S. (1997). The relations among children's social goals, implicit personality theories and response to social failure. *Developmental Psychology*, *33*.

Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance.* Chicago, IL:

University of Chicago, Consortium on Chicago School Research.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Fike, D., & Fike, R. (2008). Predictors of first-year student retention in community college. *Community College Review*, *36*(2), 68–88.

Finkel, E. J., Slotter, E. B., Luchies, L. B., Walton, G. M., & Gross, J. J. (in press). A brief intervention to promote conflict reappraisal preserves marital quality over time. *Psychological Science*.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, *59*(2), 117–142.

Fishbach, A., & Trope, Y. (2005). The substitutability of external control and self-control. *Journal of Experimental Social Psychology*, *41*(3), 256–270.

Fishbach, A., Zhang, Y., & Trope, Y. (2010). Counteractive evaluation: Asymmetric shifts in the implicit value of conflicting motivations. *Journal of Experimental Social Psychology*, *46*(1), 29–38.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., ... Ji., P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, *6*(3), 151–175.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, *4*(2), 196–202.

Fryer, R. G., & Torelli, P. (2010). An empirical analysis of 'acting White'. *Journal of Public Economics*, *94*(5), 380–396.

Fullan, M. (2007). The new meaning of educational change (4th ed.). New York, NY: Teachers College Press.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148–162.

Gainor, K. A., & Lent, R. W. (1998). Social cognitive expectations and racial identity attitudes in predicting the math choice intentions of Black college students. *Journal of Counseling Psychology*, *45*, 403–413.

Gallup and Lumina Foundation. (2013). America's call for higher education redesign: The 2012 Lumina foundation study of the American public's opinion on higher education. Retrieved from `http://www.gallup.com/poll/160298/americans-say-postsecondary-degree-vital-barriers.aspx`

Giner-Sorolla, R. (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science*, *7*(6), 562–571.

Gladwell, M. (2000). *The tipping point: How little things can make a big difference*. Boston, MA: Little, Brown.

Godes, O., Hulleman, C. S., & Harackiewicz, J. M. (2007). Boosting students' interest in math with utility value: Two experimental tests. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Goldrick-Rab, S., Harris, D., Benson, J., & Kelchen, R. (2011). Conditional cash transfers and college persistence: Evidence from a randomized needbased grant program (Discussion Paper, no. 1393-11). Institute for Research on Poverty.

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*(1), 17–28.

Good, C., Aronson, J. A., & 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.

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *Journal of Early Adolescence*, *13*(1), 21–43.

Graczvk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling up of social-emotional and academic innovations in public schools. *School Psychology Review*, *32*(3), 303–319.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*(3), 541–53.

Grubb, W. N. (2010, September). *The quandaries of basic skills in community colleges: Views from the classroom* (NCPR Working Paper). Paper presented at the National Center for Postsecondary Research developmental education conference, New York, NY.

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. Unpublished manuscript.

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological Science*, *23*(8), 899–906.

Harlow, C. W. (2003). Education and correctional populations (Bureau of Justice statistics special report). Washington, DC: U.S. Department of Justice.

Hattie, J. A. (1999, June). Influences on student learning (Inaugural professorial address). University of Auckland, New Zealand. Retrieved from `http://www.arts.auckland.ac.nz/staff/index.cfm?P=8650`

Hattie, J. A., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Hembee, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal*

*for Research in Mathematics Education*, *21*(1).

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin*, *128*(774).

Heppen, J. B., & Therriault, S. B. (2008). Developing early warning systems to identify potential high school dropouts (Tech. Rep.). American Institutes for Research, National High School Center.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of educational research*, *70*(2), 151–179.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Greenwich, CT: Information Age.

Hiltz, S. R. (1997). Impacts of college-level courses via asynchronous learning networks: Some preliminary results. *Journal of Asynchronous Learning Networks*, *1*(1).

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–599.

