



HHS Public Access

Author manuscript

Am Psychol. Author manuscript; available in PMC 2021 July 23.

Published in final edited form as:

Am Psychol. 2020 December ; 75(9): 1269–1284. doi:10.1037/amp0000794.

What Can Be Learned from Growth Mindset Controversies?

David S. Yeager^{*,1}, Carol S. Dweck²

¹University of Texas at Austin

²Stanford University

Abstract

The growth mindset is the belief that intellectual ability can be developed. This article seeks to answer recent questions about growth mindset, such as: Does a growth mindset predict student outcomes? Do growth mindset interventions work, and work reliably? Are the effect sizes meaningful enough to merit attention? And can teachers successfully instill a growth mindset in students? After exploring the important lessons learned from these questions, the article concludes that large-scale studies, including pre-registered replications and studies conducted by third parties (such as international governmental agencies), justify confidence in growth mindset research. Mindset effects, however, are meaningfully heterogeneous across individuals and contexts. The article describes three recent advances that have helped the field to learn from this heterogeneity: standardized measures and interventions, studies designed specifically to identify where growth mindset interventions *do not* work (and why), and a conceptual framework for anticipating and interpreting moderation effects. The next generation of mindset research can build on these advances, for example, to begin to understand and perhaps change classroom contexts in ways that can make interventions more effective. Throughout, the authors reflect on lessons that can enrich meta-scientific perspectives on replication and generalization.

Keywords

Implicit theories; growth mindset; adolescence; educational psychology; motivation; meta-science

A growth mindset is the belief that personal characteristics, such as intellectual abilities, can be developed, and a fixed mindset is the belief that these characteristics are fixed and unchangeable (Dweck, 1999; Dweck & Leggett, 1988; Yeager & Dweck, 2012). Research on these mindsets has found that people who hold more of a growth mindset are more likely to thrive in the face of difficulty and continue to improve, while those who hold more of a fixed mindset may shy away from challenges or fail to meet their potential (see Dweck & Yeager, 2019). There has been considerable interest among researchers, policymakers, and educators in the use of growth mindset research to improve educational outcomes, in part because research on mindsets has yielded effective, scalable interventions. For instance, the National Study of Learning Mindsets (NSLM; Yeager, 2019) evaluated a short (<1hr), online

*Address correspondence to David Yeager (dyeager@utexas.edu), Department of Psychology and Population Research Center, University of Texas at Austin, 305 E. 23rd Street, Mail Stop G1800, Austin, Texas 78712-1699.

Author contributions: D.S.Y. conceptualized the paper, conducted data analysis, produced data visualizations, and wrote the first draft of the paper. C.S.D. collaborated on the conceptualization and writing of the paper.

growth mindset intervention in a nationally representative sample of 9th graders in the U.S. ($N = 12,490$). Compared to the control condition, the intervention improved grades for lower-achieving students and improved the rate at which students overall chose and stayed in harder math classes (Yeager, Hanselman, Walton, et al., 2019). These or similar effects have appeared in independent evaluations of the NSLM (Zhu et al., 2019) and international replications (Rege et al., in press).

With increasing emphasis on replication and generalizability has come an increased attention to questions of *when, why, and under what conditions* growth mindset associations and intervention effects can be expected to appear. Large trans-disciplinary studies have yielded insights into contextual moderators of mindset intervention effects, such as the educational cultures created by peers or teachers (Rege et al., in press; Yeager, Hanselman, Walton, et al., 2019), and has begun to document the behaviors that can mediate intervention effects on achievement (Gopalan & Yeager, 2020; Qin et al., 2020). At the same time, researchers and practitioners have begun to test a variety of mindset measures, manipulations, and interventions. These have sometimes failed to find growth mindset effects and other times have replicated mindset effects.

The trend toward testing “heterogeneity of effects” and the boundary conditions of effects has coincided with the rise of the meta-science movement. This latter movement has asked foundational questions about the replicability of key findings in the literature (see Nelson et al., 2018). The emergence of these two trends in parallel—increased examination of heterogeneity and increased focus on replicability—has made it difficult to distinguish between studies that call into question the foundations of a field and studies that are a part of the normal accumulation of knowledge about heterogeneity (about where effects are likely to be stronger or weaker; see Kenny & Judd, 2019; Tipton et al., 2019). Many scientists and educational practitioners now wish to know: Which key claims in the growth mindset literature still stand? How has our understanding evolved? And what will the future of mindset research look like?

In this article we examine different controversies about the mindset literature and discuss the lessons that can be learned from them.¹ To be clear, we think controversies are often good. It can be useful to reexamine the foundations of a field. Furthermore, controversies can lead to theoretical advances and positive methodological reforms. But controversies can only benefit the field in the long run when we identify useful lessons from them. This is what we seek to do in the current paper. We highlight why people’s growth versus fixed mindsets or why growth mindset interventions should sometimes affect student outcomes and what we can learn from the times when they do not.

Preview of Controversies

In this paper we review several prominent growth mindset controversies:

¹In the present paper we focus on the intelligence mindsets, which form the core of much of the mindset literature. Of course, people can have mindsets about other characteristics as well (e.g., personality or social qualities; see Schroder et al., 2016) and similar mediators and mechanisms have often emerged in those domains as well.

1. Do mindsets predict student outcomes?
2. Do student mindset interventions work?
3. Are mindset intervention effect sizes too small to be interesting?
4. Do teacher mindset interventions work?

As we discuss below, different authors can stitch together different studies across these issues to create very different narratives. One narrative can lead to the conclusion that mindset effects are not real and that even the basic tenets of mindset theory are unfounded (Li & Bates, 2019; Macnamara, 2018). Another selective narrative could lead to the opposite conclusion. Although we are not disinterested parties, we hope to navigate the evidence in a way that is illuminating for the reader—that is, in a way that looks at the places where mindset effects are robust and where they are weak or absent. Below, we first preview the key concepts at issue in each controversy and the questions we will explore. After that, we will review the evidence for each controversy in turn.

1. Do mindsets predict student outcomes?

A fixed mindset, with its greater focus on validating one's ability and drawing negative ability inferences after struggle or failure, has been found to predict lower achievement (e.g., grades and test scores) among students facing academic challenges or difficulties, compared to a growth mindset with its greater focus on developing ability and on questioning strategy or effort after failure (e.g. Blackwell et al., 2007). However, some studies have not found this same result (e.g. Li & Bates, 2019). Meta-analyses have shown overall significant associations in the direction expected by the theory, but the effects have been heterogeneous and thus call for a greater understanding of where these associations are likely and where they may be less likely (Burnette et al., 2013; Sisk et al., 2018).

Recently, Macnamara (2018) and her co-authors (Burgoyne, Hambrick, & Macnamara, 2020; Sisk, et al., 2018) and Bates and his co-author (Li & Bates, 2019) have looked at this pattern of results and called into question the basic correlations between mindsets and outcomes. Below, we look at the differences across these correlational studies to try and make sense of the discrepant results. We ask: Are the effects found in the large, generalizable studies (i.e. random or complete samples of defined populations), particularly those conducted by independent organizations or researchers? But perhaps more important, how can we know whether a given result is a “failure to replicate” or a part of a pattern of decipherable heterogeneity?

2. Do student mindset interventions work?

The effects of mindset interventions on student outcomes have been replicated but they, too, are heterogeneous (Sisk et al., 2018; Yeager, Hanselman, Walton, et al., 2019). Some studies (and subgroups within studies) have shown noteworthy effects, but other studies or subgroups have not. Macnamara (2018), Bates (Li & Bates, 2019), and several commentators (Shpancer, 2020; Warne, 2020) have argued that this means mindset interventions do not work or are unimportant. Others have looked at the same data and concluded that the effects are meaningful and replicable but they are likely to be stronger or

weaker for different people in different contexts (Gelman, 2018; Tipton et al., 2019). We ask whether the heterogeneity is in fact informative and whether there is an over-arching framework that could explain (and predict in advance) heterogeneous effects.

3. Are mindset intervention effect sizes too small to be interesting?

Macnamara and colleagues have argued that growth mindset intervention effect sizes are too small to be worthy of attention (Macnamara, 2018; Sisk et al., 2018). Others have disagreed (Gelman, 2018; Kraft, 2020; The World Bank, 2017). Here we ask: What are the appropriate benchmarks for the effect sizes of educational interventions? How do mindset interventions compare? And how can an appreciation of the moderators of effects contribute to a more nuanced discussion of “true” effect sizes for interventions (also see Tipton et al., 2020; Vivaldi, 2020)?

4. Do teacher mindset interventions work?

Correlational research has indicated a role for educators’ mindsets and mindset-related practices in student achievement (Canning et al., 2019; Leslie et al., 2015; Muenks et al., 2020). Nevertheless, two recent mindset interventions, delivered to and by classroom teachers, have not had any discernable effects on student achievement (Foliano et al., 2019; Rienzo et al., 2015). We examine the issue of teacher-focused interventions and ask: Why might it be so difficult to coach teachers to instill or support a growth mindset? How can the next generation of mindset research make headway on this?

Controversies

Controversy #1: Do Mindsets Predict Student Outcomes?

Background.

