Why Does a Growth Mindset Intervention Impact Achievement Differently across Secondary Schools? Unpacking the Causal Mediation Mechanism from a National Multisite Randomized Experiment

Xu Qin, Stephanie Wormington, Alberto Guzman-Alvarez & Ming-Te Wang

📄 View supplementary material ⎘

📅 Published online: 15 Jun 2021.

✎ Submit your article to this journal ⎘

📊 Article views: 612

🔍 View related articles ⎘

⬛ View Crossmark data ⎘

Routledge
Taylor & Francis Group

THEORY, CONTEXTS, AND MECHANISMS

Check for updates

Why Does a Growth Mindset Intervention Impact Achievement Differently across Secondary Schools? Unpacking the Causal Mediation Mechanism from a National Multisite Randomized Experiment

Xu Qin[a] (ID), Stephanie Wormington[b], Alberto Guzman-Alvarez[a] (ID) and Ming-Te Wang[a]

[a]University of Pittsburgh, Pittsburgh, Pennsylvania, USA; [b]Center for Creative Leadership, Greensboro, North Carolina, USA

ABSTRACT

The growth mindset or the belief that intelligence is malleable has garnered significant attention for its positive association with academic success. Several recent randomized trials, including the National Study of Learning Mindsets (NSLM), have been conducted to understand why, for whom, and under what contexts a growth mindset intervention can promote beneficial achievement outcomes during critical educational transitions. Prior research suggests that the NSLM intervention was particularly effective in improving low-achieving 9th graders' GPA, while the impact varied across schools. In this study, we investigated the underlying causal mediation mechanism that might explain this impact and how the mechanism varied across different types of schools. By extending a recently developed weighting method for multisite causal mediation analysis, the analysis enhances the external and internal validity of the results. We found that challenge-seeking behavior played a significant mediating role, only in medium-achieving schools, which may partly explain the reason why the intervention worked differently across schools. We conclude by discussing implications for designing interventions that not only promote students' growth mindsets but also foster supportive learning environments under different school contexts.

Introduction

Students' motivation—or opinions toward and reasons for engaging in schoolwork—are critical correlates of their academic achievement, adjustment, and success (Linnenbrink-Garcia et al., 2016; Wigfield & Cambria, 2010). When students are positively motivated, they are more likely to experience various positive achievement outcomes, including academic achievement, engagement, persistence, and long-term educational success (Maehr & Zusho,

2009; Wang & Eccles, 2013). However, students' motivation in school tends to decline significantly during adolescence, which can have enduring consequences for later educational and career outcomes (Archambault et al., 2010; Corpus et al., 2009; Wang & Degol, 2013).

To counteract motivational decline and support students' academic achievement, researchers and educators have increasingly turned to brief social-psychological interventions (Dweck & Yeager, 2019; Yeager & Dweck, 2012). One of the most rigorously studied of these interventions (Yeager et al., 2016, 2019) has targeted students' incremental beliefs about intelligence, also referred to as growth mindset. A growth mindset represents the belief that one's basic intelligence or skills are malleable and can improve through effort (Dweck, 1999; Dweck & Leggett, 1988). Despite the intuitive association that such a belief would result in positive educational outcomes, there is conflicting evidence about the impact of growth mindset interventions on academic achievement. Some experimental studies report notable positive and lasting effects of the intervention, such as increases of approximately half of a letter grade (Yeager et al., 2016; for a meta-analysis of growth mindset and other interventions targeting motivational variables, see Lazowski & Hulleman, 2016). In contrast, some meta-analytic syntheses suggest little to no overall effects on average (Sisk et al., 2018). Nevertheless, across these studies and meta-analytic syntheses, findings have consistently indicated that growth mindset interventions tend to be particularly effective for historically disenfranchised and low-achieving students (e.g., Paunesku et al., 2015; see Sisk et al., 2018 for a meta-analytic review).

In recent years, leading researchers and funding agencies have consistently encouraged researchers to extend beyond assessing whether an intervention has the intended effect and evaluate why an intervention is successful (Hulleman & Cordray, 2009; Irwin & Supplee, 2012; e.g., LoCasale-Crouch et al., 2018). There have been theories for why the growth mindset intervention may promote achievement outcomes. The most prominent of these theories assert that engaging a growth mindset changes the "meaning system" of attributions, goals, and responses to challenges (Dweck & Yeager, 2019) and then sets in motion a self-sustaining recursive process of motivation and behavior, which ultimately improves student achievement (Yeager & Walton, 2011). Despite substantial theorizing, potential mediation mechanisms underlying the growth mindset impact have been tested mainly in correlational research and not under rigorous causal frameworks. Thus, it has been difficult to draw any causal conclusions about why growth mindset interventions impact academic outcomes of interest.

An intervention may generate various impacts across different contexts due to natural variations in participant composition, local implementation, and organizational setting (Weiss et al., 2014). Hence, it is essential to further evaluate whether the impact of an intervention is generalizable across contexts. If not, we must determine in what contexts the intervention is effective and why (e.g., Qin et al., 2021; Weiss et al., 2017). Growth mindset interventions in the United States have been found to generate heterogeneous impacts on academic outcomes across schools nationwide (Yeager et al., 2019); however, little is known about the reason.

Hence, it has become important to assess the underlying causal mediation mechanisms and their heterogeneity to deepen researchers' understanding of effective intervention implementation and provide opportunities to confirm, refute, and revise guiding theoretical models. To fill this gap, we assessed whether the impact of a growth mindset intervention

on achievement is significantly transmitted through a theoretically-meaningful but empirically understudied mediator: challenge-seeking behaviors. We also examined how the mediation mechanism varies across schools at different achievement levels.

Growth Mindset, or Incremental Theories of Intelligence

Academic challenges and setbacks are part of the student experience. However, what is essential for students' subsequent academic success is not necessarily the presence or absence of academic challenges and setbacks. Rather, it is paramount that we understand how students perceive and react subjectively to such challenges. Recent studies have indicated that cultivating a growth mindset about learning may help students better cope with adversity and, as a result, achieve academic success. In contrast to a fixed mindset (i.e., an entity belief of intelligence), a growth mindset enables students to view intelligence as mutable and responsive to internal forces, such as effort and differential strategy use (Dweck, 2006; Dweck & Leggett, 1988). Compelling correlational and experimental evidence has suggested that adopting a growth mindset is positively associated with adaptive outcomes such as grades and persistence, particularly for students with a history of low achievement (e.g., Paunesku et al., 2015; Yeager et al., 2016).

Although malleability of intelligence serves as the core belief associated with a growth mindset, the endorsement of a growth mindset gives rise to a series of interconnected beliefs and behaviors, all of which have been theorized to affect achievement-related actions and outcomes (Dweck, 2006; Dweck & Leggett, 1988; Dweck & Yeager, 2019). It is essential to acknowledge these downstream views and behaviors that arise from growth mindset endorsement and how they might be associated with academic outcomes. For example, students who endorse a growth mindset believe that intelligence is malleable. Therefore, they do not interpret failure as a threat to their innate ability. Instead, they perceive it as feedback that they need to change their approach or strategy (Mueller & Dweck, 1998). Consequently, students with a growth mindset are likely to persist in the face of failure, acknowledge the importance of practice, and respond to failure with increased effort (Dweck, 2006; Hong et al., 1999; Nussbaum & Dweck, 2008; Yeager & Dweck, 2012). Students who endorse a growth mindset are also likely to seek out more challenging materials to further increase their skills if they succeed (Dweck, 2006, 2008). By contrast, students who endorse a fixed mindset are likely to avoid failure or demonstrating their lack of skill to others (e.g., decreased help-seeking behavior; Shively & Ryan, 2013).

Fostering Growth Mindset Endorsement Through Social-Psychological Interventions

Even if students do not naturally endorse a growth mindset, experimental evidence in K-12 and higher education settings suggests that growth mindset beliefs are malleable and can be developed. While earlier work focused on the influence of praise or feedback (e.g., Haimovitz & Dweck, 2017), more recent research has explored the potential impact of intervention activities on growth mindset development (e.g., Aronson et al., 2002; Blackwell et al., 2007; Boaler, 2013; Good et al., 2003; Paunesku et al., 2015; see Dweck & Yeager, 2019). Growth mindset interventions are designed to help students conceptualize intelligence and skills as malleable, recognize that trying difficult tasks provides an

opportunity to learn and grow, and understand that applying different strategies when they struggle can help them succeed (Dweck, 2008; Yeager et al., 2016, 2019).

Often, growth mindset interventions take the form of lessons about neuroplasticity and strengthening neurological pathways through learning. After reading about the growth mindset, students solidify their understanding through a writing activity, traditionally by describing the mindset and its potential benefits to a fellow student. Versions of the growth mindset intervention have been implemented in both online (e.g., Paunesku et al., 2015) and face-to-face environments (e.g., Yeager et al., 2016). These versions vary in length, with some interventions administered over eight separate sessions (Aronson et al., 2002; Blackwell et al., 2007; Good et al., 2003) and others as brief as growth mindset messages built into Khan Academy lessons (Paunesku et al., 2015).

Evidence has shown that growth mindset interventions are more impactful for low-achieving students who are vulnerable to future academic challenges and setbacks (Blackwell et al., 2007; Broda et al., 2018; Paunesku et al., 2015; Sisk et al., 2018). The effects of endorsing a growth mindset are thought to be most pronounced when students face challenges, failures, or setbacks because they tend to react with adaptive, approach-oriented beliefs and behaviors (e.g., help-seeking, making internal and controllable attributions for failure) rather than maladaptive, avoidance-oriented beliefs and behaviors (e.g., endorsing performance-avoidance goals, giving up on tasks; Dweck & Leggett, 1988). At the beginning of high school, low-achieving students experience academic challenges more frequently than their high-achieving counterparts, and these failure experiences can begin a self-perpetuating cycle that gives rise to future failure experiences and negative achievement trajectories (i.e., recursive processes; Cohen et al., 2009). Hence, these students are expected to benefit more from growth mindset interventions that reframe how students perceive and interpret challenge and failure in a more positive and adaptive manner (Yeager et al., 2019).