Horn, L., & Nevill, S. (2006). Profile of undergraduates in U.S. postsecondary education institutions: 2003–04 with a special analysis of community college students (184th ed.) Washington, DC: U. S. Department of Education, National Center for Education Statistics.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; psychological studies of opinion change*. New Haven, CT: Yale University Press.

Hulleman, C. S., Durik, A. M., Schweigert, S. B., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology*, *100*(2), 398–416.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2009). Enhancing interest and performance with a utility value intervention. *Journal of Education Psychology*, *102*(4), 880–895.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410–1412.

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological bulletin*,

*136*(3), 422–49.

Hurley, E. A., Boykin, A. W., & Allen, B. A. (2005, November). Communal versus individual learning of a math-estimation task: African american children and the culture of learning contexts. *The Journal of Psychology*, *139*(6), 513–27.

Ilon, L., & Normore, A. H. (2006). Relative cost-effectiveness of school resources in improving achievement. *Journal of Education Finance*, *31*(3), 238–254.

Institute of Education Sciences. (2012). Search of funded research grants and contracts for 2012. Retrieved June 25, 2013, from `ies.ed.gov/funding/grantsearch/`

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.

Jamieson, J. P., & Harkins, S. G. (2009). The effect of stereotype threat on the solving of quantitative GRE problems: A mere effort interpretation. *Personality and Social Psychology Bulletin*, *35*(10), 1301–14.

Jamieson, J. P., Harkins, S. G., & Williams, K. D. (2010). Need threat can motivate performance after ostracism. *Personality and social psychology bulletin*.

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology*, *46*(1), 208–212.

Jang, H. (2008). Supporting students' motivation, engagement, and learning during

an uninteresting activity. *Journal of Educational Psychology*, *100*(4), 798–811.

Jenkins, D., & Cho, S. W. (2011). Get with the program: Accelerating community college students' entry into and completion of programs of study (CCRC Working Paper No. 32). New York, NY: Columbia University, Teachers College, Community College Research Center.

Johnson, E., & Goldstein, D. (2003). Do defaults save lives? *Science*, *302*, 1338–1339.

Juvonen, J., Wang, Y., & Espinoza, G. (2011). Bullying experiences and compromised academic performance across middle school grades. *Journal of Early Adolescence*, *31*(152).

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal*, *102*(1), 59–80.

Khan, S. (2011). *Let's use video to reinvent education.* Retrieved from `http://www.ted.com/talks/salman_khan_let_s_use_video_to_reinvent_education.html`

Kolbe, T., Partridge, M., & O'Reilly, F. (2011). Time and learning in schools: A national profile. Boston, MA. Retrieved from `http://www.timeandlearning.org/?q=sass`

Kolesnikova, N. (2009). Community Colleges: A Route of Upward Economic Mobility (Community Development Report). Federal Reserve Bank of St. Louis. Retrieved from `www.stlouisfed.org/community_development/assets/pdf/CommunityColleges.pdf`

Koole, S. L., & Lakens, D. (2012). Rewarding replications: A sure and simple way to improve psychological science. *Perspectives on Psychological Science*, *7*(6), 608–614.

Labaree, D. F. (1998). Educational researchers: Living with a lesser form of knowledge. *Educational Researcher*, *27*(8), 4–12.

Lee, J. Q., McInerney, D. M., Liem, G. A. D., & Ortiga, Y. P. (2010). The relationship between future goals and achievement goal orientations: An intrinsic–extrinsic motivation perspective. *Contemporary Educational Psychology*, *35*(4), 185–202.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology*, *38*(4), 424–430.

Lesgold, A. M., & Welch, R. M. (Eds.). (2012). *Improving adult literacy instruction: Options for practice and research.* Washington, DC: National Academies Press.

Levin, H., & Calcagno, J. C. (2008). Remediation in the community college: An evaluator's perspective. *Community College Review*, *35*(3), 181–207.

Lewin, K. (1947). Group decision and social change. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in Social Psychology* (pp. 459–473). New York, NY: Holt.

