# The Effects of Two Mindset Interventions on Low-Income Students' Academic and Psychological Outcomes

Jill Gandhi[1], Tyler W. Watts[2], Michael D. Masucci[3], & C. Cybele Raver[1]

[1] Steinhardt School of Culture, Education and Human Development, New York University, New York, New York, United States of America

[2] Teachers College, Columbia University, New York, New York, United States of America

[3] Department of Psychology, Louisiana State University, Baton Rouge, Louisiana, United States of America

Paper published in the Journal of Research on Educational Effectiveness, March 6, 2020

**Acknowledgement**

Research reported in this publication was supported by the Institute of Education Sciences, U.S. Department of Education, under award number R305A160176 ($3,210,436). This project was 100% financed with federal money. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences. We are grateful to David Yeager, Alejandro Ganimian, Chen Li, Javanna Obregon, and Alaa Khader for their helpful contributions to this project. We would also like to thank Chicago Public Schools and the dedicated center directors, teachers, families, and students who made the Chicago School Readiness Project possible.

*Corresponding Author:*
jill.gandhi@nyu.edu

Academic achievement has historically been the primary marker of college readiness for high school students. However, there are also socioemotional, motivational, and behavioral factors that can boost, or limit, students' psychological preparation for higher education, and their chances of completing high school on time (Farrington et al., 2012; Nagaoka et al., 2013). "Wise" mindset interventions have captured the attention of researchers, policymakers, and educators as a "light-touch" approach to increasing students' motivation for academic persistence into and through college (Broda et al., 2018; Yeager & Walton, 2011). Wise mindset interventions often consist of a set of simple classroom writing assignments that have been designed to covertly alter students' self-representations as learners with the high potential to support their orientation towards coping with academic hurdles, and to shape their construal of schools as places to which they can belong (Walton, 2014). Recent research on wise mindset interventions has tested the feasibility of one-time, self-administered treatments that attempt to shift student mindsets, with several studies offering compelling evidence of long-term impacts on academic outcomes for adolescents and college students (Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Dweck, Walton, & Cohen, 2011; Yeager & Walton, 2011).

Although initial efficacy research on mindset interventions has been promising (e.g., Farrington, 2013; Paunesku et al., 2015; Walton & Cohen, 2011), the theoretical mechanism through which these interventions move student behavior remains unclear. Some have hypothesized that mindset interventions work in a recursive manner, providing an initial nudge that then cascades into positive, self-reinforcing behavior that alters achievement trajectories over time (DeBacker et al., 2018; Yeager & Walton, 2011). In many cases, these brief interventions may not demonstrate academic benefits until months or years following the intervention (Dweck et al., 2011).

Previous research has suggested that mindset interventions should target at-risk populations, as the strongest treatment effects have been observed on the academic performance of low-income students (Broda et al., 2018; Sisk, Burgoyne, Sun, Butler, & Macnamara, 2018; Walton & Cohen, 2011). However, evidence for this mechanism is limited. Recent studies have argued that to better understand the mechanism through which these interventions operate, more research should focus on proximate outcomes collected before academic achievement measures – such as emotion regulation and inhibitory control – particularly for students from diverse circumstances and populations (Hanselman, Bruch, Gamoran, & Borman, 2014).

In the current study, we aimed to shed light on the hypothesized psychological mechanisms through which mindset interventions might operate by testing two different wise mindset interventions on the same sample of treatment group students, one year apart: 1) an intervention targeting the students' purpose for learning; 2) a second intervention one year later that sought to increase the students' growth mindset beliefs. Further, we tested these interventions with a sample of low-income, ethnic minority (65% identified as Black) high school students from the Chicago area, and collected a host of potential outcome measures to observe whether the interventions affected underlying psychological processes as well as academic achievement. As such, we sought to answer two questions: 1) Can two brief mindset interventions effectively boost academic achievement in a population of low-income youth living in urban communities with high violence exposure? 2) In addition to grade point average (GPA), do these mindset interventions shift specific psychological processes that may contribute to students' academic achievement, including task diligence, anxiety, critical motivation, or sense of belonging uncertainty in their schools?

**Background**

Broadly, mindset interventions seek to target a set of psychological processes (including attitudes, strategies, and behaviors) that have been hypothesized to substantially undergird students' academic achievement and motivation (Farrington et al., 2012; Nagaoka et al., 2013). Mindset interventions can be intensive, such as a growth mindset intervention that operated as a workshop taught over eight classroom sessions (Blackwell, Trzesniewski, & Dweck, 2007), or as brief as a one-time 15 to 20 minute writing assignment (Cohen et al., 2009). The appeal of one-time, brief interventions in school settings is clear: such wise interventions are affordable, scalable, and hold the potential to increase student performance without the need for costly teacher training or repeated visits. Intervention developers have tested a variety of brief writing assignments, many of which attempt to alter related, yet distinct, aspects of student emotion regulation, motivation, and self-efficacy. Here, we review two popular light-touch approaches to shifting student mindset: 1) purpose for learning and 2) growth mindset.

**Purpose for learning interventions.** Motivation and achievement are thought to rise when students find personal value and meaning in their schoolwork (Hulleman, Godes, Hendricks, & Harackiewicz, 2010). Research on motivation has long identified the benefits of purpose in adolescence on a variety of outcomes, including affective well-being, academic performance, and academic persistence (Hill, Burrow, & Sumner, 2013). These benefits have been harnessed in several kinds of interventions, including utility value interventions that trigger an association between the academic task at hand and the student's enduring personal interests (Hulleman et al., 2010), and goal-setting interventions that ask students to set personally meaningful objectives for their lives that can be attained through academic achievement (Morisano, Hirsh, Peterson, Pihl, & Shore, 2010). These interventions are based on the theory that posits that students who have "purpose" see their academic work as a contribution to their personal transformation, so they can contribute something of personal value to the world – even if the schoolwork does not seem immediately applicable to their goals (Dweck et al., 2011).

Purpose for learning interventions target students' core beliefs to answer the question "why should I learn?" Prior research explored students' spontaneous responses to this question and found that students who expressed prosocial, self-transcendent motives for learning (e.g., "I want to become an educated citizen who can contribute to society") demonstrated more effective academic self-regulation and greater college persistence than students who expressed extrinsic motives (e.g., "I want to earn more money") or self-oriented, intrinsic motives (e.g., "I want to expand my knowledge of the world") (Yeager et al., 2014). Findings on the motivational benefits of purpose have led researchers towards examining how to foster a sense of purpose in youth, and with which populations and contexts this kind of intervention would be most effective.

A subsequent randomized control trial of a one-time purpose for learning intervention was conducted with upper- and middle-class high school and college students and found that the purpose intervention improved overall grades in math and science courses with an effect size of $d = 0.11$ (Yeager et al., 2014). Furthermore, research on an online purpose for learning intervention found that this intervention was particularly effective at improving the GPAs of academically at-risk students (Paunesku et al., 2015). As a result of that prior study and evidence of benefits for at-risk students from other mindset intervention trials, the present study focused on supporting ethnic minority students in under-resourced schools.

**Growth mindset interventions.** Growth mindset interventions are based on research suggesting that people develop implicit beliefs about their own intelligence (Dweck, 2008b), and these beliefs can be broadly categorized as entity beliefs and incremental beliefs. Entity beliefs describe the belief that intelligence is fixed and cannot be increased. Incremental beliefs refer to the belief that intelligence is malleable and can be grown through effort and practice (Yeager & Dweck, 2012). Growth mindset interventions seek to encourage students to adopt a mindset that views academic ability and competence as malleable through their efforts. One form that these interventions take is using the “my brain is like a muscle” analogy to convey how taking on academic challenges are useful in developing intelligence. A self-administered, online growth mindset intervention given to a diverse sample of students transitioning to high school resulted in a decrease in reported fixed mindset beliefs, as well as an intervention benefit of 0.13 grade points for students who were one standard deviation below the mean of prior performance (i.e., the lower-achieving students) (Yeager et al., 2016).

Growth mindset interventions may be especially beneficial for struggling students’ academic performance due to the way that they reframe challenges as a means to promote learning and resilience (Burnette, Russell, Hoyt, Orvidas, & Widman, 2018). Moreover, there is some evidence that students from lower-income families who hold a growth mindset are buffered from the negative effects of poverty on academic achievement (Claro, Paunesku, & Dweck, 2016). A small-scale growth mindset intervention was tested on low-income, racial/ethnic minority students in an urban middle school and resulted in positive effects of about 0.30 grade points for students in the treatment group over students in the control group, although it is worth noting that this was an intensive intervention (Blackwell et al., 2007). A meta-analysis of growth mindset interventions found that these interventions were most effective for students at high academic risk and from economically disadvantaged backgrounds (Sisk et al., 2018). However, there were few studies contributing to both of these results, the sample sizes tended to be small, and the high-risk students’ effect did not differ significantly from the low-risk students’ effect, which were null.

**Intervention outcomes.** Course grades and GPA are the most commonly examined outcomes of mindset intervention, because grades are often strong predictors of educational attainment, as well as other positive adult outcomes (Easton, Johnson, & Sartain, 2017; Farrington et al., 2012). Grades, by proxy, reflect not just content knowledge and academic skills, but academic behaviors and attitudes such as work habits, motivation, and attendance. However, some researchers have called for further examination behind the theorized recursive mechanism of these wise interventions (Burnette et al., 2018; Yeager, et al., 2014) – what measurable psychological processes explain the positive shift in students' academic outcomes?

Rooted in interpersonal and social psychological theories of identity threat, recent mindset intervention studies have demonstrated that students of marginalized status experience academic challenges as stigmatizing and emotionally uncomfortable or distressing (Good, Aronson & Inzlicht, 2003; Oyserman, Bybee & Terry, 2006). A wealth of studies highlight that, as a result of that emotional distress, students for whom identity threat has been triggered are likely to withdraw effort (including disengaging from classroom discussion, failing to turn in homework assignments, and dropping out of courses) (see Hembree, 1990; Nguyen & Ryan, 2008; Spitzer & Aronson, 2015). Recent work suggests the promise of theories of emotional self-regulation (particularly in the context of distress, frustration, and anxiety) for supporting adolescents' academic performance (McCrae, Ciesielski, & Gross, 2012; Ochsner, et al., 2004). Thus, we examined students' level of anxiety, academic self-regulation (or diligence), belonging uncertainty, and critical motivation as proximal outcomes of the mindset interventions.