Where Is a Growth Mindset Intervention Effective, and Why? Assessing Heterogeneity in Mediation Mechanisms

Growth mindset interventions have primarily been conducted at single sites and focused on overall intervention effects, which mirrors patterns in research on social-psychological and educational interventions generally. Researchers have recently argued that this is not sufficient for informing policy and practice; rather, it is critical to determine whether intervention impacts vary by contextual factors. As such, the importance of evaluating the between-site heterogeneity of intervention impacts has become increasingly valued (e.g., Heckman et al., 1997; Olsen, 2017; Raudenbush & Bloom, 2015; Weiss et al., 2017).

To address this concern and produce more generalizable results surrounding the efficacy of growth mindset interventions, a team of leading researchers launched the National Study of Learning Mindsets (NSLM), a nationally representative multisite randomized evaluation of an online growth mindset intervention for 9th grade students during the 2015–2016 school year in U.S. public high schools (Yeager et al., 2019). With students randomly assigned to treatment and control groups within each school and schools purposefully sampled based on their sizes, average achievement levels, and

demographic compositions, the NSLM study offers unique opportunities for investigating whether the intervention impact is generalizable across different contexts. Findings have indicated that the growth mindset intervention significantly improved low-achieving students' academic returns, while this impact varied significantly across schools.

Where?—Moderating Role of School Achievement Levels

In particular, the positive impact was most pronounced in medium-achieving schools[2] (Yeager et al., 2019). Tipton, Yeager, and colleagues (Tipton et al., 2016, 2019; Yeager et al., 2019) have argued that the potential positive effects of motivation for students in low-achieving schools are suppressed by limited resources and concerns about fulfilling basic needs (e.g., safety). Conversely, positive impacts of students' motivation in high-achieving schools may be negligible because access to high-quality learning opportunities already bolsters student achievement in these settings. As a result, students in medium-achieving schools are most likely to display benefits from a growth mindset intervention.

This finding provides novel information about *where* growth mindset interventions may be most impactful. However, it does not address *why* the growth mindset intervention functioned differently in different school contexts (Dweck & Yeager, 2019). Assessing the underlying causal mediation mechanism and its heterogeneity across school settings may (1) be crucial for unpacking and understanding the variation in the total intervention impact, (2) inform a necessity to revisit the theory behind growth mindset interventions or growth mindset more generally, and (3) suggest school-specific modifications of the intervention practice and implementation itself.

Why?—Mediating Role of Challenge-Seeking Behaviors

In the current study, we focused on challenge-seeking as a potential mediator that transmits the impact of the growth mindset intervention. In our hypothesized mediation model, as represented in Figure 1, the growth mindset intervention may influence low-achieving students' challenge-seeking behaviors, which may subsequently impact their academic outcomes (Horng, 2016; Yeager et al., 2018). In other words, challenge-seeking behaviors may partially mediate the impact of the growth mindset intervention. Therefore, the total intervention impact can be decomposed into an indirect effect transmitted via challenge-seeking behaviors and a direct effect that functions directly or through other unspecified pathways, such as expectations for success, help-seeking behaviors, internal attributions for success and failure, or mastery-approach goal orientations.

Challenge-seeking is a crucial element of Dweck's (1986) original conceptualization of the growth mindset and a downstream behavioral consequence of endorsing a growth mindset; however, it is not yet clear whether challenge-seeking serves as a significant mediator. Growth mindset endorsement and, by extension, growth mindset interventions, were expected to foster more adaptive reactions to challenge, such as persisting in response to failure or seeking out more challenging materials to spur future growth.

---

[2]Yeager et al. (2019) generated school achievement level as a latent variable based on publicly available indicators of school performance in state and national tests and other related factors. Low-, medium-, and high-achieving schools are respectively the bottom 25%, middle 50%, and top 25%.
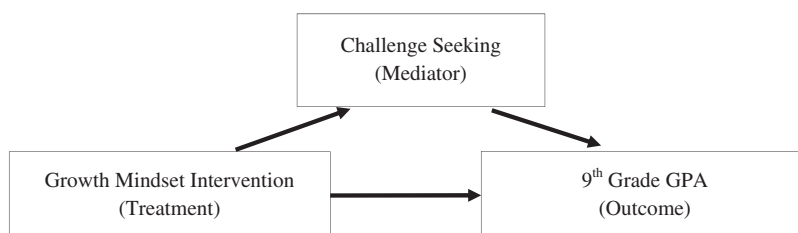
Figure 1. Diagram showing the hypothesized mediation mechanism of the growth mindset intervention impact. The total intervention impact can be decomposed into an indirect effect transmitted via challenge-seeking behaviors, represented by the arrow from the treatment to the mediator and that from the mediator to the outcome, and a direct effect that functions directly or through other unspecified pathways, represented by the arrow from the treatment to the outcome.

Despite the historical roots of growth mindset research (i.e., understanding the phenomenon of learned helplessness) and early work indicating that a growth mindset should give rise to adaptive reactions to challenge, comparatively few correlational or intervention-based studies have assessed the relationship between growth mindset and challenge-seeking behaviors. Foundational work examining the effect of praise on the growth mindset has indicated that there is an association between growth mindset and challenge-seeking behaviors, operationalized as choosing more difficult problems to solve and avoiding attributing their struggle on difficult tasks to a lack of ability (e.g., Gunderson et al., 2013; Mueller & Dweck, 1998). Given its theoretical importance and underrepresentation in empirical intervention research, findings from the current study on the mediating role of challenge-seeking will contribute to both theoretical and empirical understanding of growth mindset.

Sociological theory highlights the role that a school's curricula and instruction quality play in sustaining or restraining students' motivation following a growth mindset intervention (Yeager et al., 2019). In addition, decades of motivational theory and research have underscored the importance of educational context in the motivation development and its impact on academic outcomes (e.g., Linnenbrink-Garcia et al., 2016; Wang & Eccles, 2013). Hence, the hypothesized mediation mechanism was expected to vary across schools, which may give rise to the unevenness in the intervention impact across schools.

Due to its multi-site data collection, the NSLM study provides a unique opportunity to investigate the important between-school difference in the mediation mechanism. Advancement in this line of research has been underdeveloped due to the lack of analytic tools. The conventional methods, multilevel path analysis and SEM, rely on strong functional and distributional assumptions, ignore potential confounding bias, and have difficulties in estimating and testing the between-site variance of the mediation effects and especially in evaluating the how the mediation mechanism varied by school characteristics. To overcome the limitations, Qin and colleagues (Qin et al., 2019, 2021; Qin & Hong, 2017) have recently developed conceptual frameworks and statistical tools for investigating the population average and the between-site variation of causal mediation mechanisms in multisite randomized trials. By extending this approach, the current study explored whether the school achievement level moderated the indirect effect via challenge-seeking in a pattern similar to prior findings concerning the heterogeneity of

the total impact. In other words, we investigated if the mediating role of challenge-seeking was most pronounced in medium-achieving schools.

Research Questions of the Current Study

Our study sought to examine challenge-seeking as a mediator underlying the growth mindset intervention's impact and evaluate the between-school variation of its mediating role. To be specific, we investigated the following research questions:

Population Average Causal Mediation Mechanism

We evaluated the extent to which the growth mindset intervention improved low-achieving students' GPA in core classes (mathematics, science, English or language arts, and social studies) via enhancing their challenge-seeking behaviors. We did so by decomposing the total intervention impact into an indirect effect transmitted through challenge-seeking behaviors and a direct effect that works directly or through other unspecified pathways. We also assessed whether the growth-mindset-induced change in challenge-seeking behaviors generated a greater impact on academic outcomes in the intervention group than in the control group.

Between-School Heterogeneity of the Causal Mediation Mechanism

To explain the between-school variation of the growth mindset intervention's impact on low-achieving students' academic outcomes, we assessed whether the phenomenon was partly due to the variation in the mediating role of challenge-seeking behaviors across schools, and if so, what features of schools might explain such variation. In particular, we tested whether the mediating role of challenge-seeking behaviors might be less pronounced within the low-achieving schools (which, in theory, may not be able to provide students with the same quality or consistency of learning opportunities and supports and may disproportionately serve students who are less likely to have basic nonacademic needs met regularly) and high-achieving schools (which in theory may have already provided students with the high-quality learning opportunities and supports such that any additional intervention may only have diminishing returns in terms of additional motivational gains) (Tipton et al., 2019; Yeager et al., 2019).

Findings from this study may inform efforts to design interventions that not only promote students' growth mindsets, but also foster supportive learning environments at the school level. It will also contribute to our theoretical understanding of where growth mindset interventions may be most effective and what potential role challenge-seeking plays in that effect across different school contexts.