Lewin, K. (1958). Group decision and social change. In E. E. Maccoby, T. Newcomb, & E. Hartley (Eds.), *Readings in Social Psychology* (pp. 196–211). New York: Holt, Rinehart, and Winston.

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), 197–213.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, *65*(3), 272–292.

Logel, C., Walton, G. M., Peach, J., Spencer, S. J., & Zanna, M. P. (2010). An intervention to improve women's grades in masculine engineering domains. Paper presented at the 8th biennial convention of the Society for the Psychological Study of Social Issues, New Orleans, LA.

Lohr, S. (2013, June 20). Sizing up big data. *The New York Times*, F1.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098.

Ludwig, J., & Phillips, D. (2007). *The Benefits and Costs of Head Start* (NBER Working Paper no. 12973). Cambridge, MA: National Bureau of Economic

Research. Retrieved from http://www.nber.org/papers/w12973

Mangels, J. A., Butterfield, B., Lamb, J., Good, C., & Dweck, C. S. (2006). Why do beliefs about intelligence influence learning success? A social cognitive neuroscience model. *Social Cognitive and Affective Neuroscience*, *1*(2), 75–86.

Markow, D., & Pieters, A. (2011). *The MetLife survey of the American teacher: Preparing students for college and careers*. New York, NY: MetLife.

Markus, H. R. (2005). On telling less than we can know: The too tacit wisdom of social psychology. *Psychological Inquiry*, *16*(4), 180–184.

Marrow, A. J. (1957). *Making management human*. New York, NY: McGraw-Hill.

Marryshow, D., Hurley, E., Allen, B., Tyler, K., & Boykin, A. (2005). Impact of learning orientation on African American children's attitudes toward high-achieving peers. *The American Journal of Psychology*, *118*(4), 603–618.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, *42*(2), 236–243.

McCord, J. (1978). A thirty-year follow-up of treatment effects. *American Psychologist*, *33*(3), 284–289.

McKnight, P. E., & Kashdan, T. B. (2009). Purpose in life as a system that creates and sustains health and well-being: An integrative, testable theory. *Review of General Psychology*, *13*(3), 242.

Merton, R. K. (1946). *Mass persuasion: The social psychology of a war bond drive*. New York, NY: Harper & Brothers.

Midgley, C., & Urdan, T. (2001). Academic self-handicapping and achievement goals: A further examination. *Contemporary Educational Psychology*, *26*(1), 61–75.

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations*, *18*(1), 57–76.

Miller, R. L., Brickman, P., & Bolen, D. (1975). Attribution versus persuasion as a means for modifying behavior. *Journal of Personality and Social Psychology*, *31*(3), 430–41.

Mischel, W., Cantor, N., & Feldman, S. (1996). Principles of self-regulation: The nature of willpower and self-control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 329–360). New York, NY: Guilford Press.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science*, *330*, 1234–1237.

Molden, D. C., & Dweck, C. S. (2006). Finding "meaning" in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, *61*(3), 192–203.

Moore, M. G., & Kearsley, G. (2005). *Distance education: A systems view (2nd ed.)*.

Belmont, CA: Wadsworth.

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.

Mullin, C. M. (2012). Why access matters: The community college student body (Policy Brief 2012-01PBL). Washington, DC: American Association of Community Colleges.

Nasir, N. S., McLaughlin, M. W., & Jones, A. (2008, July). What does it mean to be African American? constructions of race and academic identity in an urban public high school. *American Educational Research Journal*, *46*(1), 73–114.

National Center for Education Statistics. (2001). Table 5: Overview of public elementary and secondary schools and districts: School year 1999-2000 (Table). Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved from `http://nces.ed.gov/pubs2001/overview/table05.asp`

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Newcomb, T. (1943). Personality and social change: Attitude formation in a student community. New York, NY: Dryden.

Nisbett, R. E., & Valins, S. (1972). Perceiving the causes of one's own behavior. In

E. E. Jones, D. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the Causes of Behavior* (pp. 63–78). Morristown, NJ: General Learning Press.

