Title

Whether Students Benefit from a Social and Emotional Learning Intervention Depends on their Motivation Profile

Author

Catalan Molina, Diego

Publication Date

2021

Peer reviewed|Thesis/dissertation

Whether Students Benefit from a Social and Emotional Learning Intervention Depends on their Motivation Profile

By

DIEGO CATALAN MOLINA

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Human Development

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____
Dr. Kali H. Trzesniewski, Chair

____
Dr. Jennifer Falbe

____
Dr. Adrienne Nishina

____
Dr. Michal Kurlaender

____
Dr. Kevin Gee

Committee in Charge

2021

Table of Contents

Acknowledgements

This work was partially supported by an award from the Agency for Research and Development (ANID) / Programa Formacion de Capital Humano Avanzado / DOCTORADO BECAS CHILE / 2015 – 72160147. In addition, the data presented here derive from a World Bank-initiated activity that is financed by the Australian Department of Foreign Affairs and Trade (DFAT) through the Partnership for Knowledge-Based Poverty Reduction Trust Fund. Indonesian data collection was supported by the World Bank Group's Umbrella Facility for Gender Equality in partnership with DFAT. The Umbrella Facility for Gender Equality (UFGE) is a multi-donor trust fund administered by the World Bank to advance gender equality and women's empowerment through experimentation and knowledge creation to help governments and the private sector focus policy and programs on scalable solutions with sustainable outcomes. The UFGE is supported with generous contributions from Australia, Canada, Denmark, Finland, Germany, Iceland, Latvia, the Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom, United States, and the Bill and Melinda Gates Foundation.

I thank Akbar Nikmatullah Dachlan for field coordination and supervision of data collection in Indonesia. In addition, I would like to acknowledge the valuable feedback and contributions to this manuscript from Tenelle Porter, Hillary C. Johnson, Noel Mueller, and Samantha De Martino. Last, I would like to acknowledge the contribution of many more research team members from the East Asia and Pacific Innovation Lab and the Mind, Behavior, and Development Unit at the World Bank, who helped to design and manage the larger efficacy study of All Can Succeed.

Abstract

One of the main challenges in intervention research today is understanding who benefits from universal school programs. Partly, this challenge involves capturing the diversity of experiences, needs, and traits in across students that may explain "who benefits" from interventions. Here, we studied *motivation profiles* (i.e., systems of beliefs, goals, and behaviors) to understand variation in treatment effects of social and emotional school interventions. Using data from a large intervention study (2,097 schools), we found four motivation profiles (*growth*, *multiple goals, disconnected*, and *severely disconnected*). Moreover, we found differential effects of a short social and emotional intervention, such that growth and severely disconnected, but not multiple goals or disconnected, students increased their test scores. In addition, each profile increased their use of a unique set of learning strategies. These findings suggest that motivation profiles can expand what we know about "who benefits" from social and emotional interventions and generate insights on how to improve interventions to reach a wider range of students.

Keywords: motivation profiles, person-centered approach, heterogeneity, school interventions

Whether Students Benefit from a Social and Emotional Learning Intervention Depends on their Motivation Profile

One of the main challenges in intervention research today is understanding who benefits from school programs (Bloom & Michalopoulos, 2013; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). This challenge is especially relevant to *universal* school interventions, which expose all students in a school or a classroom to an intervention, while only benefiting a subset of students (Greenberg & Abenavoli, 2017). To address this challenge, the field has called for researchers to study variation in treatment effects across students, teachers, and schools (Tipton, 2020; Yeager et al., 2019). This call is motivated by the fact that average treatment effects (i.e., effects of an intervention for the average student) conceal the natural variation in benefits across students, classrooms, and schools (Bloom, Raudenbush, Weiss, & Porter, 2017; Greenberg & Abenavoli, 2017; Raudenbush & Bloom, 2015). We need to learn more about who benefits from universal interventions and who does not, so that we can inform intervention development to better reach a wider range of students (Jiang, Santos, Josephson, Mayer, & Boyd, 2018; Lerner, 2015; Solmeyer & Constance, 2015).