Method

Research Design and Target Population

The NSLM used a stratified random sample of 139 schools selected from around 12,000 public high schools in the United States. 65 schools, including 12,490 9th grade students, agreed to participate and provide student records to the research team. Yeager

et al. ([2019](#)) have verified based on the Tipton generalizability index (Tipton, [2014](#)) that the analytic sample featured a nationally representative probability sample of regular U.S. public high schools (Yeager et al., [2019](#)). Participants were asked to complete two 25-min self-administered online sessions during regular school hours, spaced around 20 days apart. Within each school, students were randomly assigned to an intervention group (for which the sessions were designed to reduce negative effort beliefs, fixed-trait attributions, and performance avoidance goals and motivate challenge-seeking behaviors) and a control group (for which the sessions focused on brain functions while not emphasizing intelligence beliefs). Students and teachers were blind to the study goals and group assignments throughout the study. The study had a high rate of fidelity. Students viewed 95% of screens during the on-line sessions. The attrition rate between the first and second sessions was lower than 10% for both the intervention and control groups.

Because the randomization was conducted within schools, this study used a multisite randomized design with schools as experimental sites. As discussed in Raudenbush and Bloom ([2015](#)) and Raudenbush and Schwartz ([2020](#)), there are two potential targets of inference in a multisite study. In the NSLM, one is the population of students, and the other is the population of schools. The former is of more interest when the focus is on the implementation of the intervention among all the students and each student is equally representative of the population. The latter is of more interest when attention is paid to the performance of the intervention at the school level and how the intervention impact varies across schools. Because our goal is to evaluate if the growth mindset intervention impacts and the underlying mediation mechanisms are generalizable across schools, we choose the population of public high schools in the U.S as the target of inference.

Measures and Study Sample

Students completed the baseline survey before randomization and completed the follow-up survey immediately following the second online session. Both baseline and follow-up surveys captured student demographic and psychometric measures, along with additional student achievement measures captured from school administrative files.

Outcome—GPA

We considered 9th grade GPA in core classes (mathematics, science, English or language arts, and social studies) at the end of the academic year as the outcome of interest. Schools provided students' grades in each course. All numeric and letter grades were standardized across all the schools to a scale of 0–4.3. The average grade of all the core courses taken from the intervention term to the end of the 9th grade was calculated for each student.

Mediator—Challenge-Seeking Behaviors

We considered challenge-seeking behavior measured after students completed both sessions as the focal mediator. Students were asked to choose between an easy mathematics assignment, in which they were more likely to get most problems right but not learn

anything new, and a difficult mathematics assignment, in which they were more likely to get more problems wrong but learn something new.

Moderator—School Achievement Level

We evaluated how the mediation mechanism varied by school achievement level, which was generated as a latent variable based on publicly available indicators of school performance in state and national tests and other related factors. Low-, medium-, and high-achieving schools were considered the bottom 25%, middle 50%, and top 25%, respectively (Yeager et al., 2019).

In this study, we focused on 6,258 low-achieving students. Yeager et al. (2019) defined students as low-achieving "if they were earning GPAs at or below the school-specific median in the term before random assignment" or "if they were below the school-specific median on academic variables used to impute prior GPA" for those who were missing prior GPA data. The sample reflects a diverse background of low-achieving 9th graders in the U.S. (41% female, 40% white, 12% black/African-American, 31% free or reduced-price lunch, and 25% reported that their mother had a bachelor's degree or higher). It includes 65 schools in which the number of low-achieving students ranged from 15 to 338, with a mean of 96. Some students failed to provide information on their challenge-seeking behaviors or GPA. We define them as non-respondents. The proportion of non-respondents within each school varies from 20% to 100%, with a mean of 86%.

To ensure that the results of the present study are generalizable to the entire population of public high schools in the U.S., we employed sample weights to adjust for sample and survey designs and nonresponse weights to account for nonresponse. The latter is designed to safeguard against, for example, a situation in which only highly-engaged students responded, which would otherwise skew study results so that they would only apply to highly-engaged students.

Causal Estimands

In drawing causal conclusions, it is necessary to clarify the definitions of the causal effects first. We define the population average and between-school variance of the causal direct and indirect effects under the potential outcomes causal framework (Neyman et al., 1935). Let T_{ij} denote the treatment assignment of student i at school j. It takes values $t = 1$ for an assignment to the growth mindset intervention and $t = 0$ for the control group. Let M_{ij} denote the focal mediator, challenge-seeking, and Y_{ij} denote the outcome, 9th grade GPA. We view the potential mediator as a function of the treatment assignment, $M_{ij}(t)$, which represents the student's potential challenge-seeking behavior if assigned to treatment group t. $M_{ij}(t) = 1$ if the student chose a difficult mathematics assignment under treatment condition t, and $M_{ij}(t) = 0$ if the student chose an easy one. We view the potential outcome as a function of the treatment assignment and the potential mediator. $Y_{ij}(t, M_{ij}(t))$ represents the student's potential GPA under treatment condition t, and $Y_{ij}(t, M_{ij}(t'))$ for $t \neq t'$ represents the student's potential GPA if assigned to treatment t while his or her challenge-seeking behavior took the value under the counterfactual condition t'. The potential mediators and outcomes are defined under the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1980, 1986, 1990),

which assumes that, within each school or across schools, a student's potential mediators are unrelated to the treatment assignments of other students, and a student's potential outcomes are independent of the treatment assignments and the mediator values of other students.

The intention-to-treat (ITT) effect is defined as a contrast of the potential outcome between the two treatment conditions for each student, i.e., $ITT_{ij} = Y_{ij}(1, M_{ij}(1)) - Y_{ij}(0, M_{ij}(0))$. This can be decomposed into a natural indirect effect (NIE) of the growth mindset intervention on GPA transmitted via one's challenge-seeking behavior, $NIE_{ij} = Y_{ij}(1, M_{ij}(1)) - Y_{ij}(1, M_{ij}(0))$, and a natural direct effect (NDE) of the intervention on GPA, $NDE_{ij} = Y_{ij}(1, M_{ij}(0)) - Y_{ij}(0, M_{ij}(0))$ (Pearl, 2001; Robins & Greenland, 1992). The former represents the impact of the growth mindset intervention on GPA attributable to the intervention-induced change in one's challenge-seeking behavior while all the other elements are held at the level under the growth mindset condition. The latter indicates the growth mindset impact on GPA when one's challenge-seeking behavior is held at the level that he or she would have under the control condition. Alternatively, the ITT effect can be decomposed into a pure indirect effect, $PIE_{ij} = Y_{ij}(0, M_{ij}(1)) - Y_{ij}(0, M_{ij}(0))$, and a total direct effect $TDE_{ij} = Y_{ij}(1, M_{ij}(1)) - Y_{ij}(0, M_{ij}(1))$ (Robins & Greenland, 1992). NIE_{ij} may not be equal to PIE_{ij}, and equivalently, NDE_{ij} may not be equal to TDE_{ij}. A discrepancy between the two decompositions exists if the intervention-induced change in challenge-seeking behaviors influences students' 9th grade GPA differently between the growth mindset condition and the control condition. Hong et al. (2015) defined the difference as a natural treatment-by-mediator interaction effect.

Above we have defined the effects for each student. By taking an average of each effect over all the low-achieving students at a given school, we then define the corresponding school-specific effect, based on which it is straightforward to define the population average and between-school variance of the effect among low-achievers.

Identification Assumptions

$Y_{ij}(t, M_{ij}(t))$ was observed only if student i at school j was selected into the sample, was assigned to treatment t, and provided information on the outcome, while $Y_{ij}(t, M_{ij}(t'))$ for $t \neq t'$ is never observable. To equate their expectations with the observed quantities at each school, we make the following assumptions proposed by Qin et al. (2019):

1. *Strongly ignorable sampling mechanism.* Sample selection is independent of the potential mediators and outcomes within levels of the observed pretreatment covariates (i.e., covariates preceding the treatment assignment) at each school. This assumption is satisfied by the sampling design.
2. *Strongly ignorable treatment assignment.* The treatment assignment of each sampled student is independent of the potential mediators and outcomes within levels of the observed pretreatment covariates at each school. This assumption is guaranteed by the random treatment assignment.

3. *Strongly ignorable nonresponse.* Whether a sampled student provided information on both the mediator and the outcome in a given treatment group is independent of the potential mediators and outcomes under the same treatment condition within levels of the observed pretreatment covariates at each school. In other words, conditional on the observed pretreatment covariates, a participant is as if randomized to respond to the mediator and the outcome in each treatment group at each school. A violation of this assumption would not only change the representativeness of the sample but also induce systematic pretreatment discrepancy between the intervention group and the control group in the remaining sample.

4. *Strongly ignorable mediator.* The mediator value of each respondent is independent of the potential outcomes under either treatment condition within levels of the observed pretreatment covariates at each school. This assumption can be alternatively expressed as, among the students with the same observed pretreatment covariates, their mediator values are as if randomized in each treatment group or across treatment groups at each school.

Assumptions (3) and (4) are particularly strong because they assume no unmeasured pretreatment confounders and no posttreatment confounders (i.e., confounders affected by treatment) of the relationship between the response status and the mediator (or the outcome) or that between the mediator and the outcome. These assumptions cannot be guaranteed by design. For example, among the low-achieving students with the same observed pretreatment covariates, those whose parents were engaged in tutoring with them before the intervention might be more likely to seek challenges, have better academic performance, and provide information on both measures. In other words, parental engagement prior to treatment might confound the response-mediator, response-outcome, or mediator-outcome relationship. In addition, students assigned to the intervention group might engage more with the treatment messages and thus became more inclined to seek out challenges, have higher GPA, and respond to the related questions. Hence, treatment engagement, which is affected by treatment, might also confound the response-mediator, response-outcome, or mediator-outcome relationship. Failures to account for such pretreatment or posttreatment confounders would lead to violations of Assumptions (3) and (4). It is almost impossible to observe all the confounders in the real world. Nevertheless, as explicated later, one can use sensitivity analysis to evaluate the extent to which potential violations of the assumptions would change the initial conclusions.