Nussbaum, A. D., & Dweck, C. S. (2005, April). Managing self-esteem: An implicit theories perspective. Paper preseneted at the annual meeting of the Western Psychological Association, Portland, OR.

Nussbaum, A. D., & Dweck, C. S. (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality and Social Psychology Bulletin*, *34*(5), 599–612.

Osborne, J. W. (1997). Race and academic disidentification. *Journal of Educational Psychology*, *89*(4), 728–735.

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), 188–204.

Packard, B. W., & Nguyen, D. (2003). Science career-related possible selves of adolescent girls: A longitudinal study. *Journal of Career Development*, *29*(4), 251–263.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology & Society*, *12*(4), 207–217.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis of confidence? *Perspectives on Psychological Science*, *7*(6), 528–530.

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*(1), 83–98.

Paunesku, D., Yeager, D., Romero, C., & Walton, G. (2012). A brief growth mindset intervention improves academic outcomes of community college students enrolled in developmental mathematics courses. Unpublished Manuscript, Stanford University.

Petty, R. E., & Cacioppo, J. T. (1986). Communication and persuasion: Central and peripheral routes to attitude change. New York, NY: Springer-Verlag.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. New York, NY: Springer.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin*, *112*(1), 160–164.

Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*(2), 243–56.

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to

stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology*, *40*(2), 152–168.

R Development Core Team. (2006). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org`

Rachlin, H. (2000). *The science of self-control*. Cambridge, MA: Harvard University Press.

Reis, H. T., & Gosling, S. D. (2010). Social psychological methods outside the laboratory. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 82–114). New York, NY: Wiley.

Richard, F., Bond, C., & Stokes-Zoota, J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, *7*(4), 331–363.

Robins, R. W., & Pals, J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, *1*, 313–336.

Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2013). Academic and emotional functioning in middle school: The role of malleability beliefs. (Manuscript submitted for publication)

Romero, C., Paunesku, D., & Dweck, C. (2011). Brainology in the classroom: An

online growth mindset intervention affects GPA, conduct, and implicit theories. *Poster session presented at Society for Research in Child Development Biennial Meeting*, Montreal, Canada.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *33*(1), 135–146.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301.

Ross, L., Lepper, M., & Ward, A. (2010). History of social psychology: Insights, challenges, and contributions to theory and application. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), Handbook of social psychology (5th ed., Vol. 1, pp. 3–50). Hoboken, NJ: Wiley.

Ross, L., & Nisbett, R. (1991). *The Person and the situation: Perspectives of social psychology.* New York, NY: Mcgraw-Hill.

Rouse, C. E. (2005). *The labor market consequences of an inadequate education.* New York, NY: Columbia University.

Roush, W. (2012). *Can Anyone Catch Khan Academy? The Fate of the U in the YouTube Era.* Xconomy. Retrieved from `http://www.xconomy.com/national/2012/07/20/can-anyone-catch-khan-academy-the-fate-of-the-u-in-the-youtube-era`

Rush, A. (2009). The role of efficacy and effectiveness trials. *World Psychiatry*, *8*(1), 34–35.

Rutschow, E. Z., Richburg-Hayes, L., Brock, T., Orr, G., Cerna, O., Cullinan, D., & Martin, K. (2011). Turning the tide: Five years of achieving the dream in community colleges. New York, NY: Manpower Demonstration Research Corporation.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, *85*(3), 440–452.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434.

Schwab, K. (2012). *The global competitiveness report 2012–2013*. Geneva: World Economic Forum.

Shah, N. H., & Tenenbaum, J. D. (2013). The coming age of data-driven medicine: Translational bioinformatics' next frontier. *Journal of the American Medical Informatics Association*, *19*(1), e2–e4.

Sherman, D. K., Bunyan, D. P., Creswell, J. D., & Jaremka, L. M. (2009). Psychological vulnerability and stress : The effects of self-affirmation on sympathetic nervous system responses to naturalistic stressors. *Health Psychology*, *28*(5),

554 –562.