***Anxiety.*** For some students, the combined psychological experience of anxiety, disappointment, and frustration can undermine performance and effort expenditure in academic pursuits (Morisano et al., 2010). One potential proximate effect of mindset interventions is a reduction in anxiety and improved regulation of negative emotions that might underlie low motivation and academic avoidance. Yeager and colleagues found that 9th graders who were randomly assigned to mindset interventions were less likely than their control group-assigned counterparts to experience negative emotions in the context of school-related stressors (Miu, Yeager, Sherman, Pennebaker, & Trzesniewski, 2014; Yeager et al., 2014). Additionally, some sense-of-purpose interventions, such as self-affirmation interventions that ask students to write down personally-important values, may reduce stress and help students function more effectively (Sherman & Hartson, 2011).

Additionally, growth mindset interventions may work to reduce the anxiety related to the threat of failure when students hold what has been termed a “fixed mindset belief” that poor academic performance is immutable and intrinsic. Believing that the brain’s capacity to learn and academic performance are malleable (a “growth” mindset) offers the student the view that they can improve through effort, rather than viewing failures as confirmation of their perceived low academic status (Yeager et al., 2018; Wilson & Linville, 1985). Indeed, empirical evidence suggests that viewing social situations through an incremental framework decreases physiological stress response (as measured using cortisol and cardiovascular response) and improves performance outcomes (Jamieson, Mendes, Blackstock, & Schmader, 2010; Ramirez & Beilock, 2011; Yeager, Lee, & Jamieson, 2016). Although students’ ability to regulate negative emotions is clearly implicated in those theoretical frameworks, the role of emotion regulation in mindset intervention efficacy has not been formally tested. Mindset interventions that promote students’ focus on larger personal goals and a malleable view of intelligence may break the loop of anxiety and poor academic performance that comes from self-identifying as a failing student.

***Academic task diligence.*** Academic task diligence, or academic self-regulation, is thought to be an important skill for students who are frequently overwhelmed with tempting alternatives to schoolwork, and a potential mechanism underlying improved GPA. Games, social media, and other distractions compete for students' attention, unless they are highly motivated and adequately focused on their schoolwork. This kind of persistence may result from having a self-transcendent motivation for learning (Yeager, 2014). Indeed, past research suggests that the Academic Diligence Task (Galla et al., 2014), a task that tests students' levels of academic self-regulation, is sensitive to purpose for learning interventions, and has positive associations with students' academic outcomes. We hypothesized that task diligence could also emerge from a growth mindset intervention, due to the motivating messages that repeated practice on academic tasks can be construed as opportunities to "build the brain" and spur grade improvements.

***Belonging uncertainty.*** We examined long-term shifts in student sense of belonging uncertainty in their schools after the purpose for learning intervention. Social belonging, or a sense of having positive relationships with others in a specific context, is an important aspect of academic life. Students from marginalized communities, such as African-American students, may struggle with belonging uncertainty in school (Walton & Cohen, 2011). A sense of social belonging and feeling connected to one's school is a likely outcome of intervention-group students finding meaning in their schoolwork and reframing problems in their community as something that their education may empower them to address. Conceptually, a sense of social belongingness bolsters students' academic self-regulation by imbuing them with the confidence to work through challenges (Spitzer & Aronson, 2015).

***Critical motivation.*** Finally, we explored the role of critical motivation as a potential outcome of both purpose and growth mindset interventions. Critical motivation refers to a marginalized youth's ability to reflect on and question social structures and to transform this awareness into the motivation to produce social change (Diemer, Rapa, Voight, & McWhirter, 2016; McWhirter & McWhirter, 2016). If a purpose for learning intervention meets its goals, then students should more readily connect their academic efforts to their own long-term goals of making a difference for themselves and for society. In a self-reflective writing component of the intervention, students are given an opportunity to share ways they would like to help others and to support change in their communities, and how succeeding in school can support these goals. For low-income students, an orientation towards personally-relevant issues related to social conditions and structural constraints may increase their motivation and academic achievement. Similarly, the growth mindset intervention emphasized societally-oriented benefits of adopting a more flexible, optimistic perspective on learning strategies, increased effort, and practice by focusing on ways that the mindset can help students "to learn in school so they can give back to the community and make a difference in the world later" (Yeager et al., 2016).

**Heterogeneity.** In addition to our focus on low-income, ethnic minority students and the extension of outcomes beyond GPA, we examined heterogeneity of treatment by relevant characteristics within our sample, including baseline trait anxiety, prior academic performance, gender, and ethnic/racial identity. Consistently, findings from a range of psychosocially-focused interventions suggest that students facing the highest levels of risk show greater benefit than do those facing fewer stressors (Gormley, Gayer, Phillips, & Dawson, 2005; Jones, Brown, & Aber, 2011). Unfortunately, our field has reached few clear conclusions as to whether mindset interventions serve students who struggle with managing anxiety versus those who are relatively more skilled in emotional regulation. Given the relation between situational appraisal and task performance (see Jamieson et al., 2009), we examine whether students with high versus low trait anxiety differentially responded to both mindset interventions.

In addition, we examined whether these particular mindset interventions may offer greater or less benefit for those students struggling academically, relative to their higher performing counterparts. Previous mindset research has incorporated moderation by baseline achievement level and grades, suggesting that when students are already high-achieving, a mindset intervention may not be necessary, or may not demonstrate significant effects (Paunesku et al., 2015; Yeager, 2014). In order to test if lower-performing students benefit more from these messages, we tested moderation by higher versus lower baseline grades.

Furthermore, students' negative stereotypes of their identities as members of racial, ethnic and gender categories have consistently been found to contribute to academic underperformance (Dweck, 2008b; Nguyen & Ryan, 2008). Mindset interventions like the growth mindset intervention tested here can reduce the academic gap for women and ethnic minorities, and evidence suggests that intervention effects may be moderated by student race and gender (Broda et al., 2018; DeBacker et al., 2018).

**Current Study**

In the study presented here, we explored the influence of a one-time purpose for learning intervention and a one-time growth mindset intervention on low-income, ethnic minority adolescents' academic outcomes. We also tested the impact of those mindset interventions on a set of key psychological processes believed to support students' academic performance, including their task persistence, anxiety, belonging uncertainty, and critical motivation. At the time of this study, these two interventions were available online (http://www.perts.net)[1] as free modules for use by educators to promote students' academic motivation, resilience, and achievement.[2] These two online interventions were previously tested with a sample of 1,594 high school students in 13 geographically and socioeconomically diverse schools (Paunesku et al., 2015). That study found that academically at-risk students who received both the sense-of-purpose and growth mindset interventions in combination (2 weeks apart) earned GPAs about 0.13 points higher than at-risk students in the control condition. In our study, we examined the effect of both interventions in combination, focusing first on the short-term and long-term (one year later) effects of the purpose intervention, before turning to the short-term effects of the growth mindset intervention on the same set of students.

We conducted this research with a sample of low-income adolescents spread across an estimated 275 schools in a large urban district. We focused both interventions on a sample of ethnic-minority youth from economically disadvantaged neighborhoods in Chicago. Prior evaluations of similar one-time interventions have focused on conducting mindset interventions in classrooms or computer labs, administrated by teachers, across a smaller number of schools (DeBacker et al., 2018; Paunesku et al., 2015). The present study explores the scalability of online administered mindset interventions by having students self-administer the intervention on laptop computers (under researcher supervision) either in their school or in their home.

Based on prior research, we hypothesized that students randomized to the purpose for learning intervention would academically outperform students in the control group and would express lower levels of belonging uncertainty. Additionally, we hypothesized that both

[1] At the time of publication, the two online modules at http://www.perts.net were specific to the Growth Mindset intervention described in this paper.
[2] Minor modifications were made to update the language and to cater it to the population of study, while retaining the language and themes that made the interventions effective in prior trials. See the appendix for a detailed description of the changes made to the original Yeager et al. (2014, 2016) interventions.

interventions would result in higher academic diligence, higher levels of critical motivation, and lower state anxiety for treatment group students than for the control group. In the following sections, we present the baseline measures that were used in both treatment analyses, describe the analytic approach, and then discuss each intervention and its results separately.

# Method

## Study Design

The present study was conducted as part of an ongoing longitudinal study called the Chicago School Readiness Project (CSRP; Raver et al., 2009). The CSRP was a pre-K intervention implemented in low-income Head Start sites serving predominantly ethnic minority communities in Chicago. The intervention boosted children's self-regulation and school readiness skills (see Raver et al., 2008; 2011) and some follow-up work has suggested that the intervention had long-run effects on measures of executive function, emotion regulation, and academic achievement (Watts et al., 2018). Youth who participated in the CSRP from pre-K have been followed longitudinally through adolescence, and ten years after the original treatment program, they were re-randomly assigned to two groups: 1) receipt of both Purpose for Learning (PFL) and Growth Mindset (GM) interventions; 2) to the active control condition. Youth in the treatment group were presented with the standard PFL and GM intervention modules (described in more detail below), administered one year apart. Youth in the active control group also engaged in reading, reflection, and writing activities on unrelated but informational topics.

To help ensure that the re-randomization procedure would be balanced across the original pre-K intervention groups, we linked students back to their original Head Start center and randomized *within* each Head Start center to either the PFL/GM group or control. Indeed, original pre-K intervention group status was not related to the PFL/GM intervention status ten years later (*p* = 0.92). In all subsequent analyses, we control for pre-K intervention group.

As part of the decision to randomize the mindset intervention conditions within the students' original Head Start center, we conducted statistical power calculations to determine if power estimates were adequate for treatment effects estimated at this level of randomization. When predicting student-level outcomes with treatment nested in 18 original Head Start centers, we estimated about 27 students per original center (original cell sizes attenuated by 20% due to longitudinal attrition). Additionally, we estimated intraclass correlations of the dependent variables from urban, high-risk samples participating in experimental evaluations ranging from 0.02 for self-regulatory outcomes to 0.20 for achievement and the alpha set to 0.05. These power estimates are further strengthened by the number and strength of baseline covariates available in this study at the individual and school levels and the longitudinal nature of the data, maximizing our level of predicted variance. Using this information, we estimated that our minimum detectable effect size (MDES analyses conducted using Optimal Design software; Raudenbush et al., 2011) was 0.27 with 80% power, considered a medium effect in the context of educational interventions with ethnic/racial minority students (Hill, Bloom, Black, & Lipsey, 2008). This effect size would be on the larger side for other recent estimates of the effect of growth mindset on academic achievement (see Dweck & Yeager, 2019). Our other outcome variables have not yet been tested in conjunction with PFL or growth mindset interventions, making it difficult to project how our MDES would compare to other published trials.

Figure 1 displays the timeline of the study. As Figure 1 shows, 463 adolescents were randomly assigned to either treatment or control for the PFL intervention wave, which occurred 10 years after the original pre-K study began. Immediate outcome measures were collected after the PFL intervention or control activities. In the next school year, 404 adolescents participated (12.7% attrition from the PFL intervention wave), and treatment group status was retained. Prior to administering the GM intervention, we collected information on student behavior and academic achievement, which served as a 1-year follow-up to the PFL intervention. Immediately following the GM intervention, short-term outcome data was collected.