Identification Results

Under Assumptions (1)–(3), Qin et al. (2019) have proved that, $E\left[Y_{ij}(t, M_{ij}(t)) \mid S_{ij} = j\right]$ can be identified by a weighted average of the observed outcome among the sampled low-achievers ($D_{ij} = 1$) who were assigned to treatment group t ($T_{ij} = t$) and responded to both the mediator and the outcome ($R_{ij} = 1$) at school j ($S_{ij} = j$):

$$E\big[Y_{ij}(t, M_{ij}(t))\ |S_{ij}=j\big] = E\big[W_{Dij}W_{Rij}\ Y_{ij}\ |D_{ij}=1, T_{ij}\ =t,\ R_{ij}=1, S_{ij}=j\big].$$

The weights are constructed based on propensity scores of D_{ij} and R_{ij} within levels of pretreatment covariates $\mathbf{X}_{ij}$ at each school. The sample weight $W_{Dij} = \frac{Pr\big(D_{ij}=1|\ S_{ij}=j\big)}{Pr\big(D_{ij}=1|\mathbf{X}_{ij}=\mathbf{x},\ S_{ij}=j\big)}$ restores the sample representativeness, and the nonresponse weight $W_{Rij} = \frac{Pr\big(R_{ij}=r|T_{ij}=t, D_{ij}=1, S_{ij}=j\big)}{Pr\big(R_{ij}=r|\mathbf{X}_{ij}=\mathbf{x}, T_{ij}=t, D_{ij}=1,\ S_{ij}=j\big)}$ for $t=0,\ 1$ and $r=0,\ 1$ removes the pretreatment discrepancy between the respondents ($R_{ij}=1$) and nonrespondents ($R_{ij}=0$) under each treatment condition.

Under Assumptions (1)–(4), based on the observed outcome of the same subgroup of people as above, $E\big[Y_{ij}(t, M_{ij}(t'))\ |S_{ij}=j\big]$ for $t \neq t'$ can be identified by:

$$E\big[Y_{ij}(t, M_{ij}(t'))\ |S_{ij}=j\big] = E\big[W_{Dij}W_{Rij}W_{Mij}\ Y_{ij}\ |D_{ij}=1, T_{ij}\ =t,\ R_{ij}=1, S_{ij}=j\big]$$

where $W_{Mij} = \frac{Pr(M_{ij}=m|\mathbf{X}_{ij}=\mathbf{x},\ R_{ij}=1, T_{ij}=t',\ D_{ij}=1, S_{ij}=j)}{Pr(M_{ij}=m|\mathbf{X}_{ij}=\mathbf{x}, R_{ij}=1, T_{ij}=t, D_{ij}=1,\ S_{ij}=j)}$ for $m=0,\ 1$ and $t,\ t'=0,\ 1$, in which the denominator represents one's conditional probability of having the mediator value m in the assigned treatment group t, and the numerator is his or her conditional probability of displaying the same mediator value under the counterfactual treatment condition. The weight is known as the ratio-of-mediator-probability weight (RMPW) (Hong, 2010, 2015). It adjusts for the mediator value selection and transforms the mediator distribution in treatment group t to resemble that in treatment group t', enabling the identification of the expected potential outcome under the treatment condition t while the potential mediator takes the value under the counterfactual condition t'.

With the expectation of each potential outcome at each school identified, we can identify the school-specific causal effects as contrasts of the weighted mean outcome at each school. The identification of the ITT effect, which only involves the potential outcome under each treatment condition, relies on Assumptions (1)–(3); while the identification of the natural direct and indirect effects, which further involve a third counterfactual potential outcome, is based on Assumptions (1)−(4).

As discussed in the section introducing the NSLM data, we chose the population of schools as our target of inference, because our primary interest is in the implementation of the growth mindset intervention at the school level and how the impact varied across schools. Therefore, we identified the population average and between-school variance of each causal effect, respectively by an average and variance of the school-specific effects over all the sampled schools. To adjust for the sample selection of schools, we further applied a school-level sample weight given by design.

The identification of the between-school heterogeneity in the mediation mechanism helps answer our initial question of *why* the growth mindset intervention generates heterogeneous impacts across different contexts. It is essential to further assess *where* the mediation mechanism is significant and where it is not, by investigating how school-level characteristics moderate the mediation mechanism. Such an evaluation may help practitioners to make specific school-level modifications of the growth mindset intervention. However, this has never been discussed in the literature of multisite trials due to the lack of analytic tools. Assuming the above identification assumptions hold at each

level of a school-level moderator, we can identify the causal mediation effects at each level of the moderator by conducting the above weighting adjustment within each sub-population defined by the moderator. A contrast between levels of the moderator identifies a moderating effect.

Analytic Procedure

As shown in the above identification results, the analysis is based on the respondents to both the mediator and the outcome, i.e., those who provided information on the mediator and the outcome, and the key of the analytic procedure is to apply a series of weights for enhancing the external and internal validity of the analytic conclusions concerning the causal mediation mechanism. The NSLM data provide a sampling weight with an adjustment of nonparticipation in the surveys, which serves as W_{Dij}. We further adjusted for the participants' selection into response to both the mediator and the outcome and the respondents' selection into different mediator levels by constructing the nonresponse weight W_{Rij} and RMPW weight W_{Mij}. The estimation of the weights involves two steps. We first selected observed pretreatment covariates, based on which we fit propensity score models of the response indicator and the mediator. We then constructed the weights based on the predicted propensity scores. By applying the estimated weights to the respondents, we estimated the causal parameters under a method-of-moment framework. We finally used balance checking to evaluate whether selection bias associated with the observed pretreatment covariates has been effectively reduced and conducted sensitivity analysis to further assess if a potential unmeasured confounder would easily change the analytic conclusions.

   1. **Select observed pretreatment covariates**. To account for selection into nonresponse and the mediator (challenge-seeking behaviors), we selected pretreatment confounders of the response-mediator, response-outcome, or mediator-outcome relationship on theoretical grounds. These include self-reports of demographics (gender, race, parent education, free or reduced lunch status, gifted and talented status, special education status, English language learner status, first year freshman status, and GPA) and psychological constructs (school belonging, math interest, student-teacher trust, stress, fixed mindset, effort beliefs, and expectation for success). All these covariates were measured at the baseline. We generated a missing indicator for each pretreatment covariate with missing cases. Online Appendix A lists the description of all the selected pretreatment covariates and their summary statistics by the treatment condition, response status, and mediator.

   2. **Estimate propensity scores**. We fit the following multilevel logistic regression model of the response indicator to survey participants in each treatment group for estimating the denominator of the nonresponse weight—conditional probability that student i would provide information on both the mediator and the outcome under treatment condition t at school j, $p_{Rtij} = Pr(R_{ij} = 1 | \mathbf{X}_{ij} = \mathbf{x}, T_{ij} = t, D_{ij} = 1, S_{ij} = j)$:

$$\log\left[\frac{p_{Rtij}}{1 - p_{Rtij}}\right] = \beta_{0Rt} + \mathbf{X}'_{ij}\boldsymbol{\beta}_{Rt} + r_{Rtj}, \ r_{Rtj} \sim N(0, \sigma^2_{Rt}),$$

in which r_{Rtj} is the random intercept associated with school j. By removing $\mathbf{X}_{ij}$ from the above model, we could estimate the numerator of the nonresponse weight—average

probability of responding to both the mediator and the outcome under treatment condition t at school j, $p_{Rtj} = Pr(R_{ij} = 1 | T_{ij} = t, D_{ij} = 1, S_{ij} = j)$.

We fit the following multilevel logistic regression model of the mediator to those who responded to both the mediator and the outcome in each treatment group for estimating the RMPW weight:

$$\log\left[\frac{p_{Mtij}}{1 - p_{Mtij}}\right] = \beta_{0Mt} + \mathbf{X}'_{ij}\boldsymbol{\beta}_{Mt} + r_{Mtj}, \ r_{Mtj} \sim N(0, \sigma^2_{Mt}),$$

in which $p_{Mtij} = Pr(M_{ij} = 1 | \mathbf{X}_{ij} = \mathbf{x}, R_{ij} = 1, T_{ij} = t, D_{ij} = 1, S_{ij} = j)$ and r_{Mtj} is the random intercept associated with school j. Based on the model fitted to treatment group t, we can directly predict the denominator of the RMPW weight—conditional probability that student i in treatment group t at school j would display the observed mediator value. By applying to the same student the coefficients of the model fitted to the alternative treatment group t', where $t' \neq t$, we could predict the numerator of the RMPW weight for him or her—conditional probability that he or she would display the same mediator value under the counterfactual treatment condition t', $p_{Mt'ij} = Pr(M_{ij} = 1 | \mathbf{X}_{ij} = \mathbf{x}, R_{ij} = 1, T_{ij} = t', D_{ij} = 1, S_{ij} = j)$.

3. Construct the weights. Based on the predicted propensity scores, we constructed the nonresponse weights $\hat{W}_{Rij} = \hat{p}_{Rtj}/\hat{p}_{Rtij}$ for respondents ($R_{ij} = 1$) and $\hat{W}_{Rij} = (1 - \hat{p}_{Rtj})/(1 - \hat{p}_{Rtij})$ for nonrespondents ($R_{ij} = 0$). At the same time, we constructed the RMPW weights $\hat{W}_{Mij} = \hat{p}_{Mt'ij}/\hat{p}_{Mtij}$ for the respondents who are challenge seekers ($M_{ij} = 1$) and $\hat{W}_{Mij} = (1 - \hat{p}_{Mt'ij})/(1 - \hat{p}_{Mtij})$ for the respondents who are not challenge seekers ($M_{ij} = 0$). We can then apply the product of the given weight W_{Dij} and the estimated weights $\hat{W}_{Rij}$ and $\hat{W}_{Mij}$ to the respondents, for the estimation of the causal estimands.