Sherman, D. K., Cohen, G., Nelson, L., Nussbaum, A., Bunyan, D., & Garcia, J. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology*, *97*(5), 745–764.

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In *Advances in Experimental Social Psychology* (Vol. 38, pp. 183–242). San Diego, CA: Academic Press.

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.

Sherman, D. K., Nelson, L., & Steele, C. (2000). Do messages about health risks threaten the self? Increasing the acceptance of threatening health messages via self-affirmation. *Personality and Social Psychology Bulletin*, *26*(9), 1046–1058.

Silva, E., & White, T. (2013). Pathways To Improvement: Using Psychological Strategies To Help College Students Master Developmental Math. Washington, D.C.: Carnegie Foundation. Retrieved from `http://www.carnegiefoundation.org/spotlight/productive-persistence-pathways-improvement`

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology:

Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Snyder, T. D., & Dillow, S. A. (2011). Digest of education statistics, 2010. *Statistics*, *015*, 744.

Spencer Foundation. (2011). Initiative on data use and educational improvement. Retrieved from `http://www.spencer.org/content.cfm/data-use-and-educational-improvement-details`

Spies, J. R., Nosek, B. A., Bartmess, E., Lai, C., Galak, J., Cohn, M., ... Alvarez, L. C. (2013). Reproducibility Project. Retrieved from `http://opensciencframework.org/project/EZcUj/wiki/home`

Stake, J. E., & Mares, K. R. (2001). Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation. *Journal of Research in Science Teaching*, *38*(10), 1065–1088.

Stam, H. J. (2006). Theory and practice. In C. W. Tolman, F. Cherry, R. V. Hezewijk, & I. Lubek (Eds.), *Problems in Theoretical Psychology*. New York, NY: Captus.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261–302). New York, NY: Academic Press.

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

Steele, C. M. (2010). Whistling Vivaldi: And other clues to how stereotypes affect us. New York, NY: Norton & Company Inc.

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

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, *45*(1), 5.

Stephan, J. J. (2004). State prison expenditures, 2001 (Special Report No. NCJ 202949). Bureau of Justice Statistics.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, *77*(6), 1213–1227.

Storms, M. D., & McCaul, K. D. (1976). Attribution processes and emotional exacerbation of dysfunctional behavior. *New directions in attribution research*, *1*, 143–164.

Storms, M. D., & Nisbett, R. E. (1970). Insomnia and the attribution process. *Journal of Personality and Social Psychology*, *16*(2), 319.

Taylor, V., & Walton, G. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin*, *37*(8), 1055–1067.

Thaler, R. H., & Shefrin, H. M. (1981). An economic theory of self-control. *The*

*Journal of Political Economy*, *80*(02), 392–406.

Thorndike, E. L. (1903). *Educational Psychology*. New York: Science Press.

Trope, Y., & Fishbach, A. (2000). Counteractive self-control in overcoming temptation. *Journal of Personality and Social Psychology*, *79*(4), 493.

Turner, J. C., Midgley, C., Meyer, D. K., Gheen, M., Anderman, E. 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*(1), 88–106.

Tyack, D. B., & Cuban, L. (1995). Tinkering toward utopia: A century of public school reform. Cambridge, MA: Harvard University Press.

Tyler, K. M., Uqdah, A. L., Dillihunt, M. L., Beatty-Hazelbaker, R., Conner, T., Gadson, N., ... Stevens, R. (2008). Cultural discontinuity: Toward a quantitative investigation of a major hypothesis in education. *Educational Researcher*, *37*(5), 280–297.

Tyson, K., Darity, W., & Castellino, D. R. (2005, August). It's not "a Black thing": Understanding the burden of acting White and other dilemmas of high achievement. *American Sociological Review*, *70*(4), 582–605.