**Baseline Measures**

All baseline measures were collected prior to PFL/Mindset intervention administration. In Table 1, we present a complete list of baseline measures, and we provide a brief overview of each measure and the timing of data collection in the sections that follow.

**Demographics.** Student demographics such as gender, age, and race/ethnicity were collected from parents during the Head Start year of the study, or Year 1. Additionally, parent age, parent race/ethnicity, and mother's education (binary measure with "1" indicating at least a high school completion) were collected during the Head Start year. The family's income to needs ratio was collected during the year of the PFL intervention (Year 10) and was calculated as the total family income from the previous year divided by that same year's federal poverty threshold. Parents also self-reported the number of children in their home, whether they were a single parent (is a single parent = 1), and how many hours they work in a typical week.

**Cognitive skills.** The Hearts and Flowers task (Diamond, Barnett, Thomas, & Munro, 2007), a commonly used measure of executive function, was used as our primary indicator of the youth's baseline cognitive skill and was collected immediately prior to the PFL intervention tasks. The task asks students to press a key ("Q" or "P") in response to stimuli (either hearts or flowers) that appear in succession on opposite sides of a computer screen. Hearts were associated with a congruent response, in which students pressed on the key that was on the same side of the screen as the stimulus, and flowers were associated with an incongruent response, prompting students to press on the key on the opposite side of the stimulus. The "mixed trials" block, in which students were presented with both hearts and flowers at random, is considered the most cognitively demanding trial. We used the proportion of correct responses (i.e., the number of trials with a correct response divided by the total number of trials) during mixed block as a measure of working memory, cognitive flexibility and inhibitory control. We also used mean reaction time on mixed trials minus mean reaction time on "hearts only" trials (i.e., the easiest trials) as a measure of the effect of increased cognitive demand on basic processing speed.

**Grades.** To adjust for any baseline differences in academic achievement, we included students' self-reported GPA, which was measured just prior to the administration of the PFL intervention. Students were prompted to respond to the question "How would you describe your grades in school?" with "mostly A's," "mostly B's," etc. Due to the low number of students who responded "mostly F's," (0.66%) and "mostly D's" (3.51%), these two values were combined.

Although we hoped to use district-reported GPA for our measures of grades, administrative data were missing for the majority of students. For the 172 students in our sample who had non-missing data for both self-reported GPA and district GPA, these two measures of student grades had a correlation of 0.67 ($p < .001$; see Watts et al., 2018 for full description of efforts to validate self-reported grades against district-reported GPA). This correlation was identical for our variable of GPA in which F's were combined with D's and for the original GPA variable. For the 456 non-missing responses to the self-report of GPA, 43 students (9.43%) responded with "none of these grades" or "I am not sure." Of these 43 students, 15 of them had available district-reported GPA that we imputed for their outcome grades. The remaining 28 students had their responses recoded as missing.

**Trait anxiety.** Stable, trait-like anxiety was collected as an indicator of individual differences in students' capacity to manage feelings of anxiety and worry using the State-Trait Anxiety Inventory for Children – Trait (STAIC-T), administered during a data collection wave of CSRP that occurred four years prior to the PFL intervention wave (Time 6). The STAIC-T consists of a 10-item aggregate on a 3-item Likert scale that measures trait anxiety in children between the ages of 8 and 14 (Spielberger, Edwards, Montuori, & Lushene, 1973).

**Mindset beliefs.** Baseline student mindset was collected through self-report survey items immediately prior to the PFL intervention. All survey items were on a 1 (strongly disagree) to 6 (strongly agree) Likert scale and began with the statement "How much do you agree or disagree with this statement?" Students' level of fixed intelligence was derived from their response to the question "Your intelligence (smartness) is something very basic about you and you can't change it very much." The belonging uncertainty rating was created as a mean aggregate ($\alpha = .63$) of responses to the questions "Sometimes I feel like I belong at school, and sometimes I don't feel like I belong," and "When something bad happens, I feel like maybe I don't belong at school." Teacher trust was a mean aggregate ($\alpha = .80$) of the students' responses to the questions "I am treated fairly by teachers and other adults at my school," "Teachers and other adults at my school treat me with respect," "Teachers at my school care about their students," and "Teachers at my school treat students in my racial/ethnic group fairly."

## Analytic Approach

For both analytic models, we regressed each dependent variable on a dummy variable for the mindset treatment status (which remained the same for each student who completed both interventions) and fixed effects for the students' original Head Start site to account for the unit of randomization (recall that random assignment was tied to students' original Head Start center to ensure no correlation between the pre-K treatment and the mindset interventions).

$$\text{Equation 1: } Outcome_{ij} = a_1 + \beta_1 Tx_{ij} + \sum_{j=1}^{18} \beta_2\, Site_j + e_{ij}$$

In the model above, $\beta_1$ represents the treatment impact, *Outcome* represents the dependent variables specific each to each study (described below) for the *i*th student who attended Head Start center *j*, and *Tx* represents the treatment status dummy indicator (coded "1" for treatment and "0" for control). All models were run with robust standard errors adjusted for Head Start site-level clustering in Stata 15.0.

In addition to this main effects model, we ran models that included covariates for demographic, family, and cognitive skills (see Table 1 for the full list of control variables).

$$\text{Equation 2: } Outcome_{ij} = a_1 + \beta_1 Tx_{ij} + \sum_{j=1}^{18} \beta_2\, Site_j + \chi Demographics_{ij} + \Omega Cognitive_{ij} + e_{ij}$$

For all models, we included students who had had non-missing data on each respective outcome measure. We used multiple imputation to account for all missing data on baseline covariates. For multiple imputation, we generated 10 multiply imputed datasets using the multivariate normal regression procedure in Stata 15.0. The majority of the participants in the sample (56%) had 1 or fewer missing baseline measures, and there was no significant difference between the treatment and control groups on the overall rate of missing data ($p = 0.18$).

## Purpose for Learning Intervention

### Participant Characteristics

Participants of the PFL intervention were 463 adolescents who participated in the original pre-K study (Raver et al., 2009) and the participants' parents consented them to be re-randomly assigned to the PFL intervention or control condition at the 10-year follow-up wave. The majority (68%) of the students identified as Black and 25% identified as Hispanic. Twenty-nine percent of the participants were in middle school and 71% in high school, with the majority (44%) in 9th grade. Further, most students (77%) came from families with income to needs ratios less than 1 (i.e., below the poverty line).

**Procedures**

Participants completed all tasks in a quiet area of their schools (57%) or in their homes (34%) on laptops provided by the research team[3]. Both conditions of PFL had optional audio recording of all text for students with vision impairments or reading difficulties. Headphones were provided for all participants. The intervention lasted 20-40 minutes.

Students in the PFL treatment condition were first asked to write briefly about problems in the world and/or their community that they were interested in solving. They were then presented with information about how students work hard in school because they want to grow up to "make a positive impact on the world," or to be "a good example for other people." Participants were then asked to think about their own goals and to write about how learning and working hard in school could help them achieve these goals.

In the control condition, students were asked to reflect on how their lives have changed between middle school and high school[4] and were presented with scientific information about how the brain learns through classroom assignments. This was followed by short vignettes from other students on the difference between middle and high school. At the end of the control condition, students wrote a letter to a hypothetical incoming middle school student about what has changed in their lives since middle school.

[3] Another 8% of the sample completed the tasks on their own computers at home with continuous guidance from an assessor over the phone. A remaining 1% (4 participants) completed the tasks in a location such as a restaurant/café or a parent's workplace. This flexibility in administration setting allowed us to retain as large a sample size as possible.
[4] The 29% of students who were still in middle school during administration were verbally instructed by assessors to focus on the differences between elementary school and middle school.

**Outcomes**

The short-term outcomes of the PFL intervention included indicators of task persistence (the "College Knowledge Task") and a measure of state anxiety. The long-term outcomes included self-reported overall GPA, self-reported grades in math classes, and a measure of the student's sense of belonging uncertainty in their school.

**College Knowledge Task.** The College Knowledge Task (CKT; Masucci & Raver, 2017) was used to test students' diligence on an applied real-world academic task. The overall structure and administration of the CKT was adapted from the Academic Diligence Task (ADT; Galla et al., 2014). The ADT demonstrates a measurement profile consistent with self-report measures of grit, conscientiousness, and self-control. The ADT was used in previous trials of PFL and mindset interventions to provide a behavioral test of self-regulation by pitting students' desire to achieve their learning goals against the temptation to disengage from relatively boring academic material and play games instead.

The CKT is similar to ADT in many ways, with the primary modification of the learning content, which focused on information about applying to college, rather than on completing math problems. Thus, the CKT attempted to simulate the process of managing and persisting through college applications and financial aid materials. Participants were given 10 minutes to complete the task, split into two "Blocks." In Block 1, students were introduced to the reading material and answered comprehension questions about college applications without any distractors. In Block 2, students were given the choice to switch between playing computer games (such as Tetris or Pac-Man) and engaging with the reading topics and comprehension questions related the college application process. With the modified focus of the CKT, we aimed to decrease the risk of confounding math ability or math anxiety with performance on the ADT, and we hoped to provide students participating in the study with useful information regarding college enrollment.

We examined three outcomes related to the CKT in order to measure the ability to continue on with a tedious task while ignoring the option to play a video game: 1) if the students played any games during Block 2 of the task (1 = played any games, 0 = played no games); 2) the amount of time spent on Block 2; 3) the percentage of time spent on-task. After the PFL intervention, 35% (SD = 48%) of students played at least one game during Block 2, students spent an average of 7.22 minutes (SD = 1.44 minutes) on Block 2, and students spent an average of 89% (SD = 24%) of their Block 2 time on-task (i.e., not playing games).

**State anxiety**. We measured short-term, situational anxiety (called "state anxiety") immediately following the CKT. State anxiety was measured using the State-Trait Anxiety Inventory for Children – State (STAIC-S). The STAIC-S consists of two 20-item aggregates, with 10 survey items on a 3-item Likert scale that measure state anxiety in children between the ages of 8 and 14 (Spielberger et al., 1973). After the PFL intervention, the state anxiety mean across both groups was 1.48 (SD = 0.26), indicating a low average level of anxiety, with an alpha of .86 across the 20 items.

**Grades.** Students' self-reported overall grade average and their average grade in math class were collected one year following PFL intervention, prior to the GM intervention. As with our baseline measure of self-reported GPA (described above), we combined responses to "mostly F's" (1.47%) and "mostly D's" (5.39%) into one category, and we again used district-reported GPA for any students who did not respond to the self-reported GPA survey question but had non-missing district-reported grades ($n$= 8).