4. Estimate and test the causal estimands. We adopted a method-of-moments procedure that Qin and Hong (2017) and Qin et al. (2019) developed for the estimation of the population average and between-school variance of the effects. We first estimated each causal effect through a weighted mean contrast of the outcome school by school. Subsequently, we estimated the population average and between-school variance of the effects over the population of schools. The estimation procedure incorporates the sampling uncertainty of the weights estimated in the previous step. The hypothesis testing of the population average effects is based on t-tests and that of the between-school variances is based on permutation tests. The same procedure applies to the estimation and inference of the effects in each subpopulation defined by the school-level moderator, enabling the estimation and inference of the moderating effects. This estimation procedure does not require an outcome model specification and thus avoids the risk of possible misspecifications of the outcome model's functional or distributional form.

5. Balance checking. By applying $\hat{W}_{Rij}$, we expected respondents and nonrespondents to be comparable in their observed pretreatment covariates under each treatment condition at all the schools. To verify this, we assessed if the imbalance in the distribution of each observed pretreatment covariate between respondents and nonrespondents is removed after applying $\hat{W}_{Rij}$. A covariate is considered balanced if the magnitude of the standardized weighted mean difference in the covariate between respondents and

nonrespondents is smaller than 0.25 and preferably smaller than 0.10 (Harder et al., 2010). To evaluate if balance is achieved over all the schools, we constructed the 95% plausible value range of the school-specific standardized weighted mean difference in each covariate. Similarly, we adopted the same procedure to evaluate if $\hat{W}_{Mij}$ removes the difference in each observed pretreatment covariate between challenge seekers and non-challenge seekers in each treatment group at all the schools.

6. Sensitivity analysis. An application of the product of W_{Dij}, $\hat{W}_{Rij}$, and $\hat{W}_{Mij}$ is expected to remove selection bias in identifying the causal estimands, under the ignorability assumptions of sampling mechanism, treatment assignment, nonresponse, and mediator. While the first two assumptions are guaranteed by design, the latter two would be violated if there were an unmeasured pretreatment confounder or a posttreatment confounder of the response-mediator, response-outcome, or mediator-outcome relationship. It is always possible that at least one unmeasured confounder exists. Hence, it becomes essential to conduct a sensitivity analysis to evaluate if removing the hidden bias due to unmeasured confounding would lead to a substantial change in the magnitudes of the causal effect estimates or flip their signs or significance. Sensitivity analysis for posttreatment confounding is still underdeveloped. A discussion of the sensitivity analysis for assessing the influence of posttreatment confounding can be found in the discussion section. Here we focused on assessing the influence of unmeasured pretreatment confounding. Assuming that the confounding role of an unmeasured pretreatment confounder U is comparable to that of an observed pretreatment covariate X, we estimated the plausible bias due to the omission of U by comparing the results before and after controlling for X in the analysis. Through such an evaluation for each observed pretreatment covariate, we could then obtain a plausible range of bias contributed by unmeasured pretreatment confounding.

Results

By applying the analytic procedure to the subsample of low-achieving students, we obtained the estimation results for the population average and between-school standard deviation of each causal effect, as shown in Table 1.

ITT Effects of the Growth Mindset Intervention

The growth mindset intervention significantly increased the 9th grade GPA of low-achieving students by 0.213 (SE = 0.104, $p = 0.04$) grade points, while a typical low-achieving student in the control group had a 9th grade GPA of 1.895. This impact amounted to 22% of a standard deviation of the outcome and varied significantly across schools at the significance level of 0.1. The estimated between-school standard deviation of the impact was 0.240 ($p = 0.065$). If the impact followed a normal distribution in the population of schools, it would range from −0.257 to 0.683 in 95% of schools. This finding indicates that, even though the growth mindset intervention significantly increased low-achieving students' 9th grade GPA on average, it might have generated negative impacts at some schools. In contrast, the intervention did not significantly

Table 1. Mediation mechanism underlying the growth mindset intervention impact.

	Population average effect			Between-school standard deviation		95% Plausible value range of school-specific effects
	Estimate	Effect size	p-Value	Estimate	p-Value	
ITT effect on the outcome	0.213 (0.104)	0.224	0.040	0.240	0.065	[−0.257, 0.683]
NDE	0.174 (0.095)	0.183	0.069	0.195	0.111	[−0.208, 0.556]
NIE	0.040 (0.024)	0.042	0.103	0.026	0.050	[−0.011, 0.091]
T-by-M interaction effect	0.030 (0.027)	0.032	0.258	0.000	0.688	–

increase non-low achievers' GPA across all the schools: The average effect was 0.132 (SE = 0.076, $p = 0.08$), with a between-school standard deviation of 0.135 ($p = 0.323$).

Mediation Mechanism Underlying the Growth Mindset Impact

To understand why the growth mindset intervention significantly increased low-achieving students' 9th grade GPA, we focused on the sample of low-achieving students and decomposed the ITT effect into an indirect effect via challenge-seeking behaviors and a direct effect transmitted through any other possible pathways. The average natural indirect effect was estimated to be 0.040 (SE = 0.024, $p = 0.103$), which indicates the growth mindset impact that was solely attributable to the intervention-induced increase in one's challenge-seeking behaviors under the growth mindset condition. It is similar in magnitude to the pure indirect effect, which captures the mediating role of challenge-seeking behaviors under the control condition. Their difference was estimated to be 0.030 (SE = 0.027, $p = 0.258$), indicating no significant interaction between the treatment and the mediator.

Even though the natural indirect effect was insignificant in the overall population of schools and only accounted for 19% of the ITT effect, it varied significantly across schools. Its between-school standard deviation was estimated to be 0.026 ($p = 0.05$). If the school-specific natural indirect effect via challenge-seeking followed a normal distribution, it would range from −0.011 to 0.091 in 95% of schools. As explicated earlier, the indirect effect is transmitted via the path from the intervention to challenge-seeking and then from challenge-seeking to GPA, as represented in Figure 1. To further unpack the source of the between-school heterogeneity in the indirect effect, we look into the two paths separately. The growth mindset intervention stimulated an average increase in the proportion of challenge seekers among low-achievers from 37% to 53%. The impact was statistically significant on average (SE = 0.05, $p = 0.003$) and did not vary significantly. In other words, the path from the intervention to challenge-seeking was consistently significant across schools. Hence, the between-school variation in the indirect effect via challenge-seeking is mainly due to the inconsistency in the path from challenge-seeking to GPA, representing the impact of the intervention-induced increase in challenge-seeking on GPA. This may be largely due to the heterogeneity in contextual supports.

The average natural direct effect was estimated to be 0.174 (SE $= 0.095$, $p = 0.069$), with an estimated between-school standard deviation of 0.195 ($p = 0.111$). This indicates that the growth mindset impact transmitted through all the other possible pathways, though insignificant, accounts for most of the ITT effect on average. Nevertheless, this effect did not vary significantly across schools. Hence, to understand the reason for the between-school heterogeneity in the ITT effect, it is crucial to evaluate how the natural indirect effect via challenge-seeking varied by school contexts.

Moderating Role of School Achievement Level

As found in the above analysis, the growth mindset intervention heterogeneously changed low-achieving students' GPA across schools, partly because challenge-seeking behaviors played different mediating roles in the underlying mechanism of the growth mindset intervention's effect across schools. It is important to further investigate features of schools that may moderate the mechanism. Previous studies have found that the impact of a growth mindset intervention may vary by schools at different achievement levels. Low-achieving schools may lack resources for students to benefit from the intervention, while the intervention would not add much in the high-achieving schools that likely already had abundant resources. As a result, students in medium-achieving schools were hypothesized to be the most likely to capitalize on a growth mindset intervention.

In the current study, we tested if the hypothesis about the moderating role of school achievement level also held for the mediation mechanism underlying the growth mindset impact.

As shown in Figure 2, challenge-seeking significantly mediated the impact of growth mindset on low-achieving students' GPA in medium-achieving schools ($NIE_M = 0.084$, SE $= 0.044$, $p = 0.058$). This indirect effect amounted to about 9% of the outcome's standard deviation and about half of the total intervention impact among medium-achieving schools. In low-achieving schools, the ITT effect was barely transmitted through challenge-seeking behaviors ($NIE_L = -0.002$, SE $= 0.039$, $p = 0.963$). Similarly, the mediating role of challenge-seeking in high-achieving schools was not significant ($NIE_H = -0.021$, SE $= 0.029$, $p = 0.472$).