What is mindset theory?: Let us briefly review the theoretical predictions before assessing the evidence. Mindset theory (Dweck, 1999; Dweck & Leggett, 1988; see Dweck & Yeager, 2019) grows out of two traditions of motivational research: attribution theory and achievement goal theory. Attribution theory proposed that people’s explanations for a success or a failure (their attributions) can shape their reactions to that event (Weiner & Kukla, 1970), with attributions of failure to lack of ability leading to less persistent responses to setbacks than attributions to more readily controllable factors, such as strategy or effort (see Weiner, 1985). Research by Diener and Dweck (1978) and Dweck and Reppucci (1973) suggested that students of similar ability could differ in their tendency to show these different attributions and responses. Later, achievement goal theory was developed to answer the question of *why* students with roughly equal ability might show different attributions and responses to a failure situation (Elliott & Dweck, 1988a). This line of work suggested that students who have the goal of validating their competence or avoiding looking incompetent (a performance goal) tend to show more helpless reactions in terms of (ability-focused) attributions and behavior, relative to students who have the goal of developing their ability (a learning goal) (Elliott & Dweck, 1988b; Grant & Dweck, 2003).

The next question was: Why might students of equal ability differ in their tendency toward helpless attributions or performance goals? This is what mindset theory was designed to

illuminate. Mindset theory proposes that situational attributions and goals are not isolated ideas but instead are fostered by more situation-general mindsets (Molden & Dweck, 2006). These more situation-general mindset beliefs about intelligence—whether it is fixed or can be developed—are thought to lead to differences in achievement (e.g., grades and test scores) because of the different goals and attributions in situations involving challenges and setbacks.

In sum, mindset is a theory about responses to challenges or setbacks. It is not a theory about academic achievement in general and does not purport to explain the lion's share of the variance in grades or test scores. The theory predicts that mindsets should be associated with achievement particularly among people who are facing challenges.

How are mindsets measured?: Mindsets are typically assessed by gauging respondents' agreement or disagreement with statements such as “You have a certain amount of intelligence, and you really can't do much to change it” (Dweck, 1999). Greater agreement with this statement corresponds to more of a fixed mindset, and greater disagreement corresponds to more of a growth mindset. Importantly, mindsets are not all-or-nothing—they are conceptualized as being on a continuum from fixed to growth, and people can be at different parts of the continuum at different times. In studies in which survey space has been plentiful, mindsets are often measured with 3–4 items framed in the fixed direction and 3–4 items framed in the growth direction, with the latter being reverse-scored (e.g., Blackwell et al., 2007). In recent large-scale studies in which space is at a premium (e.g. Yeager et al., 2016; 2019), a more streamlined measure of the 2 or 3 strongest fixed-framed items is used for economy, simplicity and clarity. Recently, several policy-oriented research studies have adapted the mindset measures for use in national surveys (e.g., Claro & Loeb, 2019).

Controversy 1a. Do measured mindsets predict academic achievement?

Preliminary evidence.: An initial study (Blackwell et al., 2007, Study 1, $N = 373$) found that students reporting more of a growth mindset showed increasing math grades over the two years of a difficult middle school transition, whereas those with more of a fixed mindset did not (even though their prior achievement did not differ). Later, a meta-analysis by Burnette, O'Boyle, VanEpps, Pollack, & Finkel (2013) summarized data from many different kinds of behavioral tasks and administrative records, in the laboratory and the field, and from many different participant populations, and found that mindsets were related to achievement and performance.

Large-scale studies.: Many large studies conducted by governments and non-governmental organizations have found that mindsets were correlated with achievement. First, all 4th to 7th grade students in the largest districts of California were surveyed as part of the state accountability system, (the “CORE” districts; $N = 300,629$). Growth mindset was associated with higher English/Language Arts scores, $r = .28$, and higher math scores, $r = .27$ (Claro & Loeb, 2019, Table 2, Column 1; also see West et al., 2017).² A follow-up analyses of the

²The Claro and Loeb (2019) study also did an analysis focused on *learning* – i.e. test score *gains*. Such analyses will reduce the estimated correlations with mindset because they would be controlling for any effect of mindset on prior achievement, but these analyses too were significant and in the direction expected by mindset theory.

same dataset found that the association between mindset and test scores was stronger among students who were struggling the most (those who were medium-to-low achieving students; Kanopka et al., 2020), which is consistent with the theory. Next, three large survey studies—including the National Study of Learning Mindsets (NSLM) and the *U-say* study in Norway ($N = 23,446$)—collectively showed a correlation of mindset with high school grades of $r = .24$ (see Figure 2). In Chile, all 10th grade public school students ($N = 168,533$) were asked to complete growth mindset questions in conjunction with the national standardized test. Mindsets were correlated with achievement test scores at $r = .34$, and correlations were larger among students with greater risk of low performance (those facing socioeconomic disadvantages; Claro et al., 2016). In another large study, the Programme for International Student Assessment (PISA), conducted by the OECD (an international organization that is charged with promoting economic growth around the world through research and education), surveyed random samples of students from 74 developed nations ($N = 555,458$) and showed that growth mindset was significantly and positively associated with test scores in 72 of those nations (OECD, 2019; the exceptions were China and Lebanon). (If the three separate regions of China are counted as different nations, the positive associations were significant in 72 out of 76 cases.) The OECD correlations for the U.S. and Chile approximated the previously noted data (see Figure 1).

Unsupportive evidence.: In a recent study of $N = 433$ Chinese students in 5th and 6th grade, Li and Bates (2019) found no significant correlation between students' reported mindsets and their grades or intellectual performance.³ Bahnik and Vranka (2017) found no association between mindsets and aptitude tests in a large sample of university applicants in the Czech Republic ($N = 5,653$). There can be many reasons for these null effects. One can be that the China sample is a smaller sample of convenience and so the results may be less informative. Another can be the choice and translation of the mindset items. However, the PISA data may provide another possible, perhaps more interesting, explanation. Looking at Figure 1, we see that there was no effect in the PISA data for China overall (and there was even a negative effect for the mainland China students); moreover, the Czech Republic was 71st out of 74 OECD nations in the size of their correlation. Thus, the PISA data may suggest that the Li and Bates (2019) and Bahnik and Vranka (2017) correlational studies happened to be conducted in cultural contexts that have the weakest links between mindsets and achievement in the developed world. Of course, this does not mean that every sample from these countries would show a null effect or that every sample from the countries with the higher positive correlations would show a positive effect, but the cross-national results help us to see whether a given study fits a pattern or not.

A recent meta-analysis, presented as unsupportive of mindset theory's predictions, was, upon closer inspection, consistent with the conclusions from the large-scale studies. Macnamara and colleagues meta-analyzed many past studies' effect sizes, with data from different nations, age groups, and achievement levels, and using many different kinds of

³Li and Bates (2019) also tested the effects of intelligence praise on students' motivation, similar to what was done by Mueller and Dweck (1998). The authors reported three experiments, one of which was a significant replication at $p < .05$ and one of which was a marginally significant replication at $p < .10$ (two-tailed hypothesis tests). However, each study was under-powered. The most straightforward analysis with individually under-powered replications is to aggregate the data. When this was done across the three studies, Li and Bates (2019) clearly replicated Mueller and Dweck (1998), $p < .05$.

mindset survey items (Sisk et al., 2018, Study 1). They found an overall significant association between mindsets and academic performance. However, this overall association was much smaller ($r = .09$) than the estimates from the generalizable samples presented above ($r = .27$ to $r = .34$). This puzzle can be solved when we realize that the Sisk et al. (2018) results were heterogeneous ($I^2 = 96.29\%$), and that this caused the random effects meta-analysis to give about the same weight in the average to the effect sizes of very small, outlying study from convenience samples as it gave to the generalizable studies. Thus, the data from three generalizable datasets (the CORE data, the Chile data, and the U.S. National Study of Learning Mindsets), which contributed 296,828, or 81%, of all participant data in the Sisk et al. (2018) meta-analysis, were given roughly the same weight as correlations with a sample size of $n = 10$ or fewer participants (of which there were 19).

Conclusions.—There is a replicable and generalizable association between mindsets and achievement. A lesson from this controversy is to underscore that this association cannot be captured with a single, summary effect size. Mindset associations with outcomes were expected to be, and often were, stronger among people facing academic difficulties or setbacks. Further, there was (and always will be) some unexplained heterogeneity across cultures and almost certainly within cultures as well. This unexplained heterogeneity should be a starting point for future theory development. To quote Kenny and Judd (2019), “rather than seeking out the ‘true’ effect and demonstrating it once and for all, one approaches science ... with a goal of understanding all the complexities that underlie effects and how those effects vary.” (pg. 10).

In line with this recommendation, we are deeply interested in cultures in which a growth mindset does *not* predict higher achievement, such as mainland China. Multi-national probability sample studies are uniquely suited to investigate this. On the PISA survey, students in mainland China reported spending 57 hours per week studying (OECD, 2019, Figure 1.4.5), the second-most in the world. Perhaps a growth mindset cannot increase hours of studying or test scores any further when there is already a cultural imperative to work this hard. But that does not mean mindset has no effect in such cultures. Mindset has an established association with mental health and psychological distress (Burnette et al., 2020; Schleider et al., 2015; Seo et al., in press). Interestingly, of all OECD nations, mainland China had the *strongest* association between fixed mindset and “fear of failure,” a precursor to poor mental health (OECD, 2019, Table III.2). This suggests that perhaps a growth mindset might yet have a role to play in student well-being there.