**Belonging uncertainty.** One year post-treatment, prior to the start of the GM intervention, students responded to a question regarding their sense of belonging uncertainty in their school. This item was used in previous trials of PFL (Yeager et al., 2014). Students were prompted to respond to the following question on a scale from 1 (strongly disagree) to 6 (strongly agree): “When something bad happens, I feel like maybe I don’t belong at school.”

## Results

**Baseline equivalence.** In Table 1, we present results from tests designed to evaluate whether the random assignment procedure produced groups that were equivalent on observable characteristics measured at or before study baseline. We regressed each baseline characteristic on a dummy variable for treatment status and a set of Head Start site fixed effects (i.e., the unit of random assignment), and adjusted standard errors for site-level clustering. All continuous baseline covariates were sample standardized, so coefficients can be interpreted as effect sizes.

As Table 1 reflects, we found no indication of baseline imbalance across the 17 covariates examined. We further tested whether the set of baseline characteristics differed between the treatment groups by regressing treatment status on the entire set of baseline covariates with site fixed effects again included to adjust for the within-site randomization design of the study. For this model, we employed a joint F-test, which tested whether the set of baseline covariates jointly statistically significantly differed from "0." Indeed, we did not find evidence of imbalance across all the covariates ($F(17, 91) = 0.66$, $p = 0.83$).

These balance tests suggest that no statistically significant differences between participants in the treatment and control conditions could be detected at baseline. Although we detected no statistically significant differences ($p < 0.05$) on any single test of non-equivalence, several observed differences fell between 0.05 and 0.25 standard deviations, suggesting some indication of possible non-equivalence. Thus, we follow recommendations for best practice by beginning with models that include no baseline covariates before turning to models that control for the full set of baseline characteristics (What Works Clearinghouse, 2017). By comparing the treatment impact coefficient across models that do and do not include controls, we can gauge the robustness of the unconditional model to the inclusion of other variables.

**Manipulation check.** Immediately following the PFL intervention, students responded to four manipulation check survey items to determine the short-term effects of the intervention on students' level of meaning making within their academic environment ("meaningfulness of schoolwork" measure taken from Yeager et al., 2014). Participants viewed a picture of an academic object or scenario and were prompted to "choose the description that more naturally comes to your mind when you see the picture." The participants had a choice between a description that aligned with the concrete utility of the object and a description that aligned with personally meaningful goals that would have been activated by the intervention. These four responses were summed, with the correct responses (reflecting the personally-relevant meaning of the picture rather than the concrete meaning) coded as "1." This value was regressed on PFL treatment condition and the full set of controls, including site, child/family demographics, and cognitive skills (see Equation 2). Intervention students had more correct responses than control students ($\beta = 0.27$, $p = 0.001$), suggesting that the intervention was effectively administered.

**Treatment impacts.** The treatment impact results are described in Table 2, with standardized outcomes (with the exception of College Knowledge: Played Any Games, which is a binary outcome [1 = played any games]) so coefficients can be interpreted as effect sizes. Each column of the table presents regression estimates from a different model, with the first column presenting estimates from the model with no baseline covariates (see Equation 1). In the next two columns, we progressively add in groups of baseline measures, starting first with demographic characteristics, before adding cognitive skills and baseline mindset measures. Column 3 of Table 2 contains all control variables (i.e., Equation 2; see full list of controls in Table 1).

For all CKT measures, state anxiety, and belonging uncertainty, our models failed to detect any significant effects of treatment. For our overall measure of student self-reported GPA, the estimated treatment impact was not statistically significant when no covariates were included ($\beta = -0.14$, $p = 0.07$), but when only demographic covariates were added to this model, the estimated treatment impact was negative and statistically significant ($\beta = -0.19$, $p = 0.04$). However, this estimate shrunk and did not reach statistical significance when all covariates were added ($\beta = -0.13$, $p = 0.08$). When only math grades were considered, effects were nearly identical (bivariate model: $\beta = -0.14$, $p = 0.08$; demographic controls: $\beta = -0.19$, $p = 0.04$; fully-controlled model: $\beta = -0.14$, $p = 0.11$).

**Sensitivity checks.** Although the treatment impact estimates on student GPA were largely inconsistent, we conducted sensitivity checks to further investigate what might have driven the counter-intuitive finding. We identified several potential threats to the interpretability of the self-reported grades finding, including the possibility that attrition between the intervention and the follow-up one year later could have affected the results, and the possibility that PFL-treated students took more difficult classes, resulting in lower grades on average. These sensitivity checks and their results are described in more detail in the supplementary file. Table S3 demonstrates that a regression model that included weights for students based on their likelihood of having follow-up data on self-reported grades yielded nearly identical results as shown in Table 2. Furthermore, there was no evidence that treated students were more likely to self-report selecting a more difficult academic trajectory (see Table S4).

**Treatment impact heterogeneity.** We ran models examining treatment impact heterogeneity by student gender, race, high/low grades prior to the treatment, and high/low trait anxiety. These models included the full set of covariates (i.e., Model 3 from Table 2).

Our heterogeneity tests for students with high grades (or students with "A" or "B" averages) prior to intervention and for female students largely indicated that the treatment did not differentially affect these groups. We found that treated Black students played fewer games during CKT than Black students in the control group, whereas Hispanic/Latino students in the treatment group played more games than those in the control group ($\beta = -0.20$, $p = 0.005$). The intervention had a negative impact on grades overall, but highly anxious students (cutoff at the sample mean) in the treatment group had higher math grades relative to their control-group assigned counterparts, while low-anxiety students in the treatment group reported lower grades than their control-group counterparts ($\beta = .32$, $p = 0.08$), although this effect did not reach statistical significance.

## Growth Mindset Intervention

### Participant Characteristics

Participants of the GM intervention were 404 adolescents who participated in the original pre-K study. Of the original 463 participants from the PFL sample, 59 of these students did not return for the Growth Mindset module. Students who returned from the previous year's PFL intervention retained their treatment status. Thus, 211 participants were included in GM treatment condition (and thus received both the Growth Mindset intervention and the PFL intervention) and 193 to the control condition. The demographic characteristics for this sample were nearly identical as the demographic characteristics of the students included in the PFL intervention, and these characteristics are displayed in Appendix Table S1. The majority (78%)

of the students completed the tasks in a quiet area of their schools, 10% of the students completed the tasks in their homes, and 11% over the phone with guidance from an assessor.

**Attrition.** We checked for balance on baseline characteristics and treatment status between the 59 students who dropped from the sample after the PFL intervention and the 404 we retained for the GM intervention. Dropout status was regressed on each one of these characteristics, with site fixed effects and robust standard errors for site-level clustering. We found that PFL treatment status was negatively associated with dropout status ($\beta = -0.08$, $p = .01$), indicating that slightly more students in the control group dropped out of the sample. Additionally, we found that baseline GPA was slightly negatively associated with dropout status ($\beta = 0.04$, $p = .03$), suggesting that more students with lower GPAs at baseline dropped out of the sample. However, despite these differences in the characteristics of the 59 students that left the sample, we still found no differences in observed baseline characteristics between the treatment and control groups for the 404 remaining students (see Table 1).

**Outcomes**

The short-term outcomes for the GM intervention included an abridged version of the CKT, state anxiety, and a measure of critical motivation.

**College Knowledge Task**. See the PFL outcomes section for a full description of this measure. Immediately following the GM intervention, an abbreviated version of the CKT was administered to all participants, which removed the introductory part of the task without any game playing distractors (Block 1) so that they only completed Block 2. The students had 8 minutes to complete this task, and 40% (SD = 49%) of students played at least one game during Block 2, students spent an average of 7.61 minutes (SD = 0.95 minutes) on Block 2, and students spent an average of 82% (SD = 30%) of their Block 2 time on-task.

**State anxiety.** See the PFL outcomes section for a full description of this measure. State anxiety was collected immediately after the CKT using the STAIC-S. Across the sample, the state anxiety mean was 1.52 (SD = 0.27), with an alpha of 0.89 across the 20 items.

**Critical motivation.** Critical motivation was measured following the GM intervention. This construct was operationalized by use of selected items from the Critical Agency subscale of the Measure of Adolescent Critical Consciousness (MACC; McWhirter & McWhirter, 2016), which seeks to measure youths' moral concern with inequity along with their perceived ability and motivation to affect social change. Participants completed a 4-item questionnaire ($\alpha = .83$) that included items like "I am motivated to try to end racism and discrimination" and "I can make a difference in my community," and we averaged student responses across these four items. Items were intended the participants' level of agreement with each statement and were on a scale of 1 ("strongly disagree") to 5 ("strongly agree").

**Procedures**

The procedures were similar to those of the PFL intervention. For this second intervention, the treatment group received the GM intervention module taken from the National Mindset Study (Yeager et al., 2016), and the control condition received educational materials from the "Brainology tool kit" (Dweck, 2008a; Yeager at al., 2013). In the treatment condition, participants were first asked to elicit what issues in the world or in their community matter to them personally. They were then presented with information and vignettes on the "learning mindset," or the belief that the brain and individual can change through practiced skills and knowledge acquisition. The participants then answered another open-ended question to explain how a person could use a learning mindset to strengthen their brain. After more information about the benefits of the learning mindset, the students wrote about how they planned to use a learning mindset in their classes.

In the control condition, students were given factual information about brain science, and how healthy behaviors can improve it. Participants were first asked what they know about the brain and how to keep it healthy. They then read information on what the brain is, how it works, different methods of keeping it healthy (e.g., sleep, nutrition), and the benefits of having a healthy brain. Students then reported how they planned on keeping their brain healthy during the upcoming year.

**Results**

**Baseline equivalence.** Because the sample was nearly identical to the sample present for the PFL intervention, baseline equivalence checks produced nearly identical results as those reported above. As before, we observed small differences on our measures of whether the child came from a single parent and the number of hours the parent reported working. The specific baseline information for the GM intervention sample is presented in supplementary information file Table S1.

**Manipulation check.** Immediately following the GM intervention, students responded to four manipulation check survey items to determine the short-term effects of the intervention on students' level of meaning-making within their academic environment. These four binary responses were summed, with the correct responses coded as "1." Regression results (using full controls from Equation 2) did not result in a statistically significant difference in response between the treatment and control groups. This suggests that the short-term effects of the GM intervention did not include a shift in student' meaning-making of their academic environment.

**Treatment effects.** The regression results are described in Table 4, and follow the same format of presentation as Study 1 (Table 2). We again found null results, as we saw almost no indication that the GM intervention had any effect on students' short term outcomes—CKT game playing and block 2 duration, state anxiety, or critical motivation. Although some of the effects were positive in direction (e.g. critical motivation), no effect was statistically significant.