Balance Checking

As shown in Online Appendix B, nonresponse weighting greatly reduced the imbalance between nonrespondents and respondents in the observed pretreatment covariates in both treatment groups. Take the intervention group as an example. Before weighting, the average magnitude of the standardized mean difference between nonrespondents and respondents was larger than 0.25 for four covariates and larger than 0.1 for 16 other covariates. After weighting, the standardized mean difference was smaller than 0.1 in magnitude for all covariates. A similar improvement in balance can be found in the control group. Meanwhile, the RMPW weighting also improved balance between challenge seekers and non-challenge seekers under both treatment conditions. Following the same procedure, we also found balance achieved in each subgroup of schools at the
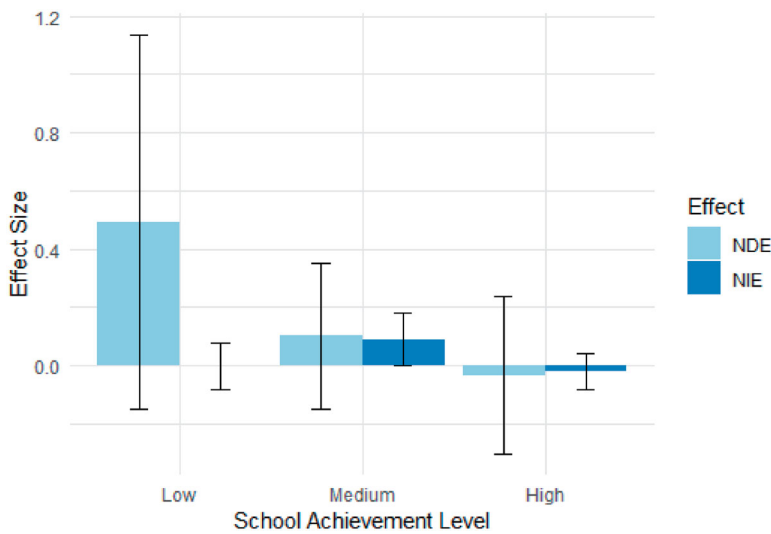
Figure 2. Bar chart showing the natural direct and indirect effect estimates at low-achieving, medium-achieving, and high-achieving schools. Each pair of error bars represents the 95% confidence interval of the corresponding effect estimate.

same achievement level. The results indicate that the weights effectively removed most of the selection bias associated with observed pretreatment covariates.

Sensitivity Analysis

Based on the assumption that an unmeasured pretreatment confounder is comparable to an observed pretreatment covariate, we assessed the bias that would be generated if any one of the observed pretreatment covariates were omitted from the original analysis. In doing so, we obtained a plausible range of effect sizes for bias as contributed by unmeasured pretreatment confounding, in the population average ITT effect, the natural direct effect, the natural indirect effect, and the treatment-by-mediator interaction effect, as shown in Figure 3. The blue dashed line indicates the mean of plausible bias values in effect size.

The initial point estimate of the effect size of the ITT effect was 0.224. As reasoned in the section introducing the identification assumptions, one may argue that parental engagement might confound the response-mediator or response-outcome relationship. However, this confounder was omitted from the initial analysis, leading to a violation of Assumption (3) and correspondingly a bias in the ITT effect estimate. If the bias that such an omission would contribute were as large as the plausible bias value of the largest magnitude, removing the bias would lead to an estimated ITT effect equal to 0.212 in effect size, almost unchanged from the initial estimate. Hence, we concluded that the original estimate of the ITT effect was insensitive to a potential violation of the identification assumptions. Similarly, the plausible effect sizes of the bias values in the natural direct effect, natural indirect effect, and treatment-by-mediator interaction effect were close to 0, respectively ranging within [−0.003, 0.011], [−0.009, 0.003], and [−0.009, 0.009]. Removing a plausible hidden bias would at most decrease the estimated effect
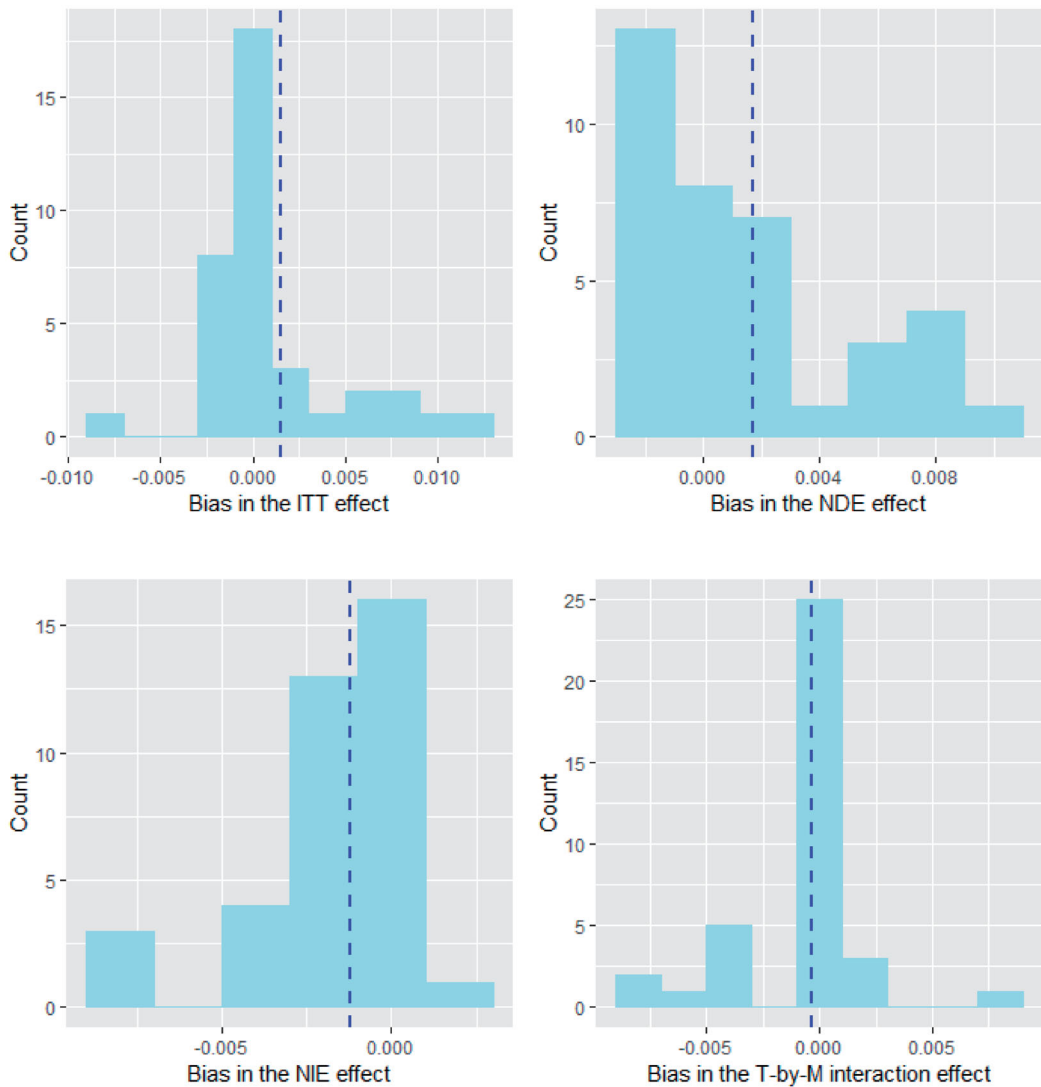
Figure 3. Plausible range of effect size of bias in the population average ITT effect, the natural direct effect, the natural indirect effect, and the treatment-by-mediator interaction effect, due to unmeasured pretreatment confounding. Each blue dashed line indicates the mean of plausible bias values in effect size for the corresponding effect.

size of the natural direct effect by 6% and increase the estimated effect sizes of the natural indirect effect and treatment-by-mediator interaction effect respectively by 21% and 28%. The signs or significance of the effects would always remain unchanged. Therefore, all the causal effects were relatively robust to the omission of an unmeasured pretreatment confounder that is comparable to the observed pretreatment confounders. Because we have adjusted for most of the theoretically important pretreatment confounders in the initial analysis, we believe that the results are not highly sensitive to unmeasured pretreatment confounding (Hong et al., 2018; Shadish et al., 2008). The same conclusion

applies to the between-school standard deviation in each causal effect and the moderating effects.

Discussion

Growth mindset interventions have been proposed as a means by which to promote student academic achievement; however, very few studies have tested theories about the behavioral mechanisms and boundary conditions for growth mindset effects on adolescent achievement outcomes. In particular, the role of educational context in the causal mechanisms through which growth mindset interventions may affect academic achievement has not been considered appropriately. In light of recent empirical evidence from the NSLM study and meta-analytic syntheses revealing that the effects of growth mindset interventions are not uniform across educational contexts (Sisk et al., 2018; Yeager et al., 2019), robust tests of theory-driven mediators and school-level moderators of intervention effects are needed to understand why and in what school contexts growth mindset interventions promote academic achievement. Doing so will yield insights into behavioral mechanisms and between-school differences in underlying mechanisms. It will also guide research and policy-related efforts around creating interventions that not only stimulate students' mindset growth, but also create school environments that support students' learning.

Consequently, this study sought to test the mediating role of challenge-seeking behaviors in the growth mindset intervention's impact across different school contexts. We focused on the critical developmental period of the transition from middle school to high school. As high school coincides with various academic and social stressors (Wang et al., 2019), helping students with this transition should have long-term effects on their educational and career trajectories (Cohen & Sherman, 2014).

We first evaluated the total intervention impact and found that the growth mindset intervention significantly promoted low-achieving 9th graders' academic achievement on average, but this impact varied significantly by schools. These findings align with those reported by Yeager et al. (2019), though our estimates are a bit larger. We estimated the effects in the population of schools, as introduced in the method section, while the target of inference in Yeager et al. (2019) study was the population of students. Even though the effect sizes of the total growth mindset impact are small, any positive effect could be worth the effort since the growth mindset intervention is relatively low-cost and easy to implement for students.

In contrast, the significant intervention impact disappeared among non-low achievers' GPA across all the schools. This finding highlights the importance of examining individual differences in students' responses to mindset interventions and aligns with prior work suggesting that students who are at risk academically stand to benefit the most from a mindset intervention (Lin-Siegler et al., 2016; Paunesku et al., 2015). Historically, academic challenges have jeopardized these students' ability to form positive beliefs about school (Binning et al., 2019). By providing a supportive narrative in which maladaptive beliefs about the school context are addressed or minimized, the growth mindset intervention provides an opportunity to bolster and support low-performing students' self-beliefs and academic behaviors. Researchers should continue this work by identifying other academically vulnerable groups of students in need of mindset

interventions, thereby aiding future efforts to develop more nuanced and tailored interventions.