Controversy #1b. Does the mindset “meaning system” replicate?—Burgoyne et al. (2020) have claimed that the evidence for the mindset “meaning system” is not robust. First, we review the predictions and measures, and then we weigh the evidence.

Meaning system predictions.: The two mindsets are thought to orient students toward different goals and attributions in the face of challenges and setbacks. Research has also found (e.g., Blackwell et al., 2007) that the two mindsets can orient students toward different interpretations of effort. For this reason, we have suggested that each mindset fosters its own “meaning system.” More specifically, mindset theory expects that the belief in fixed ability should be associated with a meaning system of *performance goals* (perhaps particularly

performance-avoidance goals, which is the goal of avoiding situations that could reveal a lack of ability), *negative effort beliefs* (e.g., the belief that the need to put forth effort on a task reveals a lack of talent), and *helpless attributions* in response to difficult situations (e.g., attributing poor performance to a stable flaw in the self, such as being “dumb”). A growth mindset meaning system is the opposite, involving learning goals, positive effort beliefs, and resilient (e.g. strategy-focused) attributions (see Dweck & Yeager, 2019). Thus, the prediction is that measured mindsets should be associated with their allied goals, effort beliefs, and attributions.

Meaning system measures.: In evaluating the robustness of the meaning system, one must carefully examine the measures used to see whether they capture the spirit and the psychology of the meaning system in question. Indeed, as a meta-analysis by Hulleman points out, research on achievement goals has often used “the same label for different constructs” (2010, pg. 422). There are two very different types of performance goal measures: those focused on *appearance/ability* (i.e. validating and demonstrating one’s abilities) and those focused on living up to a *normative standard* (e.g. doing well in school) (also see Grant & Dweck, 2003). These two different types of performance goal measures have different (even reversed) associations with achievement and with other motivational constructs (see Table 13, pg. 438, in Hulleman et al., 2010). Mindset studies have mostly asked about appearance-focused achievement goals (e.g., “It’s very important to me that I don’t look stupid in [this math] class,” as in Blackwell et al. 2007, Study 1; also see Table 1). However, correlations of mindset with normative achievement goals (the desire to do well in school) are not expected. Next, learning goals measures do not simply ask about the value of learning in general but ask about it in the context of a tradeoff (e.g., in Blackwell et al. 2007: “It’s much more important for me to learn things in my classes than it is to get the best grades;” also see Table 1). The reason is that almost anyone, even students with more of a fixed mindset, should be interested in learning, but mindsets should start to matter when people might have to sacrifice their public image as a high-ability person in order to learn. Finally, attributions are typically assessed in response to a failure situation (see Robins & Pals, 2002 and Table 1), which is where the two mindsets part ways.

Initial evidence.: The first study to simultaneously test multiple indicators of the mindset meaning system was conducted by Robins and Pals (2002) with college students ($N = 508$). They found that a fixed mindset was associated with performance goals and fixed ability attributions for failure, as well as helpless behavioral responses to difficulty, $r_s = .31, .19,$ and $.48,$ respectively (effort beliefs were not assessed). Later, Blackwell et al. (2007, Study 1, $N = 373$) replicated these findings, adding a measure of effort beliefs.

Each meaning system hypothesis has also been tested individually in experiments which directly manipulated mindsets through persuasive articles (without mentioning any of the meaning system constructs). In a study of mindsets and *achievement goals*, Cury, Elliot, Da Fonseca, and Moller (2006, Study 2) manipulated mindsets, without discussing goals, and showed that this caused differences in performance vs. learning/mastery goals in the two mindset conditions. In a study of mindsets and *effort beliefs*, Miele and Molden (2010, Study 2) crossed a manipulation of mindsets (fixed versus growth) with a manipulation of

how hard people had to work to interpret a passage of text.⁴ Participants induced to hold more of a fixed mindset lowered their competence perceptions when they had to work hard to interpret the passage, and increased their competence perceptions when the passage was easy to interpret, relative to those induced to hold a growth mindset (who did not differ in their competence perceptions in the hard vs. easy condition). It is also important to note that actual performance and effort were constant across the mindset conditions, but mindsets caused different appraisals of the meaning of the effort they expended. Finally, studies of mindsets and *resilient versus helpless reactions* difficulty have manipulated mindsets without mentioning responses to failure. These have shown greater *resilience* among those induced with a growth mindset (e.g. attributing failure to effort and/or seeking out remediation) and greater *helpless responses* among those induced with a fixed mindset (Hong et al., 1999; Nussbaum & Dweck, 2008).

Meta-analysis.: A meta-analysis by Burnette et al. (2013) synthesized the experimental and correlational evidence on the meaning system hypotheses using data from $N = 28,217$ participants. Consistent with mindset theory, the effects of mindset on achievement goals (performance versus learning) and on responses to situations (helpless versus mastery) were replicated and were statistically-significantly stronger (by approximately 50% on average) when people were facing a setback (what the authors called “ego threats”; Burnette et al., 2013).

Large-scale studies using standardized measures.: Recently, three different large studies with a total of over 23,000 participants have replicated the meaning system correlations (see Figure 2 and Table 1), two of which were generalizable to entire populations of students, either to the U.S. overall, (Yeager et al., 2019) or to entire parts of Norway (Rege et al., in press).

These studies used the “Mindset Meaning-System Index” (MMI) a standardized measure comprised of single items for each construct. To create the MMI, from past research (e.g., Blackwell et al., 2007) we chose the most prototypical or paradigmatic items that loaded highly onto the common factor for that construct and covered enough of the construct that it could stand in for the whole variable, Meeting the criteria set forth above, the performance goal item was appearance/ability focused, the learning goal involved a tradeoff (e.g. learning but risking your high grades vs. not learning but maintaining a high-ability image), and the attribution items involved a response to a specific failure situation. Each item was then edited to be clear enough to be administered widely, especially to those with different levels of language proficiency.

Understanding an unsupportive study.: In a study with $N = 438$ undergraduates, Burgoyne et al., (2020) found weak or absent correlations between their measure of mindsets and their measures of some of the meaning system variables (performance goals,

⁴Miele and Molden (2010) note that “the two versions of the [mindset-inducing] article focused solely on whether intelligence was stable or malleable and did not include any information about the role of mental effort or processing fluency in comprehension and performance” (pg. 541).

learning goals, and attributions), none of which met their $r = .20$ criterion for a meaningful effect. (They did not include effort beliefs.)

The data from Burgoyne et al. (2020) are thought-provoking. They differ from the data provided by Robins and Pals (2002), Blackwell et al. (2007) and by the participants just mentioned. Indeed, Table 1 shows that 10 out of 15 of the correlations in the three large replications exceeded the $r = .20$ threshold set by Burgoyne et al., even without adjusting for the unreliability in single-item measures that can attenuate the magnitude of associations.

One solution to the puzzling discrepancy is that Burgoyne et al. (2020) did not adhere closely to the measures used in past studies. The nature of these deviations can be instructive for theory and future research. First, Burgoyne et al. (2020) asked about normative-focused performance goals (“I worry about the possibility of performing poorly”), not appearance/ability-focused goals (cf. Grant & Dweck, 2003; Hulleman et al., 2010). Students with more of a growth mindset may well be just as eager to do well in school (and not do poorly) as those with more of a fixed mindset (indeed, endorsement of their item was near ceiling). Where they should differ is in whether they are highly concerned with maintaining a public (or private) image of being a “smart” not a “dumb” person. Second, the learning goals item used by Burgoyne and colleagues did not include any tradeoff (it simply asked “I want to learn as much as possible”) and it would not be expected to correlate with mindset. Third, the Burgoyne et al. (2020) study included a single attributional item which did not ask about an interpretation of any failure event (it simply said “Talent alone—without effort—creates success”), unlike Robins and Pals (2002).

Conclusion.—The Burgoyne et al. (2020) study was instructive because it provided an opportunity to articulate the aspects of the constructs that are most central to mindset theory. In addition, it showed the need for standardized measures, not only of the mindsets, but also of their meaning system mediators. Being short, the MMI is not designed to maximize effect sizes in small studies, but rather is designed to provide a standardized means for learning about where (and perhaps for testing hypotheses about why) the meaning system variables should be more strongly or weakly related to mindsets, to each other, and to outcomes.

Controversy #2: Do Student Mindset Interventions Work?

Based on their meta-analysis of mindset interventions, Macnamara and colleagues (Sisk et al., 2018; Macnamara, 2018) concluded that there is only weak evidence that mindset interventions work or yield meaningful effects. Here, we first clarify what is meant by a growth mindset intervention. Then we ask: Do the effects replicate, and hold up in pre-registered studies and independent analyses? And are the effects *meaningfully* heterogeneous—that is, do the moderation analyses reveal new insights, mechanisms, or areas for future research?

What is a growth mindset intervention?