**Treatment impact heterogeneity.** In addition to calculating treatment effects of the GM intervention, we ran models examining the heterogeneity of treatment effects by student gender, race, high/low grades prior to the treatment, and high/low trait anxiety. As with the PFL intervention, these models included the full set of covariates. For moderation by high grades at pre-test, race (Black), or high trait anxiety at pre-test, we saw no evidence of treatment impact heterogeneity. However, female students in the treatment group had significantly higher task-related state anxiety than female students in the control group ($\beta = 0.48$, $p = .03$), whereas male students had lower state anxiety than their control-assigned counterparts.

## Discussion

A substantial body of literature has demonstrated the promising effects of changes in mindset on students' GPAs (Blackwell et al., 2007; Claro et al., 2016; Farrington et al., 2012), and researchers have made claims that mindset interventions can have remarkable effects on educational achievement at a broader policy level (Rattan, Savani, Chugh, & Dweck, 2015; Yeager & Walton, 2011). These exciting developments have led to increased use of one-time, self-administered online mindset interventions as a way to reach a wide range of diverse student populations, with many demonstrating positive effects (Bettinger, Ludvigsen, Rege, Solli, & Yeager, 2018; Dweck & Yeager, 2018). Furthermore, even research that claims skepticism about the strength of these effects on the general student population concedes that academically high-risk and economically disadvantaged students could stand to benefit from mindset interventions

(Sisk et al., 2018). This research sought to conceptually replicate these previous findings by targeting economically disadvantaged students of color for intervention, and then aimed to extend these findings to other potential behavioral outcomes. Surprisingly, we found mostly null effects of the mindset interventions on student outcomes, and we found a small negative effect of PFL intervention alone on student grades. Analyses of treatment impact moderation suggested that certain student characteristics could play a role, but most of these tests also presented null results.

**PFL Main Effects**

Our results demonstrated that overall, the PFL intervention appeared to have little effect on hypothesized student outcomes. Although these results were surprising given earlier studies of interventions similar to the ones tested here that showed promising results (e.g., Yeager & Walton, 2011), other recent studies of various "wise" interventions have also reported null findings (Broda et al., 2018; Li & Bates, 2017; Sisk et al., 2018). In our study, we found no indication that the mindset interventions changed putative psychological processes that have been argued to play an important role in supporting or eroding students' chances of academic success. We found no differences between the treatment and control groups on key measures of academic and emotional self-regulation (task persistence, feelings of anxiety), nor on measures of belonging uncertainty or critical motivation. In short, we selected a broad set of outcomes that had each been hypothesized to connect directly to the processes targeted by the two interventions tested here, and found no evidence of either intervention's effect on that wide range of outcomes.

For our measure of overall self-reported grades, we also found surprising results. When only demographic covariates were included, we found that the PFL intervention had a significant negative effect on GPA. However, this effect was not statistically significant when all baseline covariate were added to the model. When only grades in math courses were considered, effects were similar in both direction and magnitude. Because we relied on self-reported GPA, these results could suggest actual drops in performance, or they could be indicative of changes in students' perceptions of their own academic abilities. Nevertheless, this primarily null effect was unexpected given previous evidence of mindset interventions eliciting a positive effect on low-income students' grades, and warrants further consideration.

## PFL Heterogeneity

Analyses of heterogeneity of treatment effects by trait anxiety, baseline grades, race, and gender yielded primarily null and inconsistently significant results. Our heterogeneity analyses found moderation by student race on one measure of students' academic diligence, but not others. For example, when considering whether students played any games during the CKT, our results demonstrated that Black students in the treatment group performed greater task diligence than did Black students in the control group (with no differences found among other racial/ethnic groups). However, we did not observe any evidence of moderation by race for the percentage of time spent on task or the total duration of time spent on Block 2.

Additionally, students with high trait anxiety in the PFL treatment group reported performing better (with grades that were 0.32 standard deviations higher) in math than did students with high anxiety in the control group. In turn, our results also indicate that students with low anxiety in the treatment group had lower math grades than low-anxiety control students. However, it is important to note that this small, non-significant effect was only found for math grades and not overall grades, suggesting that this finding is preliminary.

## Post-GM Main Effects

All outcomes measured post-GM yielded no significant results from any of our analytic models. It is important to note that, because the students who underwent the GM intervention were the same sample as the students in the PFL treatment condition, our outcomes represent an opportunity to test the cumulative effect of both studies, and not solely the direct effects of the GM intervention. Moreover, our manipulation check suggested that the students did not immediately respond to the intervention with a shift in meaning making of academic images. This is not surprising, given the cumulative nature of one-time growth mindset interventions that result in effects emerging several months post-intervention. However, it is interesting that this kind of meaning making of academic images did not carry over from the prior year's intervention (recall that the manipulation check was significant post-PFL).

What might explain these overall null findings for PFL and for both interventions tested in combination? The mode of implementation for the intervention could have played a role, as the adolescents participated in the intervention by interacting independently with a computer-based module either in a quiet space in their school or home. Previous studies had administered similar one-time mindset interventions in classrooms with teacher involvement (DeBacker et al., 2018; Paunesku et al., 2015), suggesting that even highly scalable interventions may still require the classroom context. In fact, research suggests that teachers who facilitate mindset interventions may reinforce or elaborate on mindset messages during daily instruction (Schmidt, Shumow, & Kackar-Cam, 2015). Moreover, the meaning of an intervention is likely to shift in different contexts, such that academic values may not be immediately salient to participants who engage with the materials in their homes (Yeager & Walton, 2011).

In addition to the novel mode of implementation, this study was conducted within the highly specific context of low-income, high-violence areas of Chicago. Low-income students of color often experience despair and hopelessness as a result of existing within a "limited opportunity structure" (Fordham & Ogbu, 1986). Previous research suggests that long-term goals are more motivating when students think that the outcome (in this case, college attendance) is personally attainable (Destin & Oyserman, 2009; Dweck et al., 2011). Although the CKT that immediately followed both interventions made no mention of the cost of attending college and only emphasized concrete steps students could take to access financial aid, we cannot be certain how these particular students may react to the mention of paying for college. The brief messages offered in the CKT may not be sufficient to counteract the realities of the high cost of college and students' often limited economic resources in trying to meet those costs. In this way, the combination of activating a higher purpose for learning or a growth mindset in students in the treatment condition without a long-term plan for attaining these goals may have left this group of students marginally more frustrated and academically disengaged.

**Post-GM Heterogeneity**

Heterogeneity analyses post-GM showed that female students in the treatment group experienced higher state anxiety immediately following the intervention than female students in the control group. This effect was not found following PFL treatment, or for any of the other outcomes tested after the GM intervention, so any interpretations should be made with caution. It remains possible that female students in the treatment group felt the burden of needing to overcome an entity mindset (or the belief that ability is innate) more strongly than male students, particularly given gender stereotypes about academic ability. But, more research should investigate this possible mechanism further before strong conclusions are drawn.

**Limitations**

It should be noted that many of our results were imprecisely estimated due to the sample size of the block-randomized design available for our analyses. Consequently, for many of our results, confidence intervals spanned ranges that would cover substantively important effects. Yet, this imprecision should be balanced against the point estimates generated across multiple outcomes tested, as we consistently estimated effects close to 0 for a range of outcomes that spanned multiple unique psychological processes. Additionally, our models show that our actual precision was better than the power analyses predicted, given that we rejected the null hypothesis at an effect size of around 0.19 on the outcome of grades post-PFL intervention.

Another potential limitation of our study is the existing relationship between our research team and the students. Mindset interventions should be sufficiently "stealthy" in order to minimize resistance to the messaging and maximize their potential impacts (Yeager & Walton, 2011). Moreover, adolescents are known to demonstrate resistance to interventions in which they perceive that adults are attempting to change their behavior (Yeager, Dahl, & Dweck, 2018). The students in this study were aware of their participation in the CSRP research study and both intervention visits included several survey batteries that made it clear that we were collecting data. To counteract this, the two interventions were presented to the participating students as an opportunity to share knowledge with other students and as a learning opportunity about the brain – not as an attempt to change the students' thinking. Despite these efforts to downplay the "treatment" oriented nature of these experiments, this kind of researcher-participant relationship with students is unusual in mindset research, leaving us open to the possibility that students may be demonstrating reactivity to the task.

**Conclusion**

This study is among a recent wave of growth mindset studies that have also found null results on measures of academic achievement (Broda et al., 2018; Li & Bates, 2017; Sisk et al., 2018). Other null studies have suggested that a variety of implementation nuances could be to blame, including attempts to customize the intervention for the population, the setting of the intervention, and the mode of administration. However, even studies that yield overall null results frequently find positive effects for subsamples of academically high-risk and economically disadvantaged students (Broda et al., 2018; Sisk et al., 2018), making our results more puzzling. Like these other studies, our results raise questions regarding the context specificity of these effects and the broader efficacy claims of light-touch, one-time interventions like the mindset interventions tested here: can these interventions really be self-administered by any students anywhere, or do we need to ensure that teachers are leading these interventions in a well-managed classroom? This study sheds light on the need for continued replication trials of mindset interventions on specific different populations of students in a variety of contexts.

## References

Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F., & Yeager, D. (2018). Increasing perseverance in math: Evidence from a field experiment in Norway. *Journal of Economic Behavior & Organization*, *146*, 1-15.

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.

Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing inequality in academic success for incoming college students: A randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness*, 1-22.

Burnette, J. L., Russell, M. V., Hoyt, C. L., Orvidas, K., & Widman, L. (2018). An online growth mindset intervention in a sample of rural adolescent girls. *British Journal of Educational Psychology*, *88*(3), 428-445.

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences*, *113*(31), 8664-8668.

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *science*, *324*(5925), 400-403.

DeBacker, T. K., Heddy, B. C., Kershen, J. L., Crowson, H. M., Looney, K., & Goldman, J. A. (2018). Effects of a one-shot growth mindset intervention on beliefs about intelligence and achievement goals. *Educational Psychology*, 1-23.

Destin, M., & Oyserman, D. (2009). From assets to school outcomes: How finances shape children's perceived possibilities and intentions. *Psychological Science*, *20*(4), 414-418.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science (New York, NY)*, *318*(5855), 1387.

Diemer, M. A., Rapa, L. J., Voight, A. M., & McWhirter, E. H. (2016). Critical consciousness: A developmental approach to addressing marginalization and oppression. *Child Development Perspectives*, *10*(4), 216-221.

Dweck, C. S. (2008a, Winter). Brainology: Transforming students' motivation to learn. Independent School. Retrieved from https://www.nais.org/magazine/independent-school/winter-2008/brainology/

Dweck, C. S. (2008b). *Mindset: The new psychology of success*. Random House Digital, Inc.

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

Dweck, C. S. & Yeager, D. S. (2019). Mindsets: A View from Two Eras. *Perspectives on Psychological Science,* 1-16.