To accomplish the goal of understanding who benefits from universal school interventions, we need to capture heterogeneity in how students may engage with and learn from such interventions. Often researchers use students' background characteristics (e.g., gender, baseline achievement, and socioeconomic status) as proxies of the type of student who may benefit from school interventions (Bloom & Michalopoulos, 2013). However, these characteristics are rough indicators for whether students will engage with and learn from interventions. A more proximal indicator of students' academic engagement and approach to learning is their achievement motivation. In particular, students' *motivation profiles* (i.e., systems of beliefs, goals, and behaviors) reflect how students make sense of their learning context, which in turn, may reflect how students engage with and learn from school interventions.

The current study had two main goals: (1) identify distinct motivation profiles among Indonesian adolescents and (2) test whether these profiles moderated the effects of a short social and emotional school intervention on student outcomes. To do so, we used data from a large randomized controlled trial of All Can Succeed (2,097 schools) conducted in Indonesia. Results from the study will help expand the evidence base on the importance of conceiving adolescents' motivation as a system of interrelated beliefs, goals, and behavior and contribute to the study of heterogeneity in intervention effects by highlighting the potential role that profiles can play in the evaluation of universal school interventions. In what follows, we review what motivation profiles are and why they are important for understanding how adolescents may engage with and learn from school interventions.

What Are Motivation Profiles and Why Do They Matter?

Traditionally, the study of adolescents' motivation has focused on understanding beliefs, goals, and behaviors as interrelated, but separate, variables (i.e., a variable-centered approach; Wormington, Corpus, & Anderson, 2012). Variable-centered approaches often focus on the unique influence of each variable over and above the others and ignore any joint influence across two or more variables. For example, in a model where beliefs, goals, and behaviors are unique predictors of academic achievement. The result of variable-centered approaches is evidence that reflects motivation as "pieces" of a person. In contrast, *person-centered* approaches cluster people to find groups of individuals with similar patterns of beliefs, goals, and behaviors (Marsh, Lüdtke, Trautwein, & Morin, 2009). For example, a study that finds three clusters of students, each one showing a unique pattern of beliefs, goals, and behaviors, which predict academic achievement together as a system.

In the study of achievement motivation, *motivation profiles* are a person-centered approach that reflects how beliefs, goals, and behaviors are organized as a system that helps students make sense of their learning context and pursue their goals (Hong et al., 1999; Linnenbrink-Garcia et al., 2018). Across multiple approaches to identifying motivation profiles, studies typically find more than one profile, suggesting that adolescents do vary in their configurations of beliefs, goals, and behaviors (e.g., Pastor, Barron, Miller, & Davis, 2007; Lau & Roeser, 2008). This finding has emerged in studies that build on self-determination theory (e.g., Gillet, Morin, & Reeve, 2017; Wormington et al., 2012), achievement goal theory (e.g., Schwinger, Steinmayr, & Spinath, 2016), mindset theory (e.g., Yu & McLellan, 2020), and studies that include constructs from more than one theory (e.g., Bae & DeBusk-Lane, 2018; Chittum & Jones, 2017; Collie, Martin, Nassar, & Roberts, 2019; Hodis, Hattie, & Hodis, 2017; Korpershoek, Kuyper, & van der Werf, 2015; Parhiala et al., 2018).

As an illustration, Lazarides and colleagues (2020) found four profiles (*high motivation beliefs, medium motivation beliefs, low motivation beliefs,* and *low intrinsic value*) by clustering three measures of math self-concept and task value. Among them, “low intrinsic value” and “medium motivation beliefs” students showed similar levels of math self-concept (i.e., whether a student sees themselves as having higher or lower math skills than their classmates). Had the authors only looked at individual differences in math self-concept, they would have pooled “low intrinsic value” and “medium motivation beliefs” students together. Pooling these profiles together would have meant assuming that “low intrinsic value” and “medium motivation beliefs” students this shared the same self-views in the school context. Yet, “low intrinsic value” students expressed much lower interest in and enjoyment during math class than "medium motivation beliefs" students. Moreover, "low intrinsic value" students were less likely to pursue math-intensive coursework than "medium motivation beliefs" students. This implies that comparing students along only one aspect of their motivation may lead researchers to wrongly pooling students with dissimilar experiences and outcomes together.

Another consistent finding in the literature is that motivation profiles emerge due to quantitative and qualitative differences (Marsh et al., 2009). Two profiles who are quantitatively different show a similar "shape" (i.e., configuration of beliefs, goals, and behaviors), but one profile shows consistently lower scores on indicators of motivation than the other does. In contrast, two profiles who are qualitatively different show distinct configurations of beliefs, goals, and behaviors. For example, Roeser and colleagues (2002) found a *multiple strengths* and a *poor mental health* profiles in their study. Both groups of students perceived their class as important and useful and that they were capable of mastering material covered in class. Yet, while multiple strengths students showed high levels of mental health, poor mental health students showed the opposite pattern (hence, their name). Therefore, the shapes of these two profiles reflect qualitative differences in how they experience their learning context.