The intervention’s core content. A growth mindset intervention teaches the idea that people’s academic and intellectual abilities can be developed through actions they take (e.g. effort, changing strategies, and asking for help) (Yeager, Romero, et al., 2016; Yeager &

Dweck, 2012). The intervention usually conveys information about neuroplasticity through a memorable metaphor; the NSLM, for instance, stated “the brain is like a muscle—it gets stronger (and smarter) when you exercise it.” Merely defining and illustrating the growth mindset with a metaphor could never motivate sustained behavior change. Thus, the intervention must also mention concrete actions people can take to implement the growth mindset, such as “You exercise your brain by working on material that makes you think hard in school.” Students also heard stories from scientists, peers, and notable figures who have used a growth mindset. Importantly, the intervention is not a passive experience but invites active engagement. In the NSLM, for example, students wrote short essays about times they have grown their abilities after struggling and how they might use a growth mindset for future goals; and they wrote a letter to encourage a future student who had fixed mindset thoughts such as “I’m already smart, so I don’t have to work hard” or “I’m not smart enough, and there’s nothing I can do about it.” This is called a “saying-is-believing” exercise and it can lead students to internalize the growth mindset relatively even in a short time, through cognitive dissonance processes (Aronson et al., 2002).

Crucial details.: Several details are key to the intervention’s effectiveness. A growth mindset is not simply the idea that people can get higher scores if they try harder. To count as a growth mindset intervention, it must make an argument that ability itself has the potential to be developed. Telling students that they succeeded because they tried hard, for example, is an attribution manipulation, not a growth mindset intervention (cf. Li & Bates, 2019). A growth mindset intervention also does not require one to believe that ability can be greatly or easily changed. It deals with the *potential* for change, without making any claim or promise about the magnitude or ease of that change. Nor does a growth mindset intervention say that ability does not matter or does not differ among students. Rather, it focuses on within-person comparisons—the idea that people can improve relative to their prior abilities.

Finally, growth mindset interventions can be poorly-crafted or well-crafted (see Yeager et al, 2016, for head-to-head comparisons of different interventions). Poorly-crafted interventions often tell participants the definition of growth mindset without suggesting how to put it into practice. As noted above, a definition alone cannot motivate behavior change. Well-crafted interventions ask students to reflect on how they might develop a “stronger” (better-connected) brain if they do challenging work, seek out new learning strategies, or ask for advice when it is needed, but they do not train students in the mediating behaviors or meaning system constructs. That is, the interventions mention these learning-oriented behaviors at a high level, so that students do not just have a definition of a growth mindset, but also have an idea of how to put it into action if they wish to do so. In fact, in some mindset experiments, the control group gets skills training (e.g. Blackwell et al., 2007; Yeager, Trzesniewski, & Dweck, 2013). Well-crafted interventions are also autonomy-supportive and are not didactic (see Yeager et al., 2018).

Crafting the interventions.: Recently, as the interventions have been evaluated in larger-scale experiments, a years-long research and development process has been followed, which included a) focus groups with students to identify arguments that will be persuasive to the

target population, b) A/B tests to compare alternatives head to head, and c) pilot studies in pre-registered replication experiments (see Yeager et al., 2016). When the intervention reliably changed mindset beliefs and short-term mindset-relevant behavior (such as challenge-seeking) consistently across gender, racial, ethnic, and socioeconomic groups, then it was considered ready for an evaluation study testing its effects on academic outcomes in the target population.

Intervention target.: Here we focus on growth mindset interventions that target students and that deliver treatments directly to students without using classroom teachers as intermediaries. We consider teacher-focused interventions below (see controversy #4).

Do student growth mindset intervention effects replicate?

Initial evaluation experiments.: Blackwell and colleagues (2007, Study 2) evaluated an in-person growth mindset intervention, delivered by highly trained facilitators (not teachers) who were trained extensively before each session and debriefed fully after each one. They found that it halted the downward trajectory of math grades among U.S. 7th grade students who had previously been struggling in school ($N = 99$). Although promising, the in-person intervention was not scalable because it required extensive training of facilitators and a great deal of class time. Later, Paunesku and colleagues (2015) administered a portion of that face-to-face intervention in an online format to U.S. high school students ($N = 1,594$). It included a scientific article conveying the growth mindset message, followed by guided reading and writing exercises designed to help students internalize the mindset. Using a double-blind experimental design, Paunesku and colleagues (2015) found effects of the intervention on lower-achieving students' grades months later, at the end of the term—a result which mirrors the findings of many correlational studies showing larger effects for more vulnerable groups (noted above). The effect size was, not surprisingly, smaller in a short, online self-administered intervention than in an in-person intervention, but the conclusions were similar and the online mode was more scalable.

Yeager, Walton and colleagues (2016, Study 2) replicated the Paunesku et al. study (using almost identical materials), but focusing on undergraduates ($N = 7,418$) and a different outcome (full-time enrollment rather than grades). They administered a version of the Paunesku et al. (2015) short online growth mindset intervention to all entering undergraduates at a large public university and found that it increased full-time enrollment rates among vulnerable groups. Broda and colleagues (Broda et al., 2018) carried out a replication ($N = 7,686$) of the Yeager, Walton et al. (2016, Study 2) study and found a similar result: growth mindset intervention effects among a subgroup of students that was at risk for poor grades.

First pre-registered replication.: The Paunesku et al. study was replicated in a pre-registered trial conducted with high school students (Yeager, Romero, et al., 2016). That study improved upon the intervention materials by updating and refining the arguments (see Study 1, Yeager et al., 2016) and more than doubled the sample size (3,676), which matches recommendations for replication studies (Simonsohn, 2015). This Yeager et al. (2016) study found effects for lower-achieving students. Of note, Pashler and Rutter (2017) have argued

that psychologists should reserve strong claims for “class 1” findings—those that are replicated in pre-registered studies. With the 2016 study, growth mindset met their “class 1” threshold.

A national, pre-registered replication. In 2013 we launched the National Study of Learning Mindsets (NSLM) (Yeager, 2019), which is another pre-registered randomized trial, but now using a nationally-representative sample of U.S. public high schools. To launch the NSLM, a large team of researchers collectively developed and reviewed the intervention materials and procedures to ensure that the replication was faithful and high-quality. Then, an independent firm specializing in nationally-representative surveys (ICF International) constructed the sample, trained their staff, guided school personnel, and collected and merged all data. Next, a different independent firm, specializing in impact evaluation studies (MDRC), constructed the analytic dataset, and wrote their own independent evaluation report based on their own analyses of the data. Our own analyses (Yeager et al., 2019) and the independent report (Zhu et al., 2019) confirmed the conclusion from prior research: there was a significant growth mindset intervention effect on the grades of lower-achieving students.

Effects on course-taking. Of note, an exploratory analysis of the NSLM found that the growth mindset intervention also increased challenge-seeking across achievement levels, as assessed by higher rates of taking advanced math (Algebra II or above). This exploratory finding was later replicated in the Norway *U*-say experiment ($N = 6,541$, Rege et al., in press) with an identical overall effect size.

Summary of supportive replications. Taken together, these randomized trials established the effects of growth mindset materials with more than 40,000 participants and answered the question of whether growth mindset interventions can, under some conditions, yield replicable and scalable effects for vulnerable groups. But this does not mean that growth mindset interventions work everywhere for all people. Indeed, there were sites within the NSLM where the intervention did not yield enhanced grades among lower achievers.

Unsupportive evidence. Rienzo, Rolfe, and Wilkensom (2015) evaluated a face-to-face growth mindset intervention in a sample of 5th grade students ($N = 286$), resembling the face-to-face method in Blackwell et al. (2007). They showed a non-significant positive effect, namely, a 4-month gain in academic achievement, $p = .07$, in the growth mindset group relative to the controls. Interestingly, the estimated effect size (and pattern of moderation results by student prior achievement reported by Rienzo et al., 2015) was larger than the online growth mindset intervention effects. Therefore the Rienzo study is not exactly evidence *against* mindset effects (see a discussion of this statistical argument in McShane et al., 2019). Rienzo et al., 2015, also reported a second, teacher-focused intervention with null effects (see controversy #4 below).

Several other studies have evaluated direct-to-student growth mindset manipulations (see Sisk et al., 2018 for a meta-analysis; also see the appendix in Ganimian, 2020). Some of these relied on non-randomized (i.e., quasi-experimental) research designs, and others were randomized trials. Some were not growth mindset interventions (e.g., involving, for

example, only an email from a professor or sharing a story of scientist who overcame great struggle). These were combined in a meta-analysis (Sisk et al., 2018). This meta-analysis yielded the same conclusion as the NSLM: an overall, significant mindset intervention effect that was larger among students at risk for poor performance. However, as with their meta-analysis of the correlational studies, the Sisk et al. (2018) meta-analysis yielded significant heterogeneity.

Is the heterogeneity in mindset effects *meaningful*?—Learning from the heterogeneity in any psychological phenomenon, especially one involving real-world behavior and policy-relevant outcomes, can be difficult (Tipton et al., 2019). In particular, it is hard to understand the source of heterogeneous findings when different studies involve different populations, different interventions, and different contexts all at once. Because meta-analyses tend to assess moderators at the study level (rather than analyzing within-study interaction effects), and because a given study tends to change all three—the population, the intervention, and the context—then moderators are confounded. For this reason, meta-regression is often poorly-suited for understanding moderators of interventions (see a discussion in Yeager et al., 2015).