To understand the underlying mediation mechanism of the intervention impact, we further investigated challenge-seeking as a focal mediator. Challenge-seeking is a malleable target behavior that educators can promote through authentic classroom and school activities, and the growth mindset intervention works directly on students' mindset about efforts in school by reappraising academic challenges and struggles (Yeager & Dweck, 2012). As such, the growth mindset intervention provides reassurance that challenges occur for every new high-school student and suggests that the challenges can be resolved with adequate effort, strategies, and time. When students understand that their academic ability can be improved and these seemingly insurmountable challenges can be overcome, they are better positioned to read negative and ambiguous cues as external and changeable and respond adaptively to stressors and failures. This growth mindset framing encourages students to seek out more challenges rather than avoid them, a behavior that is expected to eventually enhance their academic achievement.

Our analytic results indicated that challenge-seeking transmitted half of the growth mindset intervention impact among low-achieving students from medium-achieving schools, while its mediating role was relatively trivial in low-achieving and high-achieving schools. Researchers have indicated that the adoption and maintenance of a given mindset depend on contextual affordances and meaning-making experiences (Walton & Wilson, 2018; Walton & Yeager, 2020). A growth mindset is not an all-compassing panacea for improving academic achievement. Rather, a growth mindset needs to align with contextual supports—such as the provision of necessary skillsets, resources, and opportunities to experience mastery in their learning—before it is effective (see Yeager et al., 2019). In other words, a growth mindset affords students a strength-based perspective through which they can interpret their learning progress and outcomes. However, this growth mindset can only be maintained, nurtured, and promoted by contextualized messages and supports.

Beyond the theoretical and implementation insights, this study also represents an analytic procedure that may serve as a template for causal mediation analysis in multisite randomized trials. It allows researchers to ask new questions regarding not only the population average of causal mediation mechanisms but also their variations under different contexts. The careful consideration of the mechanisms of sampling, nonresponse, and mediator value selection enhances the external and internal validity of the analytic conclusions. The NSLM study collected a nationally representative sample of high schools in the U.S. so that, with the sampling weights applied, the analysis results would be generalizable to the whole nation. Besides, the randomization of participants to the intervention and control groups ensured the causal interpretation of the intervention impact. However, even within such a careful and thoughtful research design, it is critical to acknowledge that nonresponse would have changed the representativeness of the sample and introduce systematic differences between the intervention and control groups. In addition, because the mediator of challenge-seeking was naturally generated rather than experimentally manipulated, the relationship between the mediator and the achievement outcome might be confounded. To address these issues, we adjusted for the observed pretreatment covariates that confound the response-mediator, response-outcome, and mediator-outcome relationships

through propensity score-based weighting. We further conducted balance-checking to verify that the observed pretreatment covariates have greatly reduced selection bias. According to our sensitivity analysis results, the conclusions are relatively insensitive to unmeasured pretreatment confounding.

Despite its strengths, several limitations provide insights into promising directions for future research. First, our sensitivity analysis assumes that an unmeasured pretreatment confounder is comparable to the observed pretreatment covariates. Even though we have considered most of the theoretically important pretreatment confounders, this does not rule out the possibility of a change in the initial conclusions if the confounding role of an unmeasured pretreatment covariate is much stronger than those of the observed pretreatment covariates. Nevertheless, this is unavoidable in any empirical analysis that involves confounding. Second, we assume no posttreatment confounder of the response-mediator, response-outcome, or mediator-outcome relationship. Hence, the analytic conclusions are exploratory. Adjusting for post-treatment confounders has been considered infeasible in the presence of treatment-by-mediator interaction in the past decades (Avin et al., 2005; Robins, 2003). Researchers have been developing methods to address this issue most recently. Some researchers proposed to impute values of post-treatment confounders under the counterfactual treatment condition (Daniel et al., 2015; Hong et al., 2021). Others defined causal mediation effects under an interventional framework that does not involve cross-world counterfactuals; thus enabling the identification of the effects in the presence of posttreatment confounders (VanderWeele et al., 2014; Wodtke & Zhou, 2020). However, all these strategies were developed under the single-level setting. We leave it to our future research for accounting for posttreatment confounding in multisite causal mediation analysis. Third, even though a relatively important mediating role of challenge-seeking was detected in medium-achieving schools, it is marginally significant, and the effect size is small. This is partly due to the limited number of schools and the small effect size of the total intervention impact. Should more schools participate, and a larger intervention impact is detected, statistical power may be improved for investigating the mediation mechanism and its between-school heterogeneity. Fourth, we focused on a single mediator (i.e., challenge-seeking behaviors) and a single school-level moderator (i.e., average school-level achievement). As methodology continues to advance, future research may consider multiple mediators and moderators for a more thorough understanding of the between-school variation in the mediation mechanism underlying the growth mindset intervention. Fifth, the current study is focused on a single educational time point (i.e., high school transition) and does not follow students across a longer period. It will be fruitful to consider longer-term effects of the growth mindset intervention on academic outcomes in later years of high school and across the postsecondary or education-to-employment transition. Sixth, future work may wish to revise intervention materials based on students' explicit feedback about how they interact with and interpret intervention materials (cf. Yeager et al., 2016).

Should only one message be taken away from this study's findings, it is this: *Context matters*. It is essential that researchers continue this line of inquiry by accounting for contextualized stressors and supports as well as societal and historical barriers to student learning. Future study designs should be sure to consider not only who the participants are, but also the context in which they operate. In doing so, scholars can begin to better understand contextualized patterns of intervention effects. Additional research could assess other contextual factors that may enhance or dampen intervention effects, such as

assessing instructor and administrators' growth mindset beliefs and behaviors (e.g., Canning et al., 2019; Wang et al., 2020). Only by understanding contextualized nuances about which interventions work in which settings for which students can we begin to fully understand that intervention's efficacy and leverage the resources that work in the contexts where children learn and grow.

ORCID

Xu Qin 🆔 http://orcid.org/0000-0002-5907-1511
Alberto Guzman-Alvarez 🆔 http://orcid.org/0000-0002-1027-6612

Data Availability Statement

This manuscript uses data from the National Study of Learning Mindsets (doi:10.3886/ICPSR37353.v1) (PI: D. Yeager; Co-Is: R. Crosnoe, C. Dweck, C. Muller, B. Schneider, & G. Walton), which was made possible through methods and data systems created by the Project for Education Research That Scales (PERTS), data collection carried out by ICF International, meetings hosted by Student Experience Research Network at the Center for Advanced Study in the Behavioral Sciences at Stanford University, assistance from C. Hulleman, R. Ferguson, M. Shankar, T. Brock, C. Romero, D. Paunesku, C. Macrander, T. Wilson, E. Konar, M. Weiss, E. Tipton, and A. Duckworth, and funding from the Raikes Foundation, the William T. Grant Foundation, the Spencer Foundation, the Bezos Family Foundation, the Character Lab, the Houston Endowment, the National Institutes of Health under award number R01HD084772-01, the National Science Foundation under grant number 1761179, Angela Duckworth (personal gift), and the President and Dean of Humanities and Social Sciences at Stanford University.

The content is solely the responsibility of the author(s) and does not necessarily represent the official views of the Bezos Family Foundation, Student Experience Research Network, the University of Texas at Austin Population Research Center, the National Institutes of Health, the National Science Foundation, or other funders.

References

Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology*, *102*(4), 804–816. https://doi.org/10.1037/a0021075

Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38*(2), 113–125. https://doi.org/10.1006/jesp.2001.1491

Avin, C., Shpitser, I., & Pearl, J. (2005). *Identifiability of path-specific effects*. Department of Statistics, UCLA.

Binning, K. R., Wang, M. T., & Amemiya, J. (2019). Persistence mindset among adolescents: Who benefits from the message that academic struggles are normal and temporary? *Journal of Youth and Adolescence*, *48*(2), 269–286. https://doi.org/10.1007/s10964-018-0933-3

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

Boaler, J. (2013). Ability and Mathematics: The mindset revolution that is reshaping education. *FORUM*, *55*(1), 143–152. https://doi.org/10.2304/forum.2013.55.1.143

Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing inequality in academic success for incoming college students: A randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness*, *11*(3), 317–338. https://doi.org/10.1080/19345747.2018.1429037

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, *5*(2), eaau4734. https://doi.org/10.1126/sciadv.aau4734

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. https://doi.org/10.1126/science.1170769

Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology*, *65*, 333–371. https://doi.org/10.1146/annurev-psych-010213-115137

Corpus, J. H., McClintic-Gilbert, M. S., & Hayenga, A. O. (2009). Within-year changes in children's intrinsic and extrinsic motivational orientations: Contextual predictors and academic outcomes. *Contemporary Educational Psychology*, *34*(2), 154–166. https://doi.org/10.1016/j.cedpsych.2009.01.001

Daniel, R., De Stavola, B., Cousens, S., & Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, *71* (1), 1–14. https://doi.org/10.1111/biom.12248

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040–1048. https://doi.org/10.1037/0003-066X.41.10.1040

Dweck, C. S. (1999). *Self-Theories: Their role in motivation, personality and development*. Taylor and Francis/Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Dweck, C. S. (2008). Can personality be changed? The role of beliefs in personality and change. *Current Directions in Psychological Science*, *17*(6), 391–394. https://doi.org/10.1111/j.1467-8721.2008.00612.x

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, *14*(3), 481–496. https://doi.org/10.1177/1745691618804166

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. https://doi.org/10.1016/j.appdev.2003.09.002

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent praise to 1-to 3-year-olds predicts children's motivational frameworks 5 years later. *Child Development*, *84*(5), 1526–1541. https://doi.org/10.1111/cdev.12064

Haimovitz, K., & Dweck, C. S. (2017). The origins of children's growth and fixed mindsets: New research and a new proposal. *Child Development*, *88*(6), 1849–1859. https://doi.org/10.1111/cdev.12955

Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods*, *15*(3), 234–249. https://doi.org/10.1037/a0019623

Heckman, J. J., Smith, J., & Clements, N. (1997). Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *The Review of Economic Studies*, *64*(4), 487–535. https://doi.org/10.2307/2971729

Hong, G. (2010). Ratio of mediator probability weighting for estimating natural direct and indirect effects. In *Proceedings of the American Statistical Association, biometrics section* (pp. 2401–2415). American Statistical Association.