A third consistent finding in the literature is that students need to show an adaptive configuration of multiple aspects of motivation (as opposed to isolated strengths) to thrive at school (Gillet et al., 2017; Oga-Baldwin & Fryer, 2018; Wormington et al., 2012). That is, a student who shows more adaptive levels on one important belief, goal, or behavior may not have the same chances to succeed in learning contexts as a student who shows an adaptive configuration of beliefs, goals, and behaviors. This finding implies that variable-centered approaches may miss important associations between students' experiences and outcomes because of these approaches' focus on beliefs, goals, and behaviors as unique predictors of outcomes.

Though the field is rapidly accumulating knowledge on motivation profiles, two gaps need to be addressed to improve the evidence base. First, the field needs to build bridges between motivation theories to focus on profiles that emerge when a wide range of beliefs, goals, and behaviors are considered. That is, too often the study of motivation profiles is concentrated on the clusters that may emerge in the context of a single theory and a few measures that reflect such theory (e.g., profiles based on self-determination theory, Gillet et al., 2017, or achievement goals theory, Pastor et al., 2007). Building bridges across theories would help to learn about profiles that are less dependent on specific measures and more reflective of students' motivation as a broad construct.

Second, the field needs to rely more on sampling strategies that capture a wide range of students. Building the field's evidence base on small, homogeneous, and convenient samples of students may result in profiles that do not replicate well in other contexts. This is particularly important for the study of motivation profiles because the ultimate reward for the field would be to have strong predictions about which profiles may be found across contexts, what explains that students develop a specific profile, and what are the consequences of holding such profiles. Yet, these predictions rely on findings from studies that recruit students that represent a wide range of the population, instead of studies that recruit a small, and likely homogeneous, group of students.

Goal 1: To Identify Distinct Achievement Motivation Profiles Among Indonesian Adolescents

To address the gaps in the literature, we identified motivation profiles using a wide range of beliefs (*growth mindset, effort, grit,* and *challenge-seeking*), goals (*learning* and *performance-*

avoidance goals), and behaviors (*mastery behavior*) across 50,280 Indonesian adolescents. The sample was representative of public middle schools in Java and Sumatra, the two most populated islands in the country. With a wide range of measures of motivation and a large representative sample, we were able to make predictions about the profiles that may replicate from study to study or across contexts.

In addition, we focused on Indonesian youth because the literature on motivation profiles is almost entirely based on samples from Western, educated, industrialized, rich, and democratic (WEIRD) societies. The lack of research in non-WEIRD countries is a weakness in the field considering that findings from WEIRD samples do not always generalize globally (Heinrich, Heine, & Norenzayan, 2010). Therefore, more research in non-WEIRD countries can expand our understanding of what motivation profiles look like and how much they matter globally. To put this in perspective, Indonesia is the fourth most populous country in the world and approximately 85% of the US population (World Bank, 2021). Consequently, the field could gain new insights by expanding the scope of "where" we study motivation profiles.

Given that past studies have mostly relied on samples from WEIRD countries, it is plausible that our hypotheses below do not describe well the motivational dynamics of Indonesian adolescents. Nonetheless, there are two studies that have found similarities in the motivational dynamics between US and Indonesian youth, which suggests that prior evidence may be reasonably good to forecast findings in Indonesia. For example, survey measures of growth mindset, grit, and learning goals appear to tap into similar constructs in Indonesia and in the US (Napolitano et al., 2021). Furthermore, Indonesian adolescents' avoidance goals (i.e., avoiding situations where peers might evaluate one's ability) are associated with lower academic achievement (Liem, Martin, Porter, & Colmar, 2012), a similar pattern found among US-based adolescents (Hulleman, Schrager, Bodmann, & Harackiewicz, 2010).