It is far easier to find out if contextual heterogeneity is meaningful if a study holds constant key study design features (the intervention and the targeted group) and then carries out the experiment in different contexts (Tipton et al., 2019). If effects still varied, this would mean that at least some of the heterogeneity we were seeing was systematic and potentially interesting. Indeed, large, rigorous randomized trials have a long history of settling debates caused by meta-analyses that aggregated mostly correlational or quasi-experimental studies (see, e.g., the class size debate, Krueger, 2003).

Heterogeneous effects in the NSLM.: The NSLM was designed to provide exactly this kind of test, by unconfounding contextual heterogeneity from intervention design. It revealed that growth mindset intervention effects were indeed systematically larger in some schools and classrooms than others. In particular, the intervention improved lower-achieving students' grade point averages chiefly when *peers* displayed a norm of challenge-seeking (a growth mindset behavior; Yeager et al., 2019) and math grades across achievement levels when math *teachers* endorsed more of a growth mindset (Yeager et al., under review). Thus, the intervention (which changed growth mindset beliefs homogeneously across these settings), was most effective at changing grades in *populations* who were vulnerable to poor outcomes, and in *contexts* where peers and teachers afforded students the chance to act and continue acting on the mindset (see Walton & Yeager, 2020).

A proposed framework for understanding heterogeneity: The Mindset \times Context perspective.: The NSLM results led to a framework that can help to understand (and predict, in advance) heterogeneous results. We call it the *Mindset \times Context* perspective (see Figure 3) (Yeager, Hanselman, Muller, & Crosnoe, 2019). According to this view, a mindset intervention should have meaningful effects only when people are actively facing challenges or setbacks (e.g., when they are lower-achieving, or are facing a difficult course or school transition) *and* when the context provides opportunities for students to act on their mindsets

(e.g., via teacher practices that support challenge-seeking or risking mistakes) (Bryan et al., 2020; Walton & Yeager, 2020). Mindset \times Context stands in contrast to the “*mindset alone*” perspective, which is the idea that if people are successfully taught a growth mindset they will implement this mindset in almost any setting they find themselves in.

A key implication of Mindset \times Context theory is that the intervention will likely need further customization before it can be given to different populations (e.g., in the workplace, for older or younger students, or in a new domain). And even a well-crafted intervention will need to be delivered with an understanding of the context factors that could moderate its effects.

We note that future research using a typical convenience sample will have a hard time testing predictions of Mindset \times Context Theory. Individual-level risk factors and unsupportive context factors are often positively correlated in the population, but the two types of factors are expected to have opposite moderation effects. This is why it is useful to use either representative samples (like the NSLM) or carefully constructed quota samples which disentangle the two.

Conclusions.—The NSLM results show why it can be misleading to look at studies with null findings and conclude that something “doesn’t work” or isn’t “real” (also see Kenney & Judd, 2019). In fact, if researchers had treated each of the 65 schools in the NSLM as its own separate randomized trial, there could be many published papers in the literature that might look like “failures to replicate.” This is why it is not best practice to count significant results (see Gelman, 2018, for a discussion of this topic). When, instead, we had specific hypotheses about meaningful sources of heterogeneity and a framework for interpreting them (Figure 3) we found that the schools varied in their effects in systematic ways and for informative reasons. There is still much more to learn about heterogeneity in growth mindset interventions, of course, in particular with respect to classroom cultures and international contexts (cf. Ganimian, 2020).

Controversy #3: Are Mindset Effect Sizes Too Small to Be Interesting?

What are the right benchmarks for effect sizes?—Macnamara and colleagues stated that a “typical” effect size for an educational intervention is $.57 SD$ (Macnamara, 2018; Sisk et al., 2018, pg. 569). To translate this into a concrete effect, this would mean that if GPA had a standard deviation of 1, as it usually does, then a typical intervention should be expected to increase GPA by $.57$ points, for instance from a 3.0 to a 3.57. Macnamara and colleagues argue that “resources might be better allocated elsewhere” because growth mindset effects were much smaller than $.57 SD$ (Sisk et al., 2018, pg. 569).

Questioning Macnamara (2018), Gelman (2018) asked “Do we really believe that $.57$? Maybe not.” The $.57 SD$ benchmark comes from a meta-analysis by Hattie, Biggs, and Purdie (1996) of learning skills manipulations. Their meta-analysis mostly aggregated effects on variables at immediate post-test, almost all of which were researcher-designed measures to assess whether students displayed the skill they had just learned (almost like a manipulation check). However such immediate measures that are very close to precisely what was taught are well-known to show much larger effect sizes than multiply-determined

outcomes that unfold over time in the real world, like grades or test scores (Cheung & Slavin, 2016). Indeed, in the same Hattie meta-analysis cited by Macnamara (2018), Hattie and colleagues stated: “There were 30 effect sizes that included follow-up evaluations, typically of 108 days, and the effects sizes declined to an average of .10” (Hattie et al., 1996, pg. 112). Judging from the meta-analysis cited by Macnamara (2018), the “typical” effect size for a study looking at effects over time would be .10 *SD*, not .57 *SD*. But even this is too optimistic when we consider study quality.

The current standard for understanding effect sizes in educational research does not look to a meta-analysis of all possible studies regardless of quality but looks to “best evidence syntheses.” This means examining syntheses of studies using the highest-quality research designs that were aimed at changing real-world educational outcomes (Slavin, 1986). This is important because lower-quality research designs (non-experimental, non-randomized, or non-representative) do not provide realistic benchmarks against which interventions should be evaluated (Cheung & Slavin, 2016). Through the lens of the best evidence synthesis, about the *best* effect that can be expected for an educational intervention with real-world outcomes is .20 *SD*. For instance, the effect of having an exceptionally good vs. a below-average teacher for an entire year is .16 *SD* (Chetty et al., 2014). An entire year of math learning in high school is .22 *SD* (Lipsey et al., 2012). The positive effect of drastically reducing class-size for elementary school students was .20 *SD* (Nye et al., 2000).

A “typical” effect of educational interventions is much smaller. Kraft (2020) located all studies funded by the federal government’s Investing in Innovation (I3) Fund, which were studies of programs that had previously shown promising results and won a competition to undergo rigorous, third-party evaluation (Boulay et al., 2018). This is relevant because researchers had to pre-specify their primary outcome variable and hypotheses and report the results regardless of their significance, so there is no “file drawer” problem. These were studies looking at objective real-world outcomes (e.g. grades, test scores, course completion rates) some time after the educational interventions. The median effect size was .03 *SD* and the 90th percentile was .23 *SD* (See Kraft, 2020, Table 1, rightmost column). In populations of adolescents (the group targeted by our growth mindset interventions), there were no effects over .23 *SD* (see an analysis in the online supplement in Yeager et al., 2019). Kraft (2020) concluded that “effects of 0.15 or even 0.10 *SD* should be considered large and impressive when they arise from large-scale field experiments that are preregistered and examine broad achievement measures” (pg. 248).

It is interesting to note that the most highly-touted “nudge” interventions (among those whose effects unfold over time) have effects that are in the same range (see Benartzi et al., 2017). In prominent experiments, the effects of descriptive norms on energy conservation (.03 *SD*), the effects of implementation intentions on flu vaccination rates (.12 *SD*), and the effects of simplifying a financial aid form on college enrollment for first-time college-goers (.16 *SD*) (effect sizes calculated from posted *z* statistics, <https://osf.io/47f7j/>) were never anywhere close to .57 *SD* overall.

Summary of effect sizes.—We have presented this review to suggest that psychology’s effect size benchmarks, which were based on expectations of laboratory results and are often

still used in our field, have led the field astray (also see a discussion in Kraft, 2020). In the real world, single variables do not have huge effects. Not even relatively large, expensive, and years-long reforms do. If psychological interventions can get a meaningful chunk of a .20 effect size on real-world outcomes in targeted groups, reliably, cost-effectively, and at scale, that is impressive.

Comparison to the NSLM.—In this context, growth mindset intervention effect sizes are noteworthy. The NSLM showed an average effect on grades in the pre-registered group of lower-achieving students of .11 *SD* (Yeager et al, 2019). Moreover, this occurred with a short, low-cost self-administered intervention that required no further time investment from the school (indeed, teachers were blind to the purpose of the study). And these effects were in a scaled-up study.

Growth mindset effects were larger in contexts that were specified in pre-registered hypotheses. The effect size was .18 (for overall grades) and .25 (for math and science grades) in schools that were not already high-achieving and provided a supportive peer climate (Yeager et al., 2019, Figure 3). Of course, we are aware that these kinds of moderation results might, in the past, have emerged from a *post-hoc* exploratory analysis and would therefore be hard to believe. So it might be tempting to discount them. But these patterns emerged from a disciplined pre-analysis plan that was carried out by independent Bayesian statisticians who analyzed blinded data using machine-learning methods (Yeager et al., 2019), and the moderators were confirmed by an independent research firm's analyses, over which we had no influence (Zhu et al., 2019).