Easton, J. Q., Johnson, E., & Sartain, L. (2017). The predictive power of ninth-grade GPA. Chicago, IL: University of Chicago Consortium on School Research.

Farrington, C. A. (2013). Academic mindsets as a critical component of deeper learning. *University of Chicago: Consortium on Chicago School Research*.

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--A Critical Literature Review*. Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white'". *The urban review*, *18*(3), 176-206.

Galla, B. M., Plummer, B. D., White, R. E., Meketon, D., D'Mello, S. K., & Duckworth, A. L. (2014). The Academic Diligence Task (ADT): Assessing individual differences in effort on tedious but important schoolwork. *Contemporary educational psychology*, *39*(4), 314-325.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology*, *24*(6), 645-662.

Gormley Jr, W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology*, *41*(6), 872.

Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in context: School moderation of the impact of social identity threat on racial/ethnic achievement gaps. *Sociology of Education*, *87*(2), 106-124.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 33-46.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172-177.

Hill, P. L., Burrow, A. L., & Sumner, R. (2013). Addressing important questions in the field of adolescent purpose. *Child Development Perspectives*, *7*(4), 232-236.

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*, *102*(4), 880.

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.

Jones, S. M., Brown, J. L., & Lawrence Aber, J. (2011). Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. *Child Development*, *82*(2), 533-554.

Li, Y., & Bates, T. C. (2017). Does mindset affect children’s ability, school achievement, or response to challenge? Three failures to replicate. *Unpublished manuscript*.

Masucci, M.D. & Raver, C.C. (March, 2017). The “College-Knowledge” Task as a novel behavioral measure of academic diligence. Poster presented at the Annual Meeting of the Society for Research on Educational Effectiveness, Washington, DC.

McWhirter, E. H., & McWhirter, B. T. (2016). Critical consciousness and vocational development among Latina/o high school youth: Initial development and testing of a measure. Journal of Career Assessment, 24(3), 543-558.

Miu, A. S., Yeager, D. S., Sherman, D., Pennebaker, J., & Trzesniewski, K. (2014). Preventing Depression by Teaching Adolescents That People Can Change: Nine-Month Effects of a Brief Incremental Theory of Personality Intervention. Clinical Psychological Science, 3(5), 826-743.

Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, *95*(2), 255.

Nagaoka, J., Farrington, C. A., Roderick, M., Allensworth, E., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2013). Readiness for college: The role of noncognitive factors and context. *Voices in Urban Education*, *38*, 45-52.

Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*(6), 1314.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D., & Gross, J. J. (2004). For better or for worse: neural systems supporting the cognitive down-and up-regulation of negative emotion. Neuroimage, 23(2), 483-499.

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.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, *26*(6), 784-793.

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, *331*(6014), 211-213.

Rattan, A., Savani, K., Chugh, D., & Dweck, C. S. (2015). Leveraging mindsets to promote academic achievement: Policy recommendations. *Perspectives on Psychological Science*, *10*(6), 721-726.

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., Bloom, H., & Hill, C. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01)[Software]

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early childhood research quarterly*, *23*(1), 10-26.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: self-regulation as a mediating mechanism. *Child development*, *82*(1), 362-378.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of consulting and clinical psychology*, *77*(2), 302.

Schmidt, J. A., Shumow, L., & Kackar-Cam, H. (2015). Exploring teacher effects for mindset intervention outcomes in seventh-grade science classes. *Middle Grades Research Journal*, *10*(2), 17-32.

Sherman, D. K., & Hartson, K. A. (2011). Reconciling self-protection with self-improvement. *Handbook of self-enhancement and self-protection*, *128*.

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological science*, *29*(4), 549-571.

Spielberger, C.D., Edwards, C.D., Montuori, J., & Lushene, R. (1973). State-Trait Anxiety Inventory for Children. Published by Mind Garden, Inc., www.mindgarden.com.

Spitzer, B., & Aronson, J. (2015). Minding and mending the gap: Social psychological interventions to reduce educational disparities. *British Journal of Educational Psychology*, *85*(1), 1-18.

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*(1), 73-82.

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.

Watts, T. W., Gandhi, J., Ibrahim, D. A., Masucci, M. D., & Raver, C. C. (2018). The Chicago School Readiness Project: Examining the long-term impacts of an early childhood intervention. *PloS one*, *13*(7), e0200144.

What Works Clearinghouse (2017). Standards handbook (version 4.0). Washington, DC: Institute of Education Sciences.

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.

Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, *13*(1), 101-122.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational psychologist*, *47*(4), 302-314.

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of personality and social psychology*, *107*(4), 559.

Yeager, D. S., Lee, H. Y., & Jamieson, J. P. (2016). How to improve adolescent stress responses: Insights from integrating implicit theories of personality and biopsychosocial models. *Psychological science*, *27*(8), 1078-1091.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., Greene, D., Walton, G. M., & Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of educational psychology*, *108*(3), 374.

Yeager, D. S., Trzesniewski, K. H., & Dweck, C. S. (2013). An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion. *Child development*, *84*(3), 970-988.

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

*Figure 1. Data Collection Timeline*

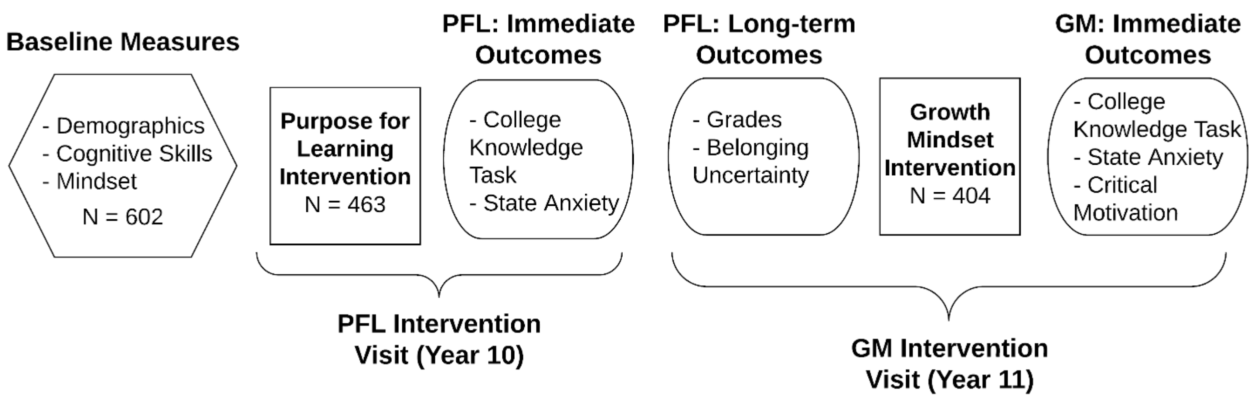
Baseline Measures
- Demographics
- Cognitive Skills
- Mindset
N = 602
Purpose for Learning Intervention
N = 463
PFL: Immediate Outcomes
- College Knowledge Task
- State Anxiety
PFL: Long-term Outcomes
- Grades
- Belonging Uncertainty
Growth Mindset Intervention
N = 404
GM: Immediate Outcomes
- College Knowledge Task
- State Anxiety
- Critical Motivation
PFL Intervention Visit (Year 10)
GM Intervention Visit (Year 11)

*Table 1. Descriptive Means of Control Variables by PFL Treatment Condition*

| | Treatment | | Control | | $p$-value | |
|---|---|---|---|---|---|---|
| | Mean/% | SD | Mean/% | SD | | N |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Demographics** | | | | | | |
| Female (Y1) | 0.55 | 0.50 | 0.54 | 0.50 | 0.972 | 463 |
| Age (Y1) | 4.96 | 0.77 | 4.89 | 0.77 | 0.548 | 463 |
| African American/Black (Y1) | 0.65 | 0.48 | 0.71 | 0.45 | 0.118 | 463 |
| Income to Needs Ratio (Y10) | 0.72 | 0.63 | 0.67 | 0.53 | 0.338 | 463 |
| Number of Children in Home (Y10) | 2.61 | 1.40 | 2.73 | 1.29 | 0.220 | 340 |
| Parent Age (Y1) | 29.41 | 7.72 | 29.97 | 7.79 | 0.592 | 443 |
| Parent African American/Black (Y1) | 0.74 | 0.44 | 0.75 | 0.43 | 0.893 | 389 |
| Mother Graduated H.S. (Y1) | 0.35 | 0.48 | 0.40 | 0.49 | 0.391 | 463 |
| Single Parent (Y10) | 0.64 | 0.48 | 0.75 | 0.44 | 0.053 | 340 |
| Parent Hours of Work (Y10) | 37.30 | 12.56 | 34.42 | 13.47 | 0.050 | 249 |
| **Cognitive Skills** | | | | | | |
| H & F Mixed Trials Accuracy (Y10) | 0.65 | 0.19 | 0.66 | 0.20 | 0.850 | 454 |
| H & F Reaction Time (ms) (Y10) | 186.43 | 95.14 | 183.36 | 107.60 | 0.714 | 453 |
| Grades (GPA) (Y10) | 2.78 | 0.82 | 2.79 | 0.85 | 0.650 | 428 |
| Trait Anxiety (Y6) | 1.87 | 0.37 | 1.83 | 0.35 | 0.457 | 321 |
| **Mindset** | | | | | | |
| Fixed Intelligence (Y10) | 3.25 | 1.65 | 3.33 | 1.65 | 0.670 | 462 |
| Belonging Uncertainty (Y10) | 3.16 | 1.22 | 3.17 | 1.22 | 0.976 | 461 |
| Teacher Trust (Y10) | 4.85 | 0.88 | 4.86 | 0.92 | 0.923 | 462 |
| Observations | 232 | | 231 | | | |

*Note*. Year of data collection from the start of the original study is in parenthesis. Y1 refers to Year 1, or the first year of data collection, when students were in preschool. Y10 variables were collected in Year 10, just prior to the administration of the PFL intervention. P-values of baseline mean differences were generated from a series of regressions in which each respective baseline characteristic was regressed on PFL treatment status and the set of Head Start site fixed effects.