Registered hypotheses. First, we expected that more than one profile would better fit our data than a one-profile model (Hypothesis 1.1). A one-profile model is interpreted as a population of students with a homogeneous pattern of motivational beliefs, goals, and behaviors. This hypothesis was based on the accumulated evidence by motivation research, which suggests that distinct patterns of organization of beliefs, goals, and behaviors can be found between individuals (Marsh et al., 2009). Specifically, based on prior research on students' beliefs about intelligence (Chen & Tutweiler, 2017; Gunderson, Park, Maloney, Beilock, & Levine, 2018; Haimovitz & Dweck, 2017; Park, Gunderson, Tsukayama, Levine, & Beilock, 2016; Yeager & Dweck, 2012; Yu & McLellan, 2020) and motivation profiles (e.g., Gillet et al., 2017; Linnenbrink et al., 2018), we hypothesized that we would find four distinct motivation profiles: A growth profile (Hypothesis 1.2), a fixed profile (Hypothesis 1.3), a multiple goals profile (Hypothesis 1.4), and a disconnected profile (Hypothesis 1.5).

The shapes of these profiles were defined by seven beliefs, goals, and behaviors: Growth mindset (i.e., belief in the malleability of intelligence, Dweck, 1999), effort beliefs (i.e., belief in exerting effort to improve ability, Blackwell, Trzesniewski, & Dweck, 2007), grit (i.e., perseverance of effort for long-term goals, Duckworth, Peterson, Matthews, & Kelly, 2007), challenge-seeking preferences (i.e., engaging in challenging or easy school-related activities, Dweck, 1999), learning goals (i.e., seeking understanding and skill development, Elliott & Dweck, 1988), performance-avoidance goals (i.e., avoiding situations in which peers may identify a student's weaknesses or knowledge gaps, Elliott & Dweck, 1988), and mastery behavior (i.e., behavior that shows persistence, effort, resilience, and challenge-seeking, Porter et al., 2020).

Growth profile. Students who show a growth profile would have positive, approach-oriented beliefs and attitudes towards learning and their behavior would align with these beliefs. That is, they would have high scores on growth mindset, learning goals, and mastery behavior, as well as low scores on performance-avoidance goals relative to other profiles. The overall shape of the growth profile would suggest that students are not afraid of seeking challenges and perfecting their skills.

Previous studies have found a similar profile (Chen & Tutweiler, 2017; Haydel & Roeser, 2002; Lau & Roeser, 2008; Roeser et al., 2002; Yu & McLellan, 2020). For example, Yu and McLelland (2020) found a *growth-focused* profile whose shape matches our hypothesized growth profile. These students showed high levels of growth mindset, effort beliefs, learning goals, and perseverance, as well as low levels of performance-avoidance goals and self-handicapping behavior. Similarly, Roeser and colleagues (2002) found a *mastery-oriented* profile, representing students high on growth mindset, self-efficacy, and learning goals. Although different names have been used in the literature, these examples capture the defining characteristic of our hypothesized growth profile: A student who believes they can improve their skills, they seek to improve skills, and they behave like they want to master skills.

Fixed profile. Students who show a fixed profile would have negative, avoidance-oriented beliefs and attitudes about learning and their behavior would align with these beliefs. That is, they would have low scores on growth mindset, learning goals, and mastery behavior, as well as high scores on performance-avoidance goals relative to other profiles. The overall shape of the fixed profile would suggest that students avoid challenging situations because they do not see a connection between challenging themselves and improving their skills.

The same studies that have found a growth profile, have also found students that resemble our hypothesized fixed profile. Yu and McLelland (2020) found an *ability-focused* profile, characterized by low levels of growth mindset, effort beliefs, learning goals, and perseverance, as well as high levels of performance-avoidance goals and self-handicapping behavior. In addition, Roeser and colleagues (2002) found a *helpless* profile, defined by low levels of growth mindset, learning goals, and self-efficacy, as well as high on performance-avoidance goals.

Multiple goals profile. Students who show a multiple goals profile would actively pursue both learning and performance-avoidance goals. That is, they would have high scores on both goals. Regardless of how they view themselves or how they behave, the defining characteristic of students with a multiple goals profile is that, although they are driven to develop their skills, they also worry about looking "dumb" in front of others. This defining characteristic may hold them back from embracing some challenges and learning opportunities, whereas growth students would embrace struggle and mistakes as an essential part in their skill development trajectory (regardless of how others might view their mistakes). In the larger literature on motivation profiles, multiple goals students are a consistent find when learning and performance goals, as well as growth mindset, are among the profile indicators (e.g., Bae & DeBusk-Lane, 2018; Pastor et al., 2007; Schwinger et al., 2016; Yu and McLelland, 2020; Roeser et al., 2002).