Effects of this magnitude can have important consequences for students' educational trajectories. In the NSLM, the overall average growth mindset effect on lower-achievers' 9th grade poor performance rates in their courses (that is, the earning of D/F grades) was a reduction of 5.3 percentage points (Yeager et al., 2019). Because there are over 3 million students per year in 9th grade (and lower-achievers are defined as the bottom half of the sample), this means that a scalable growth mindset intervention could prevent 90,000 at-risk students per year from failing to make adequate progress in the crucial first year of high school. In sum, the effect sizes were meaningful—they are impressive relative to the latest standards and in terms of potential societal impact—and the moderators meet a high bar for scientific credibility.

Controversy 4: Teacher Mindset Interventions

The final controversy concerns mindset interventions aimed at or administered by teachers. So far, teacher-focused growth mindset interventions have not worked (Foliano et al., 2019; Rienzo et al., 2015), even though they were developed with great care and were labor-intensive. A likely reason is that the evidence base for teacher-focused interventions is just beginning to emerge. Among other things, the field will need to learn 1) precisely how to address teachers' mindsets about themselves and their students, 2) which teacher practices feed into and maintain students' fixed and growth mindsets, 3) how to guide and alter the teachers' practices, and 4) how to do so in a way that affects students' perceptions and

behaviors and that enhances students' outcomes. Moreover, changing teacher behavior through professional development is known to be exceptionally challenging (TNTP, 2015).

For this reason, it might be preferable to start by administering a direct-to-student program to teach students a growth mindset, such as the (free-to-educators) program we have developed (available at www.perts.net). Then the focus can be on helping teachers to support its effects. As the field begins to tackle this challenge, it will not have to start from scratch but can build on recent studies documenting the role of teachers' mindsets and mindset-related practices in student achievement (e.g. Canning et al. 2019).

Summary of the Controversies

Three of the questions we have addressed so far—Does growth mindset predict outcomes? Do growth mindset interventions effects replicate? Are the effect sizes meaningful?—have strong evidence in the affirmative. In each case we have been inspired to learn from critiques, for instance, by learning more about the expected effect sizes in educational field experiments, or designing standardized measures and interventions. There is also evidence that speaks to the meaningful heterogeneity of the effects. As we have discussed, there are studies, or sites within studies, that do not show predicted mindset effects, but the more we are learning about the students and contexts at those sites the more we can improve mindset measures and intervention programs. The fourth controversy, about educational practitioners, highlights a key limitation of the work to date and points to important future directions.

Conclusion

This review of the evidence showed that the foundations of mindset theory are sound, the effects are replicable, and the intervention effect sizes are promising. Although we have learned much from the research-related controversies, we might ask at a more general level: Why should the idea that students can develop their abilities be controversial? And why should it be controversial that believing this can inspire students, in supportive contexts, to learn more? In fact, don't all children deserve to be in schools where people believe in and are dedicated to the growth of their intellectual abilities? The challenge of creating these supportive contexts for all learners will be great, and we hope mindset research will play a meaningful role in their creation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The writing of this essay was supported by the William T Grant Foundation Scholars Award, the National Institutes of Health (R01HD084772 and P2CHD042849), and the National Science Foundation (1761179). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

- Aronson JM, Fried CB, & 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. 10.1006/jesp.2001.1491
- Bahník Š, & Vranka MA (2017). Growth mindset is not associated with scholastic aptitude in a large sample of university applicants. *Personality and Individual Differences*, 117, 139–143. 10.1016/j.paid.2017.05.046
- Benartzi S, Beshears J, Milkman KL, Sunstein CR, Thaler RH, Shankar M, Tucker-Ray W, Congdon WJ, & Galing S (2017). Should governments invest more in nudging? *Psychological Science*, 28(8), 1041–1055. 10.1177/0956797617702501 [PubMed: 28581899]
- Blackwell LS, Trzesniewski KH, & Dweck CS (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246–263. 10.1111/j.1467-8624.2007.00995.x [PubMed: 17328703]
- Boulay B, Goodson B, Olsen R, McCormick R, Darrow C, Frye M, Gan K, Harvill E, & Sarna M (2018). The investing in innovation fund: Summary of 67 evaluations (NCEE 2018–4013). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Broda M, Yun J, Schneider B, Yeager DS, Walton GM, & 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. 10.1080/19345747.2018.1429037
- Bryan CJ, Tipton E, & Yeager DS (2020). To change the world, behavioral intervention research will need to get serious about heterogeneity. Unpublished Manuscript, University of Texas at Austin.
- Burgoyne AP, Hambrick DZ, & Macnamara BN (2020). How firm are the foundations of mind-set theory? The claims appear stronger than the evidence. *Psychological Science*, 31(3), 258–267. [PubMed: 32011215]
- Burnette JL, Knouse LE, Vavra DT, O’Boyle E, & Brooks MA (2020). Growth mindsets and psychological distress: A meta-analysis. *Clinical Psychology Review*, 101816. 10.1016/j.cpr.2020.101816
- Burnette JL, O’Boyle EH, VanEpps EM, Pollack JM, & Finkel EJ (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655–701. 10.1037/a0029531 [PubMed: 22866678]
- Canning EA, Muenks K, Green DJ, & Murphy MC (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. 10.1126/sciadv.aau4734 [PubMed: 30793027]
- Chetty R, Friedman JN, & Rockoff JE (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), 2593–2632. 10.1257/aer.104.9.2593
- Cheung ACK, & Slavin RE (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283–292. 10.3102/0013189X16656615
- Claro S, & Loeb S (2019). Students with growth mindset learn more in school: Evidence from California’s CORE school districts. EdWorkingPaper: 19–155. Retrieved from Annenberg Institute at Brown University: <http://www.edworkingpapers.com/Ai19-155>.
- Claro S, Paunesku D, & Dweck CS (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences*, 113(31), 8664–8668. 10.1073/pnas.1608207113
- Cury F, Elliot AJ, Da Fonseca D, & Moller AC (2006). The social-cognitive model of achievement motivation and the 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, 90(4), 666–679. 10.1037/0022-3514.90.4.666 [PubMed: 16649862]
- Diener CI, & Dweck CS (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, 36(5), 451. 10.1037/0022-3514.36.5.451

- Dweck CS (1999). *Self-theories: Their role in motivation, personality, and development*. Taylor and Francis/Psychology Press.
- Dweck CS, & Leggett EL (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256–273. 10.1037/0033-295X.95.2.256
- Dweck CS, & Reppucci ND (1973). Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology*, 25(1), 109–116. 10.1037/h0034248
- Dweck CS, & Yeager DS (2019). *Mindsets: A view from two eras*. Perspectives on Psychological Science. 10.1177/1745691618804166
- Elliott ES, & Dweck CS (1988a). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5–12. 10.1037//0022-3514.54.1.5 [PubMed: 3346808]
- Elliott ES, & Dweck CS (1988b). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54(1), 5–12. 10.1037//0022-3514.54.1.5 [PubMed: 3346808]
- Foliano F, Rolfe H, Buzzeo J, Runge J, & Wilkinson D (2019). *Changing mindsets: Effectiveness trial*. National Institute of Economic and Social Research.
- Ganimian AJ (2020). Growth-mindset interventions at scale: Experimental evidence from Argentina. *Educational Evaluation and Policy Analysis*, 42(3), 417–438. 10.3102/0162373720938041
- Gelman A (2018, 9 13). Discussion of effects of growth mindset: Let’s not demand unrealistic effect sizes. <https://statmodeling.stat.columbia.edu/2018/09/13/discussion-effects-growth-mindset-lets-not-demand-unrealistic-effect-sizes/>
- Gopalan M, & Yeager DS (2020). How does adopting a Growth Mindset Improve Academic Performance? Probing the Underlying Mechanisms in a Nationally- Representative Sample. OSF Preprints. 10.31219/osf.io/h2bzs
- Grant H, & Dweck CS (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, 85(3), 541–553. 10.1037/0022-3514.85.3.541 [PubMed: 14498789]
- Hattie J, Biggs J, & Purdie N (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66(2), 99–136. 10.3102/00346543066002099
- Hong Y, Chiu C, Dweck CS, Lin DM-S, & Wan W (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, 77(3), 588–599. 10.1037/0022-3514.77.3.588
- Hulleman CS, Schrager SM, Bodmann SM, & Harackiewicz JM (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, 136(3), 422–449. 10.1037/a0018947 [PubMed: 20438145]
- Kanopka K, Claro S, Loeb S, West M, & Fricke H (2020). Changes in social-emotional learning: Examining student development over time ([Working Paper] Policy Analysis for California Education). Stanford University. https://www.edpolicyinca.org/sites/default/files/2020-07/wp_kanopka_july2020.pdf
- Kenny DA, & Judd CM (2019). The unappreciated heterogeneity of effect sizes: Implications for power, precision, planning of research, and replication. *Psychological Methods*, 24(5), 578–589. 10.1037/met0000209 [PubMed: 30742474]
- Kraft MA (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, 49(4), 241–253. 10.3102/0013189X20912798
- Krueger AB (2003). Economic Considerations and Class Size. *The Economic Journal*, 113(485), F34–F63. JSTOR.
- Leslie S-J, Cimpian A, Meyer M, & Freeland E (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, 347(6219), 262–265. 10.1126/science.1261375 [PubMed: 25593183]
- Li Y, & Bates TC (2019). You can’t change your basic ability, but you work at things, and that’s how we get hard things done: Testing the role of growth mindset on response to setbacks, educational attainment, and cognitive ability. *Journal of Experimental Psychology: General*, 148(9), 1640–1655. 10.1037/xge0000669 [PubMed: 31464486]
- Li Y, & Bates TC (2020). Testing the association of growth mindset and grades across a challenging transition: Is growth mindset associated with grades? *Intelligence*, 81, 101471. 10.1016/j.intell.2020.101471