**Error! Not a valid link.**

*Table 2. Treatment Effect of the PFL Intervention on Psychological Processes*

| | Unconditional | Demographics | Cognitive Skills |
|---|---|---|---|
| | 1 | 2 | 3 |

| | | | |
|---|---|---|---|
| *Short-term Outcomes* | | | |
| College Knowledge: Played Any Games | -0.007 | 0.005 | -0.002 |
| *n=445* | (0.052) | (0.050) | (0.049) |
| College Knowledge Block 2 Duration | -0.110 | -0.137 | -0.122 |
| *n=445* | (0.106) | (0.112) | (0.110) |
| College Knowledge: Time on Task | 0.040 | 0.031 | 0.048 |
| *n=445* | (0.110) | (0.109) | (0.104) |
| State Anxiety | 0.143 | 0.141 | 0.132 |
| *n=448* | (0.105) | (0.111) | (0.109) |
| *Long-term Outcomes* | | | |
| Grades | -0.143 | -0.191* | -0.131 |
| *n=375* | (0.069) | (0.082) | (0.069) |
| Math Grades | -0.142 | -0.187* | -0.136 |
| *n=393* | (0.076) | (0.082) | (0.079) |
| Belonging Uncertainty | 0.045 | 0.059 | 0.049 |
| *n=402* | (0.130) | (0.133) | (0.138) |
| *Baseline Covariates Included* | | | |
| Head Start Site | Inc. | Inc. | Inc. |
| Demographics | | Inc. | Inc. |
| Cognitive Skills/Mindset | | | Inc. |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. The coefficients displayed in this table are standardized, representing effect sizes. Short-term outcomes were collected immediately after the intervention and long-term outcomes were collected one year later, prior to the administration of the GM intervention. For all regression models, we used multiple imputation with 10 multiply imputed datasets to adjust for missing data on covariates. Each estimate was generated from a separate regression model. Estimates in column 1 were generated from models that only contained the treatment status and Head Start site fixed effects. In column 2, baseline characteristics taken from study students and families were included, and in column 3, the students' academic performance, cognitive skills, and mindset were included.
*p < .05.

*Table 3. Moderation of PFL Treatment Impact by Relevant Student Characteristics*

| | Treatment X High Grades | Treatment X Female | Treatment X Black | Treatment X High Trait Anxiety |
|---|---|---|---|---|
| *Short-term Outcomes* | | | | |
| College Knowledge: Played Any Games | 0.082 | -0.200 | -0.198** | 0.074 |
| | (0.086) | (0.098) | (0.061) | (0.085) |
| College Knowledge Block 2 Duration | 0.119 | -0.144 | 0.037 | 0.085 |
| | (0.185) | (0.203) | (0.222) | (0.204) |
| College Knowledge: Time on Task | -0.248 | 0.260 | 0.096 | -0.217 |
| | (0.267) | (0.177) | (0.165) | (0.210) |
| State Anxiety | 0.056 | 0.091 | 0.089 | 0.043 |
| | (0.221) | (0.183) | (0.284) | (0.177) |
| *Long-term Outcomes* | | | | |
| Grades | -0.172 | 0.115 | 0.089 | 0.190 |
| | (0.171) | (0.160) | (0.160) | (0.186) |
| Math Grades | 0.059 | -0.055 | 0.118 | 0.318 |
| | (0.180) | (0.236) | (0.232) | (0.162) |
| Belonging Uncertainty | 0.201 | 0.223 | -0.162 | 0.091 |
| | (0.181) | (0.200) | (0.191) | (0.181) |
| *Baseline Covariates Included* | | | | |
| Head Start Site | Inc. | Inc. | Inc. | Inc. |
| Demographics | Inc. | Inc. | Inc. | Inc. |
| Cognitive Skills/Mindset | Inc. | Inc. | Inc. | Inc. |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. The coefficients displayed in this table are standardized, representing effect sizes. Short-term outcomes were collected immediately after the intervention and long-term outcomes were collected one year later, prior to the administration of the GM intervention. For all regression models, we used multiple imputation with 10 multiply imputed datasets to adjust for missing data on covariates. All regression models were calculated separately and the full set of covariates was used in each model. High grades and high anxiety were determined by cutoff at the sample mean. N's varied by outcome and moderator. **p<.01.

*Table 4. Treatment Effect of the GM Intervention on Psychological Processes*

| | Unconditional<br>1 | Demographics<br>2 | Cognitive Skills<br>3 |
|---|---|---|---|
| *Short-term Outcomes* | | | |
| College Knowledge: Played Any Games | -0.058 | -0.016 | -0.049 |
| *n=386* | (0.115) | (0.103) | (0.105) |
| College Knowledge: Time on Task | -0.095 | -0.124 | -0.105 |
| *n=386* | (0.108) | (0.100) | (0.100) |
| College Knowledge: Block 2 Duration | 0.117 | 0.115 | 0.120 |
| *n=386* | (0.122) | (0.120) | (0.123) |
| State Anxiety | 0.041 | 0.055 | 0.044 |
| *n=382* | (0.102) | (0.102) | (0.099) |
| Critical Motivation | 0.059 | 0.035 | 0.057 |
| *n=393* | (0.092) | (0.087) | (0.095) |
| *Baseline Covariates Included* | | | |
| Head Start Site | Inc. | Inc. | Inc. |
| Demographics | | Inc. | Inc. |
| Cognitive Skills/Mindset | | | Inc. |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. The coefficients displayed in this table are standardized, representing effect sizes. Only short-term outcomes were collected, immediately after implementation of the GM intervention. For all regression models, we used multiple imputation with 10 multiply imputed datasets to adjust for missing data on covariates. Each estimate was generated from a separate regression model. Estimates in column 1 were generated from models that only contained the treatment status and Head Start site fixed effects. In column 2, baseline characteristics taken from study students and families were included, and in column 3, the students' academic performance, cognitive skills, and mindset were included. No effects were statistically significant.

*Table 5. Moderation of GM Treatment Impact by Relevant Student Characteristics*

| | Treatment X High Grades | Treatment X Female | Treatment X Black | Treatment X High Trait Anxiety |
|---|---|---|---|---|
| *Short-term Outcomes* | | | | |
| College Knowledge: Played Any Games | 0.086 | -0.135 | -0.175 | -0.030 |
| | (0.212) | (0.241) | (0.257) | (0.234) |
| College Knowledge: Time on Task | 0.016 | 0.209 | 0.136 | -0.100 |
| | (0.197) | (0.266) | (0.226) | (0.195) |
| College Knowledge Block 2 Duration | -0.226 | 0.074 | -0.023 | 0.056 |
| | (0.194) | (0.242) | (0.210) | (0.204) |
| State Anxiety | 0.137 | 0.481* | 0.024 | 0.048 |
| | (0.193) | (0.193) | (0.196) | (0.244) |
| Critical Motivation | 0.115 | -0.043 | -0.295 | 0.110 |
| | (0.153) | (0.157) | (0.171) | (0.194) |
| *Baseline Covariates Included* | | | | |
| Head Start Site | Inc. | Inc. | Inc. | Inc. |
| Demographics | Inc. | Inc. | Inc. | Inc. |
| Cognitive Skills/Mindset | Inc. | Inc. | Inc. | Inc. |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. The coefficients displayed in this table are standardized, representing effect sizes. Only short-term outcomes were collected, immediately after implementation of the GM intervention. For all regression models, we used multiple imputation with 10 multiply imputed datasets to adjust for missing data on covariates. All regression models were calculated separately and the full set of covariates was used in each model. High grades and high anxiety were determined by cutoff at the sample mean. N's varied by outcome and moderator.

*p < .05.

**Supplementary Information for:**

**The Effects of Two Mindset Interventions on Low-Income Students' Academic and Psychological Outcomes**

## Intervention Modifications

### Purpose for Learning Intervention

The original intervention (Yeager et al., 2014) was tailored to Alief Middle School in Houston, Texas. The present version was distributed to students across multiple states in either high school or middle school. To accommodate these differences, references to Alief Middle School were removed and replaced with references to “your school.” General references to middle school students in the vignettes were replaced with “middle and high school students.” The open-ended question text in the control condition remained unchanged and asked students to explore the differences between middle and high school. To account for this, participants in middle school were verbally instructed by the assessor to reflect on the differences between middle school and elementary school instead. General references to the institution administering the intervention (originally Stanford University and University of Texas at Austin) were replaced with references to New York University and updated with appropriate graphics. However, the vignettes featuring specific researchers from UT Austin and Stanford remained unchanged in this regard. Additionally, retail careers were added as an example to the vignettes exploring the necessity of academic skills to career success. This was done to localize the intervention and provide an example more relatable to students in major urban centers (such as Chicago, where most students were assessed), as retail careers are commonly pursued by students following high school in these areas. All audio was re-recorded by our staff to reflect changes to intervention text.

### Growth Mindset Intervention

*Changes to control condition.* In the original administration (Yeager et al., 2016), the control condition taught basic information on neuroanatomy and how the different areas of the brain interact to promote learning and behavior. Open-ended questions in the control condition asked students to explain how different brain areas impact their daily lives. The CSRP control condition was adapted from modules of the Skills for Success Brain Toolkit (Yeager et al., 2013), providing information on neuroanatomy and focusing on how health-promoting behaviors (e.g. eating well, exercising, and getting an appropriate amount of sleep) can keep the brain healthy. Differences between the questions asked of participants are presented below. All audio in the GM control condition was recorded by our staff to accurately reflect the condition’s text.

| Control Condition Question | Yeager et al., 2016 | Present Study |
|---|---|---|
| Initial prompt | How can the frontal lobe be useful at home in family life, or with friends outside of school? | What do you know about the brain? How do you keep your brain healthy? |
| Second prompt | What lobes do you need when you use a phone? | N/A |
| Final prompt | What role will each part of your brain (the occipital, frontal, parietal, and temporal lobes) play in your life | In the box below, please tell us:  How are you going to keep your brain healthy this year? |

| | | |
|---|---|---|
| | outside of school over the upcoming week? | |

*Changes to intervention condition.* General references to "high school" and "high school students" were replaced with "school" and "students" to make the intervention appropriate for the present study's middle school participants. Vignette specifics and open-ended question were not changed in this regard. Two vignettes referencing "Michelle Obama, First Lady of the United States were updated to "Michelle Obama, former First Lady of the United States." Audio was only re-recorded by our staff for pages in which text had been changed; the original audio recordings were presented for all other pages.

## Method

**Baseline Equivalence: Growth Mindset.** Between the PFL intervention year and the GM intervention year, 59 students attrited from the sample. As described in the main text, there were a few differences between the attrited sample and those who return, including by treatment status and baseline GPA. Thus, we examined baseline equivalence once more for the 404 who returned for the GM intervention year.

As Table S1 reflects, we found treatment imbalance for one of the same characteristics that demonstrated imbalance prior to the PFL intervention. Adolescents randomly assigned to the treatment condition were 12 percentage points less likely to have a single parent ($p = 0.03$). No other baseline characteristics demonstrated statistically significant imbalance.

We further tested whether the set of baseline characteristics differed between the treatment and control groups by regressing treatment status on the entire set of baseline covariates with site fixed effects again included to adjust for the within-site randomization design of the study. For this model, we employed a joint F-test, which tested whether the set of baseline covariates jointly statistically significantly differed from "0." Indeed, we did not find evidence of imbalance across all the covariates ($F(17, 96) = 0.82$, $p = 0.66$).