Disconnected profile. Students who show a disconnected profile would have a misalignment between their beliefs, goals, and behaviors (i.e., a misalignment among indicators that are assumed to go in the same direction). For example, a student who believes they prefer challenging activities at school, but when presented with challenges, chooses to engage in easier tasks; or a student who believes putting in effort is the key to getting smarter, but gives up easily/quickly when challenged. Because of this misalignment, the overall shape of this profile would suggest that students' goal pursuit is disconnected from their beliefs.

Across studies on motivation profiles, another consistent finding is the type of student who shows a "low quality" profile. That is, a system of beliefs, goals, and behaviors that does not help students to adapt and thrive in their context (Gillet et al., 2017; Oga-Baldwin & Fryer, 2018). Here, disconnected students would represent a low-quality profile assuming that the misalignment of their beliefs, goals, and behaviors would reduce their chances to achieve their goals and meet their psychological needs (Dweck, 2017).

Summary. The literature on motivation profiles has shown that students thrive at school when their beliefs, goals, and behaviors are organized as adaptive systems. Yet, many studies have focused on identifying profiles using (1) measures that do not capture motivation as a broad construct and (2) samples that may not reveal the extent to which different profiles exist among students in the population. Here, we hypothesized that four profiles would emerge among a large sample of Indonesian adolescents. These profiles bridge findings from different theories aiming to capture profiles that may emerge across studies and contexts.

Can Motivation Profiles Explain Who Benefits from a Social and Emotional School Intervention?

Researchers often explain treatment effect variation based on students' background characteristics, such as gender and socioeconomic status (Bloom & Michalopoulos, 2013; Page, Feller, Grindal, Miratrix, & Somers, 2015; Raudenbush & Bloom, 2015). These characteristics are convenient tools to identify who benefits from an intervention, but they do not always help researchers and program developers to improve interventions. For example, learning that female students increase their school grades after participating in a mentoring program, while male students do not, does not directly translate into ideas to improve the intervention for future iterations. The issue is that background characteristics do not always explain who engaged with a curriculum, developed skills, and transferred their learning to new contexts.

Here, we propose that motivation profiles are a better proxy for who engaged with and learned from a universal school intervention. Therefore, the field could make greater strides in understanding the variation in school intervention effects, if such variation was explained by motivation profiles. Motivation profiles reflect how adolescents make sense of challenges in their learning context and the extent to which students are interested in and engaged at school (Grund, 2013; Haydel & Roeser, 2002; Roeser et al., 2002). In turn, we know that when students are not interested in or engaged with an intervention, they are less likely to benefit from it (Low, Smolkowski, & Cook, 2016; Zvoch, 2012). Therefore, identifying which profiles engage with and benefit from a curriculum can shed light on how a program helps (or hinders) adolescents' understanding of and approach to challenges. As a result, this information could be used to improve programs to reach a wider range of students.

For instance, learning about which motivation profiles are less likely to benefit from an intervention would lead to insights to tailor intervention content and activities to reach a wider range of participants in future iterations. Such a strategy would mirror the efforts in other fields, like medicine and public health, in which researchers seek to develop interventions that balance a universal delivery with personalized or targeted components to reach different types of people (August, Gewirtz, & August, 2018; Greenberg & Abenavoli, 2017; Lerner, Agans, DeSouza, & Hershberg, 2014).

Goal 2: To Test Whether Motivation Profiles Moderated the Effects of All Can Succeed on Students' Learning Strategies and Test Scores

All Can Succeed is a recently developed social and emotional school intervention focused on growth mindset and self-management. In its first randomized controlled trial, the program showed only a few small average effects among Indonesian adolescents (Johnson et al., 2020). These results could be taken as evidence that the program did not work, and thus, no more exploration of the program's potential impacts is needed. However, we know that average treatment effects can conceal potential positive impacts for some groups of students. All Can Succeed, for example, did show positive impacts for female students, lower-achieving students, and students in lower-performing schools (Johnson et al., 2020). Nevertheless, these findings are difficult to translate into ideas about how to improve the program to reach other students. In the present research, we explored whether the effects of All Can Succeed on test scores and learning strategies differed based on students' motivation profiles.