- Lipsey MW, Puzio K, Yun C, Hebert MA, Steinka-Fry K, Cole MW, Roberts M, Anthony KS, & Busick MD (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms (NCSEER 2013–3000). National Center for Special Education Research. <https://ies.ed.gov/ncser/pubs/20133000/>
- Macnamara B (2018). Schools are buying “growth mindset” interventions despite scant evidence that they work well. *The Conversation*. <http://theconversation.com/schools-are-buying-growth-mindset-interventions-despite-scant-evidence-that-they-work-well-96001>
- McShane BB, Tackett JL, Böckenholt U, & Gelman A (2019). Large-scale replication projects in contemporary psychological research. *The American Statistician*, 73(sup1), 99–105. 10.1080/00031305.2018.1505655
- Miele DB, & Molden DC (2010). Naive theories of intelligence and the role of processing fluency in perceived comprehension. *Journal of Experimental Psychology: General*, 139(3), 535–557. 10.1037/a0019745 [PubMed: 20677898]
- Molden DC, & Dweck CS (2006). Finding “meaning” in psychology: A lay theories approach to self-regulation, social perception, and social development. *American Psychologist*, 61(3), 192–203. 10.1037/0003-066X.61.3.192
- Mueller CM, & Dweck CS (1998). Praise for intelligence can undermine children’s motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33–52. 10.1037/0022-3514.75.1.33 [PubMed: 9686450]
- Muenks K, Canning EA, LaCrosse J, Green DJ, Zirkel S, Garcia JA, & Murphy MC (2020). Does my professor think my ability can change? Students’ perceptions of their STEM professors’ mindset beliefs predict their psychological vulnerability, engagement, and performance in class. *Journal of Experimental Psychology: General*. 10.1037/xge0000763
- Nelson LD, Simmons J, & Simonsohn U (2018). Psychology’s renaissance. *Annual Review of Psychology*, 69(1), null. 10.1146/annurev-psych-122216-011836
- Nussbaum AD, & Dweck CS (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality and Social Psychology Bulletin*, 34(5), 599–612. 10.1177/0146167207312960 [PubMed: 18276895]
- Nye B, Hedges LV, & Konstantopoulos S (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. *American Educational Research Journal*, 37(1), 123–151. 10.3102/00028312037001123
- OECD. (2019). PISA 2018 results (Volume III): What school life means for students’ lives. PISA, OECD Publishing. 10.1787/acd78851-en.
- Pashler -Hal, & de Ruiter, J. P.. (2017). Taking Responsibility for Our Field’s Reputation. *APS Observer*, 30(7). <https://www.psychologicalscience.org/observer/taking-responsibility-for-our-fields-reputation>
- Paunesku D, Walton GM, Romero C, Smith EN, Yeager DS, & Dweck CS (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, 26(6), 784–793. 10.1177/0956797615571017 [PubMed: 25862544]
- Qin X, Wormington S, Guzman-Alvarez A, & Wang M-T (2020). Why does a growth mindset intervention impact achievement differently across secondary schools? Unpacking the causal mediation mechanism from a national multisite randomized experiment. *OSF Preprints*. 10.31219/osf.io/7c94h
- Rege M, Hanselman P, Ingeborg FS, Dweck CS, Ludvigsen S, Bettinger E, Muller C, Walton GM, Duckworth AL, & Yeager DS (in press). How can we inspire nations of learners? Investigating growth mindset and challenge-seeking in two countries. *American Psychologist*.
- Rienzo C, Rolfe H, & Wilkinson D (2015). Changing mindsets: Evaluation report and executive summary. Education Endowment Foundation.
- Robins RW, & Pals JL (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self & Identity*, 1(4), 313–336. 10.1080/15298860290106805
- Schleider JL, Abel MR, & Weisz JR (2015). Implicit theories and youth mental health problems: A random-effects meta-analysis. *Clinical Psychology Review*, 35, 1–9. 10.1016/j.cpr.2014.11.001 [PubMed: 25462109]

- Schroder HS, Dawood S, Yalch MM, Donnellan MB, & Moser JS (2016). Evaluating the domain specificity of mental health–related mind-sets. *Social Psychological and Personality Science*, 7(6), 508–520. 10.1177/1948550616644657
- Seo E, Lee HY, Jamieson JP, Reis HT, Beevers CG, & Yeager DS (in press). Trait attributions and threat appraisals explain the relation between implicit theories of personality and internalizing symptoms during adolescence. *Development and Psychopathology*.
- Shpancer N (2020, 7 29). Life cannot be hacked: The secret of human development is that there is no one secret. *Psychology Today*. <https://www.psychologytoday.com/blog/insight-therapy/202007/life-cannot-be-hacked>
- Simonsohn U (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26(5), 559–569. 10.1177/0956797614567341 [PubMed: 25800521]
- Sisk VF, Burgoyne AP, Sun J, Butler JL, & Macnamara BN (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. 10.1177/0956797617739704 [PubMed: 29505339]
- Slavin RE (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, 15(9), 5–11. 10.3102/0013189X015009005
- The World Bank. (2017, 4 25). If you think you can get smarter, you will. <http://www.worldbank.org/en/results/2017/04/25/peru-if-you-think-you-can-get-smarter-you-will>
- Tipton E, Yeager DS, Iachan R, & Schneider B (2019). Designing probability samples to study treatment effect heterogeneity. In Lavrakas PJ (Ed.), *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment*. (pp. 435–456). Wiley. 10.1002/9781119083771.ch22
- TNTP. (2015). The mirage: Confronting the hard truth about our quest for teacher development. http://tntp.org/assets/documents/TNTP-Mirage_2015.pdf
- Vivalt E (2020). How much can we generalize from impact evaluations? *Journal of the European Economic Association*, 47. 10.1093/jeea/jvaa019
- Walton GM, & Yeager DS (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*. 10.1177/0963721420904453
- Warne R (2020, 1 3). The one variable that makes growth mindset interventions work. <https://russellwarne.com/2020/01/03/the-one-variable-that-makes-growth-mindset-interventions-work/>
- Weiner B (1985). An attributional theory of emotion and motivation. *Psychological Review*, 92(4), 548–573. 10.1037/0033-295X.92.4.548 [PubMed: 3903815]
- Weiner B, & Kukla A (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, 15(1), 1–20. 10.1037/h0029211
- West MR, Buckley K, Krachman SB, & Bookman N (2017). Development and implementation of student social-emotional surveys in the CORE Districts. *Journal of Applied Developmental Psychology*. 10.1016/j.appdev.2017.06.001
- Yeager DS (2019). *The National Study of Learning Mindsets*, [United States], 2015–2016. Inter-university Consortium for Political and Social Research [distributor]. 10.3886/ICPSR37353.v1
- Yeager DS, Carroll JM, Buontempo J, Cimpian A, Woody S, Crosnoe R, Muller C, Murray J, Mhatre P, Kersting N, Hulleman C, Kudym M, Murphy M, Duckworth AL, Walton G, & Dweck CS (2020). Teacher mindsets help explain where a growth mindset intervention does and doesn't work [Manuscript submitted for publication.].
- Yeager DS, Dahl RE, & Dweck CS (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, 13(1), 101–122. 10.1177/1745691617722620 [PubMed: 29232535]
- Yeager DS, & Dweck CS (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, 47(4), 302–314. 10.1080/00461520.2012.722805
- Yeager DS, Fong CJ, Lee HY, & Espelage DL (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology*, 37, 36–51. 10.1016/j.appdev.2014.11.005

- Yeager DS, Hanselman P, Muller CL, & Crosnoe R (2019). Mindset x Context Theory: How agency and structure interact to shape human development and social inequality. Working Paper, University of Texas at Austin.
- Yeager DS, Hanselman P, Walton GM, Murray JS, Crosnoe R, Muller C, Tipton E, Schneider B, Hulleman CS, Hinojosa CP, Paunesku D, Romero C, Flint K, Roberts A, Trott J, Iachan R, Buontempo J, Yang SM, Carvalho CM, ... Dweck CS (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364–369. 10.1038/s41586-019-1466-y [PubMed: 31391586]
- Yeager DS, Romero C, Paunesku D, Hulleman CS, Schneider B, Hinojosa C, Lee HY, O'Brien J, Flint K, Roberts A, Trott J, Greene D, Walton GM, & Dweck CS (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. 10.1037/edu0000098 [PubMed: 27524832]
- Yeager DS, Trzesniewski KH, & Dweck CS (2013). An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion. *Child Development*, 84(3), 970–988. 10.1111/cdev.12003 [PubMed: 23106262]
- Yeager DS, Walton GM, Brady ST, Akcinar EN, Paunesku D, Keane L, Kamentz D, Ritter G, Duckworth AL, Urstein R, Gomez EM, Markus HR, Cohen GL, & Dweck CS (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences*, 113(24), E3341–E3348. 10.1073/pnas.1524360113
- Zhu P, Garcia I, & Alonzo E (2019). An independent evaluation of growth mindset intervention. MDRC. <https://files.eric.ed.gov/fulltext/ED594493.pdf>