*Table S1. Descriptive Means of Control Variables by GM Treatment Condition*

| | Treatment | | Control | | *p*-value | |
|---|---|---|---|---|---|---|
| | Mean/% | SD | Mean/% | SD | | N |
| **Demographics** | | | | | | |
| Female (Y1) | 0.56 | 0.50 | 0.55 | 0.50 | 0.779 | 404 |
| Age (Y1) | 4.96 | 0.77 | 4.89 | 0.78 | 0.606 | 404 |
| African American/Black (Y1) | 0.65 | 0.48 | 0.71 | 0.46 | 0.070 | 404 |
| Income to Needs Ratio (Y10) | 0.73 | 0.64 | 0.66 | 0.50 | 0.278 | 404 |
| Number of Children in Home (Y10) | 2.64 | 1.42 | 2.74 | 1.32 | 0.435 | 295 |
| Parent Age (Y1) | 29.74 | 7.87 | 30.00 | 8.03 | 0.887 | 386 |
| Parent African American/Black (Y1) | 0.68 | 0.47 | 0.72 | 0.45 | 0.314 | 404 |
| Mother Graduated H.S. (Y1) | 0.34 | 0.48 | 0.38 | 0.49 | 0.535 | 404 |
| Single Parent (Y10) | 0.63 | 0.49 | 0.75 | 0.43 | 0.030* | 295 |
| Parent Hours of Work (Y10) | 37.37 | 13.01 | 33.97 | 13.48 | 0.070 | 215 |
| **Cognitive Skills** | | | | | | |
| H & F Mixed Trials Accuracy (Y10) | 0.65 | 0.19 | 0.66 | 0.21 | 0.638 | 396 |
| H & F Reaction Time (ms) (Y10) | 186.20 | 95.49 | 183.86 | 105.04 | 0.815 | 396 |
| Grades (GPA) (Y10) | 2.79 | 0.84 | 2.83 | 0.83 | 0.346 | 372 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Trait Anxiety (Y6) | 1.86 | 0.37 | 1.82 | 0.35 | 0.443 | 284 |
| **Mindset** | | | | | | |
| Fixed Intelligence (Y10) | 3.22 | 1.66 | 3.34 | 1.61 | 0.561 | 403 |
| Belonging Uncertainty (Y10) | 3.17 | 1.24 | 3.11 | 1.22 | 0.497 | 402 |
| Teacher Trust (Y10) | 4.86 | 0.86 | 4.86 | 0.93 | 0.934 | 403 |
| Observations | 211 | | 193 | | | |

*Note*. Year of data collection from the start of the original study is in parenthesis. Y1 refers to Year 1, or the first year of data collection, when students were in preschool. Y10 variables were collected in Year 10, just prior to the administration of the PFL intervention. P-values of baseline mean differences were generated from a series of regressions in which each respective baseline characteristic was regressed on GM treatment status and the set of Head Start site fixed effects.
*p < .05.

## Additional Results

### Sensitivity Checks

**PFL Grades.** As detailed in the main text, for our overall measure of student GPA, we found a negative but statistically non-significant treatment impact when no covariates were included (β = -0.14, p = 0.07), but when only demographic covariates were added to the model, the estimated treatment impact was negative and statistically significant (β = -0.19, *p* = 0.04). However, this estimate shrunk slightly and did not reach statistical significance when all covariates were added (β = -0.13, p = 0.08). We identified several potential threats to the interpretability of our self-reported grades finding. First, attrition could have affected our results if different types of students left the treatment and control groups between the administration of the PFL intervention and the measurement of self-reported grades one year later. Indeed, 19% of participating students did not report their grades or responded with “I am not sure” at the follow-up wave. Those in the PFL treatment group had a higher rate of grade reporting the following year than those in the control group (β = 0.10, p = 0.01).

To examine if this might have affected our treatment impact estimate, we ran models that weighted students based on their likelihood of having follow-up data on self-reported grades. First, we created weights by regressing a binary indicator of having non-missing *self-reported* grades (“1” = non-missing) on treatment status, site fixed effects, and the full list of baseline control variables using a probit model. Thus, in this measure of grades, we did not use the variable that substituted administrative grade data for students who responded “I don’t know” or “none of these” to this survey question. Importantly, we also included interactions between treatment status and baseline covariates to test if the characteristics of youth who left the sample differed between the treatment and control groups. In the supplementary information file, we present results from the linear probability version of this model (Table S2), and we found no statistically significant interactions between baseline characteristics and treatment status, indicating that differential attrition was not likely to be a source of substantial bias. Further, we then used the probit model to generate weights equal to the predicted likelihood of having non-missing self-reported grade data, and we re-ran our key treatment impact model with weights equal to the inverse probability of remaining in the sample. In these models, we used a missing dummy variable approach to address missing data on covariates. With this method, we imputed the mean value of each variable for any observations with missing cases, and added a dummy variable indicating whether a variable had been missing for that observation to the regression model. As Table S3 reflects, the results from this model were nearly identical to the results shown in Table 2, again indicating that differential attrition was not likely to have biased our results.

*Table S2. Predicted Value of Reporting Grades (Non-Missing on Grades)*

| | Interaction Weights |
|---|---|
| PFL Condition | 0.011 |
| | (0.096) |
| Female | 0.003 |
| | (0.055) |
| Age | -0.009 |
| | (0.033) |
| African American | -0.017 |
| | (0.114) |
| Income to Needs Ratio | -0.026 |
| | (0.032) |
| # Children in HH | 0.002 |
| | (0.034) |
| Parent Age | 0.029 |
| | (0.028) |
| Parent African American | -0.180 |
| | (0.120) |
| Mother Graduated H.S. | -0.013 |
| | (0.056) |
| Single Parent | 0.075 |
| | (0.065) |
| Hours of Work | -0.007 |
| | (0.034) |
| PFL*Female | 0.097 |
| | (0.077) |
| PFL*Age | 0.001 |
| | (0.044) |
| PFL*African American | 0.003 |
| | (0.149) |
| PFL*INR | 0.056 |
| | (0.040) |
| PFL*Number children in HH | 0.049 |
| | (0.046) |
| PFL*Parent Age | 0.006 |
| | (0.040) |
| PFL*African American parent | 0.155 |
| | (0.151) |
| PFL*Graduated HS | -0.014 |
| | (0.080) |
| PFL*Single parent | -0.103 |
| | (0.079) |
| PFL*Hours Work | -0.000 |
| | (0.053) |
| *Additional Controls* | |
| Head Start Site Fixed Effects | Inc. |
| Observations | 463 |
| R-squared | 0.083 |

*Note.* Standard errors in parentheses. Estimates here represent "weights" created by regressing a binary indicator of having non-missing self-reported grades on treatment status, site fixed effects, demographic control variables, and interactions between PFL treatment status and demographic variables.

*Table S3. Weighted Regression of Grades on PFL*

| | (1) Unweighted Regression (Original Findings) | (2) Interaction Weighted Regression |
|---|---|---|
| PFL Condition | -0.174* | -0.172 |
| | (0.076) | (0.094) |
| Female | 0.297*** | 0.313** |
| | (0.058) | (0.095) |
| Age | -0.003 | -0.027 |
| | (0.068) | (0.063) |
| African American | 0.083 | 0.052 |
| | (0.187) | (0.235) |
| Income to Needs Ratio | 0.014 | 0.011 |
| | (0.037) | (0.058) |
| # Children in HH | -0.059 | -0.056 |
| | (0.040) | (0.054) |
| Parent Age | -0.065 | -0.074 |
| | (0.048) | (0.050) |
| Parent African American | 0.051 | 0.100 |
| | (0.172) | (0.232) |
| Mother Graduated H.S. | 0.022 | 0.033 |
| | (0.109) | (0.100) |
| Single Parent | -0.140 | -0.124 |
| | (0.107) | (0.126) |
| Hours of Work | 0.109 | 0.097 |
| | (0.072) | (0.056) |
| *Additional Controls* | | |
| Head Start Site Fixed Effects | Inc. | Inc. |
| Missing Dummy Variables | Inc. | Inc. |
| Observations | 375 | 359 |
| R-squared | 0.126 | 0.124 |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. The coefficients displayed in this table are standardized, representing effect sizes. For both regression models, we used mean imputation with dummy variables to adjust for missing data on covariates. In column 2, estimates were generated from models that included weights for the probability of having non-missing self-reported grades.
*p < .05, **p<.01, ***p<.001.

Additionally, we considered if treated students were motivated to take more difficult courses the following year, resulting in lower grades on average due to the challenge. As part of a survey measure, students were asked the following questions: "In the past 6 months have you left school for a job (full time or part time)?" "In the past 6 months have you left school to attend junior college, college, or vocational training?" "In the past 6 months have you been in any AP, Advanced, Honors, or IB classes?" and "In the past 6 months have you won a school award or been on honor roll?" Supplementary information file Table S4 displays regression results predicting four self-reported decisions the following school year. This is compared against the overall sample mean for these behaviors: only 3% of students left school for a job, 2.5% left school to attend junior college, college, or vocational training, 50.9% were in any AP, Advanced, Honors, or IB classes, and 46% won a school award or were on honor roll. We found no indication that students in the PFL treatment condition were more likely to a) leave school for a job, b) leave school to attend higher education, c) enroll in any AP, advanced, honors or IB classes, or d) receive an academic award.

*Table S4. Effect of PFL on Advanced Course-Taking and Extracurricular Choices*

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Left school for a job? (Full time or part time) | Left school to attend junior college, college, or vocational training? | Been in any AP, Advanced, Honors or IB classes? | Won a school award or been on honor roll? |
| PFL Condition | -0.013 | 0.012 | 0.002 | -0.008 |
| | (0.018) | (0.017) | (0.060) | (0.039) |
| *Mean* | 0.030 | 0.025 | 0.509 | 0.460 |
| Demographics | Inc. | Inc. | Inc. | Inc. |
| Head Start Site Fixed Effects | Inc. | Inc. | Inc. | Inc. |
| Missing Dummy Variables | Inc. | Inc. | Inc. | Inc. |
| Observations | 398 | 398 | 397 | 398 |
| R-squared | 0.110 | 0.073 | 0.109 | 0.100 |

*Note.* Standard errors were adjusted for Head Start site-level clustering and are displayed in parentheses. All survey questions begin with "In the past 6 months, have you…" The coefficients displayed in this table are standardized, representing effect sizes. For all regression models, we used mean imputation with dummy variables to adjust for missing data on covariates. Demographic variables and Head Start site fixed effects were included in each model. The overall sample mean of the responses are presented in the row labeled "mean" and can be interpreted as percentage of students in the sample who responded "yes."