Hypothesis 2.1: Students with a fixed profile exposed to All Can Succeed will increase their national test scores. One of the main goals of All Can Succeed is to teach students how to reframe their experiences of struggle and challenge. Specifically, the program teaches students to see challenges as opportunities to learn new skills and that investing effort, persisting in the face of challenges, and using the right strategies are essential to successfully approach challenges. Given that students with a fixed profile are more likely to think they cannot improve their skills, they are less likely to think that investing effort and taking on challenges would be beneficial for their learning. As a result, fixed profile students are expected to show the most benefit from the program, because they are low on the skills being taught. This is consistent with previous research that shows that students who experience the most struggle in school benefit the most from growth mindset interventions (e.g., Paunesku et al., 2015; Yeager et al., 2019).

Hypothesis 2.2: Students with a fixed profile exposed to All Can Succeed will increase their use of learning strategies. As mentioned above, All Can Succeed teaches students to develop beliefs, goals, and behaviors to successfully face challenges at school. Considering that fixed students are less likely to seek challenges at school, it is plausible that they also use strategies to manage their learning less frequently than students who do seek challenges (e.g., growth students). As a result, fixed students are expected to increase their use of learning strategies after participating in the program. Perhaps, this behavioral change is the reason why students who experience the most struggle in school benefit the most from growth mindset interventions.

Current Study

This study had two main goals: (1) To identify distinct achievement motivation profiles among Indonesian adolescents and (2) to test whether these profiles moderated the effects of All Can Succeed on students' learning strategies and national test scores. To address the first goal, we used Latent Profile Analysis on seven measures of achievement motivation (Study 1). To address the second goal, we estimated whether students in each profile benefited from the intervention or not (Study 2). Throughout, we used data from a large randomized controlled trial of All Can Succeed, in which 2,097 schools were assigned to a business-as-usual control or one of two versions of the intervention. See more details about the registered hypotheses, changes to the registration, and methods in the supplementary materials.

Study 1

Method

Participants. Eight grade students (n = 50,280) participated in the efficacy study of All Can Succeed. This sample of students was representative of adolescents attending public middle schools in Java and Sumatra, the two most populated islands in the country. At baseline, 54% of students were female, 17% had experienced hunger during the last 30 days, and 40% had mothers whose highest attained educational level was lower than secondary school.

Procedure. As shown in Figure 1, a probabilistic sample of 2,097 schools in Java and Sumatra was recruited. In each school, one eight grade class was randomly selected to respond to the survey at baseline (end of eighth grade). In addition, two male and two female students in each school (n = 8,388) were chosen at random to complete the Persistence Effort Resilience Challenge-seeking (PERC) behavioral task (See description below).

At baseline and endline, students responded to a battery of survey and behavioral measures administered by a third-party research firm that was responsible for data collection. Survey administrators made every effort to ensure that students responded to the whole survey, which resulted in rates of missing data below 0.01% among survey measures. This firm was not involved in the design of the study.

Measures. To identify the motivation profiles, we used six survey measures and one behavioral task, all of which were adapted, translated, and back translated (the students took the survey in Indonesian). Below are brief description of the measures. See the online supplement for more detailed descriptions, as well as descriptive statistics (Table S1) and correlations (Table S2).