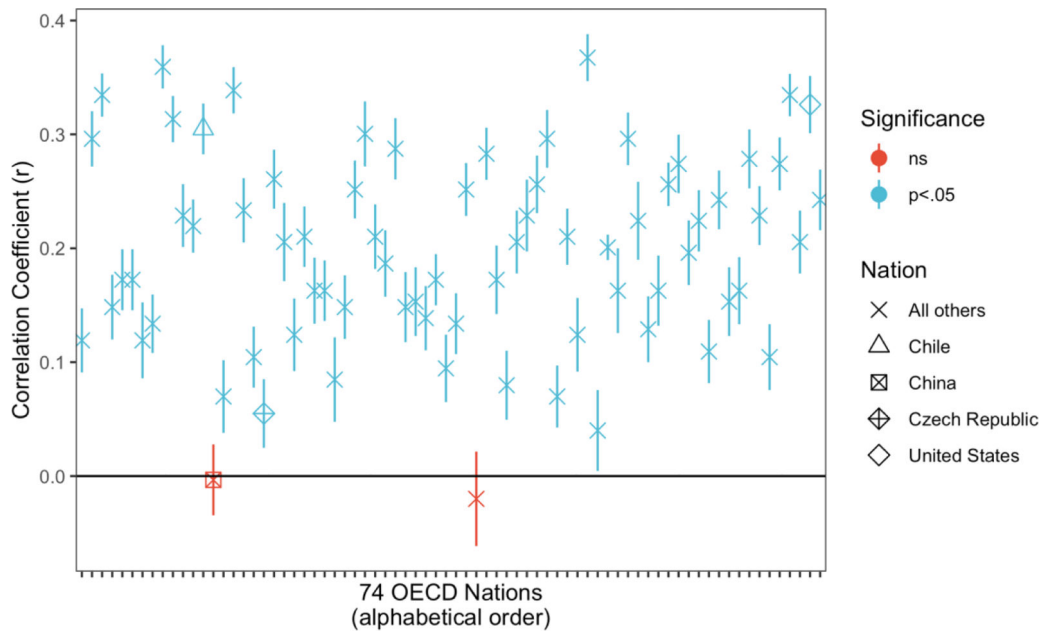


Figure 1. Correlations between mindset and test scores in 74 nations administering the mindset survey in the 2018 PISA (N = 555,458).

Note: Each dot is a raw correlation between the single-item mindset measure and PISA reading scores for each nation’s 15 year-olds. Bars: 95% CIs. Data Source: OECD, PISA 2018 Database, Table III.B1.14.5. Data were reported by the OECD in SD units; these were transformed to r using the standard “r from d” formulas. One nation contributed three different effect sizes according to the OECD (China: for Macao, for Hong Kong, and for four areas of the mainland); these three effect sizes were averaged here. Looking at the three China effects separately, each was near zero, and mainland China was negative and significant.

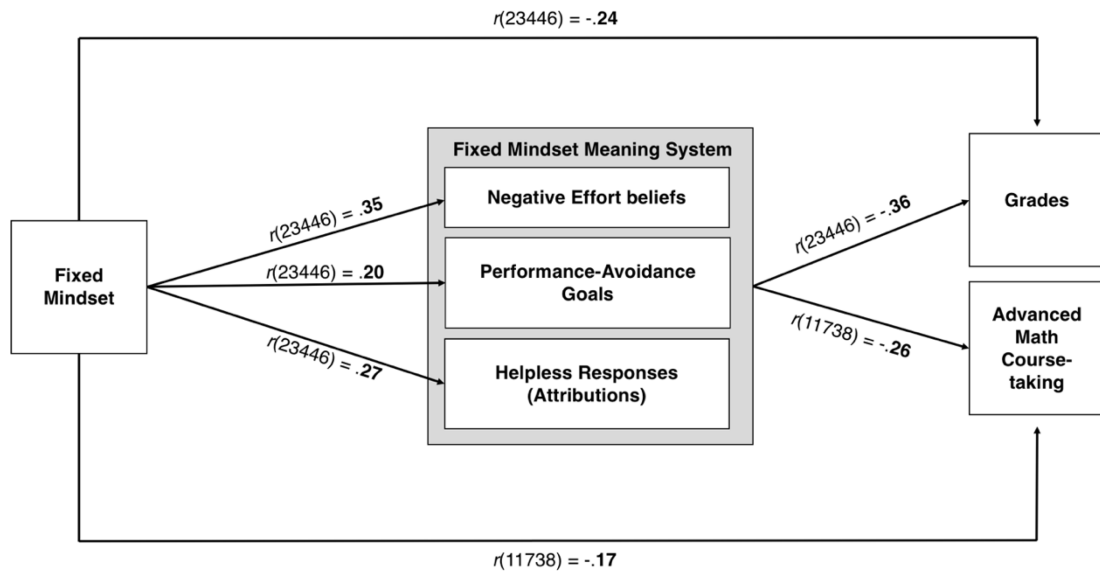


Figure 2. The links between fixed mindsets, the meaning system, and academic outcomes were replicated in three large studies of first-year high school students in two nations.

Note: Statistics are zero-order correlations estimated by meta-analytically aggregating three datasets: The U.S. National Study of Learning Mindsets (NSLM) pilot (Yeager, Romero, et al., 2016); The U.S. NSLM (Yeager et al, 2019); The Norway U-say experiment (Rege et al., in press). All p s < .001. Data from each of the items for each of the three studies are presented in Table 1 later in the paper. The meaning system paths to grades and course-taking were measured with the aggregate of the Mindset Meaning System items (see Table 1 below and the online appendix for the questionnaire). The MMI measure is brief (5 items) so that it could be incorporated into large-scale studies. Because these are single-item measures, magnitudes of correlations are likely to be under-estimates of true associations.

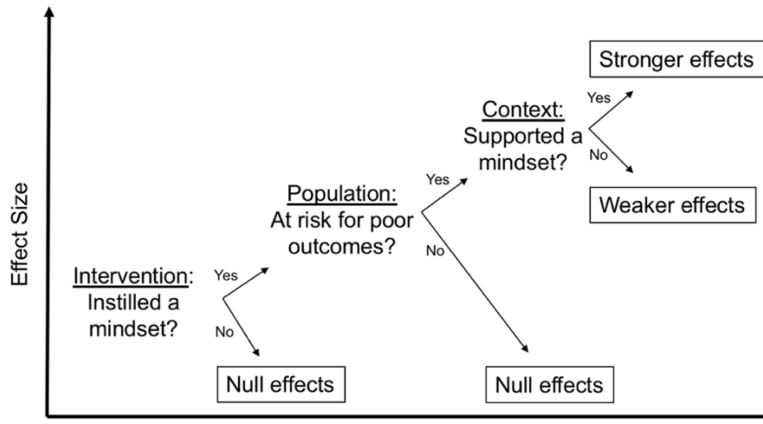


Figure 3. The mindset × context perspective: A decision tree depicts key questions to ask about a mindset intervention, and what kinds of effect to expect depending on the answer.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Correlations of Fixed Mindsets with the “Meaning System” in Replication Studies

Correlation with Fixed Mindset	Sample		
	NSLM Pilot, <i>N</i> = 3,306	NSLM, <i>N</i> = 14,894	<i>U-say</i> , <i>N</i> = 5,247
Meaning system aggregate index, <i>r</i> =	.43	.40	.41
Individual meaning system items:			
Effort Beliefs, <i>r</i> =	.47	.32	.36
Goals			
Performance-avoidance, <i>r</i> =	.21	.21	.17
Learning, <i>r</i> =	-.16	-.14	-.21
Response (attributions)			
Helplessness, <i>r</i> =	.29	.28	.23
Resilience, <i>r</i> =	-.14	-.15	-.25

Note: Correlations do not adjust for unreliability in the single items, so effect sizes are conservative relative to measures with less measurement error. NSLM = National Study of Learning Mindsets. Data sources: the NSLM (Yeager et al., 2019); the NSLM pilot study (Yeager et al., 2016), and a replication of the NSLM in Norway, the *U-say* study (Rege et al., in press). *Effort beliefs*: “When you have to try really hard in a subject in school, it means you can’t be good at that subject;” *Performance-avoidance goals*: “One of my main goals for the rest of the school year is to avoid looking dumb in my classes;” *Learning goals*: Student chose between “easy math problems that will not teach you anything new but will give you a high score vs. harder math problems that might give you a lower score but give you more knowledge;” *Helpless responses to challenge*: Getting a bad grade “means I’m probably not very smart at math;” *Resilient responses to challenge*: After a bad grade, saying “I can get a higher score next time if I find a better way to study.” All *p*s < .05.