University of California Santa Cruz. (2012). Academic Enrichment Programs. Retrieved from `http://pbsci.ucsc.edu/academics/academic-enrichment.html`

U.S. Food and Drug Administration. (1998). Guidance for industry: Providing clinical evidence of effectiveness for human drug and biological products. Rockville, MD.

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, *72*(5), 1161–76.

Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory : Another look at the quality of academic motivation. *Psychology*, *41*(1), 19–31.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, *87*(2), 246–260.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, *331*(6023), 1447–1451.

Walton, G. M., Cohen, G. L., J., G., Master, A., Cook, J. E., Purdie-Vaughns, V., & Apfel, N. (2013, April 22). A brief social-belonging intervention improves

classroom behavior over seven years among ethnic minority adolescents. *Society for Research in Child Development*, Seattle, WA.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (in prep.). Two brief social-psychological interventions transform women's experience, relationships, and achievement in engineering. Manuscript in preparation.

Walton, G. M., Spencer, S. J., & Erman, S. (2013). Affirmative meritocracy. *Social Issues and Policy Review*, *7*(1), 1–35.

Watson, J., Murin, A., Vashaw, L., Gemin, B., & Rapp, C. (2010). Keeping Pace with K-12 Online Learning: An Annual Review of Policy and Practice (Annual Review). Evergreen Education Group. Retrieved from `http://kpk12.com/cms/wp-content/uploads/KeepingPace2012.pdf`

Weiss, M. J., Visher, M. G., Wathington, H., Teres, J., & Schneider, E. (2010). Learning communities for students in developmental reading: An impact study at Hillsborough Community College. New York, NY: Columbia University, Teachers College, National Center for Postsecondary Research.

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*(3), 411–419.

Wentzel, K. R. (2002). Are effective teachers like good parents? teaching styles and student adjustment in early adolescence. *Child Development*, *73*(1), 287–301.

Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of

college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, *42*(2), 367–376.

Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, *49*(1), 287–293.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, *96*(2), 236–250.

Woolf, K., McManus, I. C., Gill, D., & Dacre, J. (2009). The effect of a brief social intervention on the examination results of UK medical students: A cluster randomised controlled trial. *BMC Medical Education*, *9*, 35.

Yamauchi, L. A. (2003). Making school relevant for at-risk students: The Wai'anae high school Hawaiian studies program. *Journal of Education for Students Placed at Risk*, *8*(4), 379–390.

Yeager, D. S., & Bundick, M. J. (2009). The role of purposeful work goals in promoting meaning in life and in schoolwork during adolescence. *Journal of Adolescent Research*, *24*(4), 423–452.

Yeager, D. S., Bundick, M. J., & Johnson, B. (2012). The role of future work goal motives in adolescent identity development: A longitudinal mixed-methods investigation. *Contemporary Educational Psychology*, *37*, 206–217.

Yeager, D. S., Henderson, M., Paunesku, D., Walton, G., Spitzer, B., D'Mello, S., & Duckworth, A. (2013). A purpose for learning. Manuscript in preparation.

Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, *75*(4), 709–747.

Yeager, D. S., Paunesku, D., Romero, C., Brown, R., Muhich, J., & Walton, G. M. (2013). Engineering psychological interventions for scale: The case of the "growth mindset" in community colleges. (Manuscript in preparation)

Yeager, D. S., Purdie-Vaughns, V., Garcia, P. P., J., Master, A., Hessert, W., Williams, M., & Cohen, G. (in press). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*.

Yeager, D. S., Trzesniewski, K. H., Tirri, K., Nokelainen, P., & Dweck, C. S. (2011). Adolescents' implicit theories predict desire for vengeance after peer conflicts: Correlational and experimental evidence. *Developmental Psychology*, *47*(4), 1090 –1107.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301.

Zhang, Y., & Fishbach, A. (2010). Counteracting obstacles with optimistic predictions. *Journal of Experimental Psychology: General*, *139*(1), 16–31.

Zimbardo, P. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In Nebraska symposium on motivation (Vol. 17, pp. 237–307). Lincoln, NE: University of Nebraska Press.