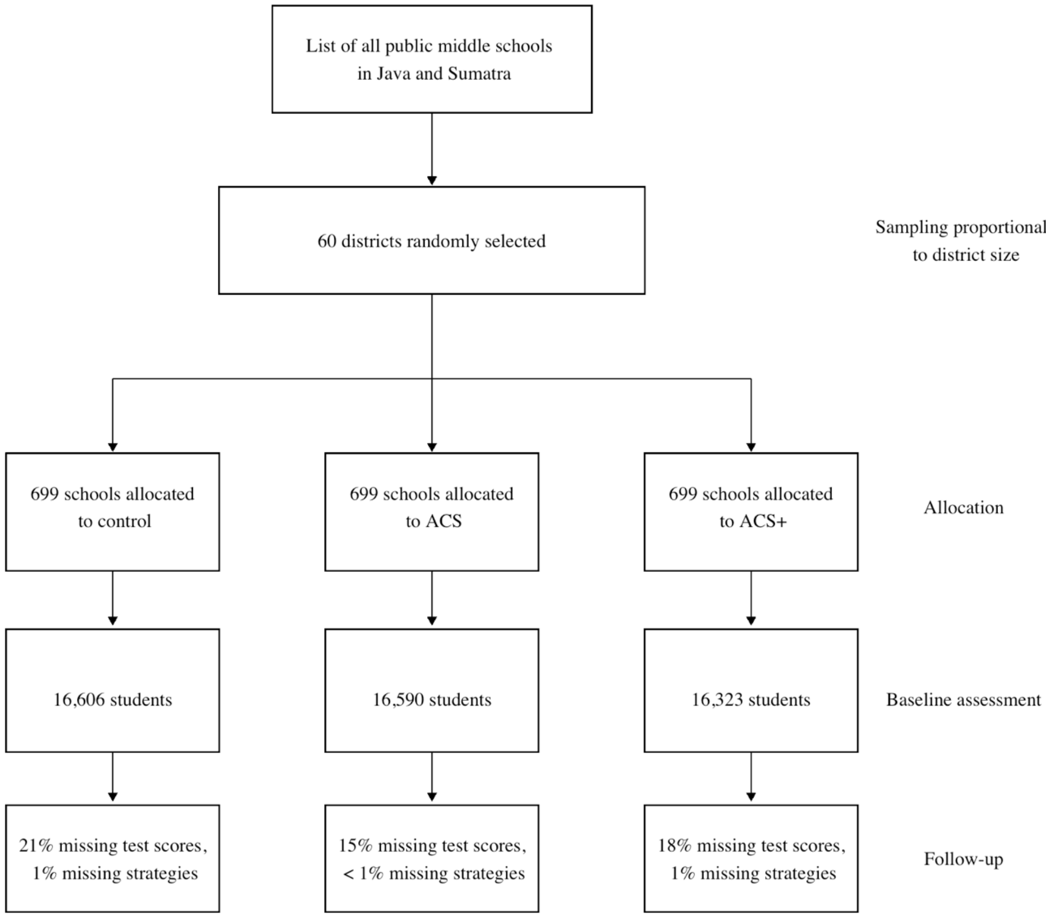

Figure 1. Recruitment and randomization chart. ACS = Six lessons; ACS+ = Six lessons plus supplemental activities. Larger districts had a higher probability of being selected from the list of all public middle schools in Java and Sumatra (i.e., sampling proportional to district size). Within each district, the same number of schools was randomized to each condition for a total of 2,097 schools. In each school, one eighth grade class was randomly selected to respond to the survey at baseline (end of 8th grade). The follow-up assessment occurred at the end of 9th grade.

Students' background: Gender, socioeconomic status, and school grades. Gender was self-reported by students. Socioeconomic status is an index that aggregated information on students' self-reported household assets (e.g., a car, a refrigerator, and cellphones with access to internet), recent experiences of hunger, and having to miss school because of work.

Baseline grades (i.e., school grades at the end of seventh grade) were obtained from schools during data collection. Grades range from 0 to 100 and the passing grade is 75 for most subjects. We averaged math, science, and English (a required second language class) grades. In addition, we used Indonesian grades as a proxy for reading comprehension.

Growth mindset. Growth mindset was measured by four items rated on a Likert-type scale, ranging from (1) Not true at all to (5) Completely true (α = .65). An example item is "You cannot change how smart you are." Items were reverse-coded so that higher scores represent the growth end of the mindset continuum. These items were adapted from the Theory of Intelligence survey (Dweck, 1999).

Effort beliefs. Effort beliefs were assessed by six items rated on a Likert-type scale, ranging from (1) Not true at all to (5) Completely true (α = .57). An example item is "Doing well in school requires hard work and effort." These items were adapted from Blackwell et al. (2007).

Grit. Grit was measured with four items rated on a Likert-type scale, ranging from (1) Almost never to (5) Almost always (α = .59). An example item is "I finish whatever I start." These items were adapted from the Short Grit Scale (Duckworth & Quinn, 2009). These items capture the *perseverance of effort* dimension of grit. However, for simplicity, we refer to the measure as *grit* throughout the manuscript.

Challenge-seeking preferences. Challenge-seeking was assessed with three dichotomous items for which students displayed a preference. An example item is "If you had to choose between having easy or difficult class work, which one would you choose? Easy class work or Difficult class work." This measure was adapted from items used in past growth mindset research (Dweck, 1999). We do not report alpha as a measure of internal consistency because these items do not represent a latent construct. Instead, these items summarize specific challenge-seeking preferences. In other words, this is a formative measure that does not rely on correlations among items (Coltman, Devinney, Midgley, & Venaik, 2008; Diamantopoulos & Winklhofer, 2001).