Hong, G. (2015). *Causality in a social world: Moderation, mediation and spill-over.* John Wiley & Sons.

Hong, G., Deutsch, J., & Hill, H. D. (2015). Ratio-of-mediator-probability weighting for causal mediation analysis in the presence of treatment-by-mediator interaction. *Journal of Educational and Behavioral Statistics*, *40*(3), 307–340. https://doi.org/10.3102/1076998615583902

Hong, G., Qin, X., & Yang, F. (2018). Weighting-based sensitivity analysis in causal mediation studies. *Journal of Educational and Behavioral Statistics*, *43*(1), 32–56. https://doi.org/10.3102/1076998617749561

Hong, G., Yang, F., & Qin, X. (2021). Post-treatment confounding in causal mediation studies: A cutting-edge problem and a novel solution via sensitivity analysis. In progress.

Hong, Y., Chiu, C., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–599. https://doi.org/10.1037/0022-3514.77.3.588

Horng, E. (2016). *Findings from the pilot for the National Study of Learning Mindsets.* Mindset Scholars Network.

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88–110. https://doi.org/10.1080/19345740802539325

Irwin, M., & Supplee, L. H. (2012). Directions in implementation research methods for behavioral and social science. *The Journal of Behavioral Health Services & Research*, *39*(4), 339–342. https://doi.org/10.1007/s11414-012-9293-z

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Linnenbrink-Garcia, L., Patall, E. A., & Pekrun, R. (2016). Adaptive motivation and emotion in education: Research and principles for instructional design. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 228–236. https://doi.org/10.1177/2372732216644450

Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F.-F A., & Luna-Lucero, M. (2016). Even Einstein struggled: Effects of learning about great scientists' struggles on high school students' motivation to learn science. *Journal of Educational Psychology*, *108*(3), 314–328. https://doi.org/10.1037/edu0000092

LoCasale-Crouch, J., Williford, A., Whittaker, J., DeCoster, J., & Alamos, P. (2018). Does fidelity of implementation account for changes in teacher–child interactions in a randomized controlled trial of banking time? *Journal of Research on Educational Effectiveness*, *11*(1), 35–55. https://doi.org/10.1080/19345747.2017.1329365

Maehr, M. L., & Zusho, A. (2009). Achievement goal theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (p. 77). Routledge.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*(1), 33–52. https://doi.org/10.1037/0022-3514.75.1.33

Neyman, J., Iwaszkiewicz, K., & Kolodziejczyk, S. (1935). Statistical problems in agricultural experimentation (with discussion). *Supplement to the Journal of the Royal Statistical Society*, *2*(2), 107–180. https://doi.org/10.2307/2983637

Nussbaum, A. D., & Dweck, C. S. (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality & Social Psychology Bulletin*, *34*(5), 599–612. https://doi.org/10.1177/0146167207312960

Olsen, R. B. (2017). Evaluating educational interventions when impacts may vary across sites. *Journal of Research on Educational Effectiveness*, *10*(4), 907–911. https://doi.org/10.1080/19345747.2017.1364554

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. https://doi.org/10.1177/0956797615571017

Pearl, J. (2001). Direct and indirect effects. In J. S. Breese & D. Koller (Eds.), *Proceedings of conference on uncertainty in artificial intelligence* (pp. 411–420). Morgan Kaufmann.

Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In P. J. Green, N. L. Hjort, & S. Richardson (Eds.), *Highly structured stochastic systems* (pp. 70–81). Oxford University Press.

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, *103*(484), 1334–1344. https://doi.org/10.1198/016214508000000733

Qin, X., Deutsch, J., & Hong, G. (2021). Unpacking complex mediation mechanisms and their heterogeneity between sites in a job corps evaluation. *Journal of Policy Analysis and Management*, *40*(1), 158–190. https://doi.org/10.1002/pam.22268

Qin, X., & Hong, G. (2017). A weighting method for assessing between-site heterogeneity in causal mediation mechanism. *Journal of Educational and Behavioral Statistics*, *42*(3), 308–340. https://doi.org/10.3102/1076998617694879

Qin, X., Hong, G., Deutsch, J., & Bein, E. (2019). Multisite causal mediation analysis in the presence of complex sample and survey designs and non-random nonresponse. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *182*(4), 1343–1370. https://doi.org/10.1111/rssa.12446

Raudenbush, S. W., & Bloom, H. (2015). Using multi-site randomized trials to learn about and from a distribution of program impacts. *American Journal of Evaluation*, *36*(4), 475–499. https://doi.org/10.1177/1098214015600515

Raudenbush, S. W., & Schwartz, D. (2020). Randomized experiments in education, with implications for multilevel causal inference. *Annual Review of Statistics and Its Application*, *7*(1), 177–208. https://doi.org/10.1146/annurev-statistics-031219-041205

Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, *3*(2), 143–155. https://doi.org/10.1097/00001648-199203000-00013

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test: comment. *Journal of the American Statistical Association*, *75*(371), 591–593. https://doi.org/10.2307/2287653

Rubin, D. B. (1986). Statistics and causal inference: Comment, which ifs have causal answers. *Journal of the American Statistical Association*, *81*(396), 961–962. https://doi.org/10.2307/2289065

Rubin, D. B. (1990). Formal mode of statistical inference for causal effects. *Journal of Statistical Planning and Inference*, *25*(3), 279–292. https://doi.org/10.1016/0378-3758(90)90077-8

Shively, R. L., & Ryan, C. S. (2013). Longitudinal changes in college math students' implicit theories of intelligence. *Social Psychology of Education*, *16*(2), 241–256. https://doi.org/10.1007/s11218-012-9208-0

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. https://doi.org/10.1177/0956797617739704

Tipton, E. (2014). How generalizable is your experiment? An index for comparing experimental samples and populations. *Journal of Educational and Behavioral Statistics*, *39*(6), 478–501. https://doi.org/10.3102/1076998614558486

Tipton, E., Yeager, D., & Iachan, R. (2016). *Developing a theory of treatment effect heterogeneity through better design: Where do behavioral science interventions work best?* Society for Research on Educational Effectiveness.

Tipton, E., Yeager, D. S., Iachan, R., & Schneider, B. (2019). Designing probability samples to study treatment effect heterogeneity. In P. J. Lavrakas & M. W. Traugott (Eds.), *Experimental methods in survey research: Techniques that combine random sampling with random assignment* (pp. 435–456). Wiley.

VanderWeele, T. J., Vansteelandt, S., & Robins, J. M. (2014). Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*, *25*(2), 300–306. https://doi.org/10.1097/EDE.0000000000000034

Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, *125*(5), 617–655. https://doi.org/10.1037/rev0000115

Walton, G. M., & Yeager, D. S. (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*, *29*(3), 219–226.

Wang, M.-T., & Degol, J. S. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Wang, M.-T., Degol, J. S., Amemiya, J. L., Parr, A., & Guo, J. (2020). Classroom climate and children's academic and psychological wellbeing: A systematic review and meta-analysis. *Developmental Review*, *57*, 100912–100927. https://doi.org/10.1016/j.dr.2020.100912

Wang, M.-T., Degol, J. L., & Henry, D. A. (2019). An integrative development-in-sociocultural-context model for children's engagement in learning. *American Psychologist*, *74*(9), 1086–1102. https://doi.org/10.1037/amp0000522

Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, *28*, 12–23. https://doi.org/10.1016/j.learninstruc.2013.04.002

Weiss, M. J., Bloom, H. S., & Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects. *Journal of Policy Analysis and Management*, *33*(3), 778–808. https://doi.org/10.1002/pam.21760

Weiss, M. J., Bloom, H. S., Verbitsky-Savitz, N., Gupta, H., Vigil, A. E., & Cullinan, D. N. (2017). How much do the effects of education and training programs vary across sites? Evidence from past multisite randomized trials. *Journal of Research on Educational Effectiveness*, *10*(4), 843–876. https://doi.org/10.1080/19345747.2017.1300719

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, *30*(1), 1–35. https://doi.org/10.1016/j.dr.2009.12.001

Wodtke, G. T., & Zhou, X. (2020). Effect decomposition in the presence of treatment-induced confounding: A regression-with-residuals approach. *Epidemiology*, *31*(3), 369–375. https://doi.org/10.1097/EDE.0000000000001168

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. https://doi.org/10.1080/00461520.2012.722805

Yeager, D. S., Hanselman, P., Paunesku, D., Hulleman, C., Dweck, C., Muller, C., & Duckworth, A. (2018). Where and for whom can a brief, scalable mindset intervention improve adolescents' educational trajectories. *PsyArXiv Preprints*.

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

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), 374–391. https://doi.org/10.1037/edu0000098

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*(2), 267–301. https://doi.org/10.3102/0034654311405999