Achievement goals. Learning goals were assessed with three items rated on a Likert-type scale, ranging from (1) Not true at all (5) Completely true (α = .58). An example item is "It's important to me that I completely understand my class work." Performance-avoidance goals were assessed with four items rated on a Likert-type scale, ranging from (1) Not true at all (5) Completely true (α = .66). An example item is "I would only answer a question in class if I knew I was right." This measure was adapted from two subscales of the Patterns of Adaptive Learning Scale (Midgley et al., 2000).

Mastery behavior. As a behavioral measure of motivation, we used the Persistence Effort Resilience Challenge-seeking (PERC) Task (Porter et al., 2020). In this task, students complete a series of puzzles of varying levels of difficulty. The puzzles are divided into different sets and student responses to each set are used to assess their levels of persistence, effort, resilience, and challenge-seeking. We used a mastery behavior composite score that summarizes students' behavioral levels of persistence, effort, resilience, and challenge-seeking. This measure is also a formative scale, and thus, we do not report alpha as an index of the reliability of this measure.

Registered Analysis Plan

We used Latent Profile Analysis (LPA) to identify motivation profiles. Broadly, LPA extracts subgroups of individuals whose responses to each measure are similar within the subgroup and different from other subgroups (Lanza & Cooper, 2016). We used Mplus 8.3

(Muthén & Muthén, 1998-2017) to identify motivation profiles and the MplusAutomation package (Hallquist & Wiley, 2018) in R (R Core Team, 2020) to automate analyses. Mplus and R scripts can be found in https://osf.io/cg4ue/?view_only=6122ff7b0ad648e2b8401a7db2cde64f.

Before conducting analyses, we registered a set of decisions about the analyses and interpretation (registration can be found at https://osf.io/4pyzn/). These decisions were: (1) we used cluster-robust standard errors (i.e., option TYPE=COMPLEX in Mplus) to account for students' clustering in schools; (2) we used mean scores as profile indicators (i.e., averaging items within each construct); (3) we standardized the mean scores to ease the interpretation of the profile shapes; (4) we only let the indicator means vary across profiles, as opposed to letting means, variances, and covariances vary; and (5) we used 10-fold cross-validation to decide on the optimal number of motivation profiles and their replicability (see Grimm, Mazza, & Davoudzadeh, 2017, for a detailed explanation of the benefits of k-fold cross-validation in mixture modeling). In addition, missing data were handled using full information maximum likelihood estimation. See the online supplement for more details on the registered analysis decisions.

After identifying motivation profiles, we planned to test whether students' background would influence how they responded to each of the profile indicators. That is, the second aim in the registered analysis plan focused on testing *measurement invariance*. We followed the procedure outlined by Masyn (2017), in which Multiple Indicator Multiple Cause (MIMIC) models are used to evaluate whether covariates introduce bias in the identification of profiles. An example of such bias is when two students who share the same profile, but have different background characteristics, show different scores on one (or more) of the profile indicators.

Results

Number of profiles. The 10-fold cross-validation process suggested that four- and five-profile solutions fit the data well. We adopted the four-profile solution because the five-profile solution was no more informative and was less parsimonious (see the supplement for details). We labelled the four profiles *growth*, *multiple goals*, *disconnected*, and *severely disconnected*. Three of these matched our hypothesized profiles (i.e., growth, multiple goals, and disconnected). We did not find our hypothesized fixed profile. Instead, we found a more pronounced form of disconnected beliefs, goals, and behaviors, which we named severely disconnected.

Profile shapes. Figure 2 shows the means for each of the profiles across measures. Students with a growth profile (12.97% of the sample; represented in gold in the figure) showed high levels of positive beliefs and behaviors and stand out because of their commitment to learning goals while not being held back by performance avoidant goals. A majority of the students showed a multiple goals profile (55.41%; light yellow in the figure). The defining characteristic of multiple goals students was their equal endorsement of both learning and performance-avoidance goals. Disconnected (27.69%) and severely disconnected (3.93%) students showed the lowest levels of motivation on five of the seven measures. Although both profiles somewhat believed that their intelligence may change (i.e., average growth mindset scores), they showed low belief in effort investment as a necessary aspect of improving their

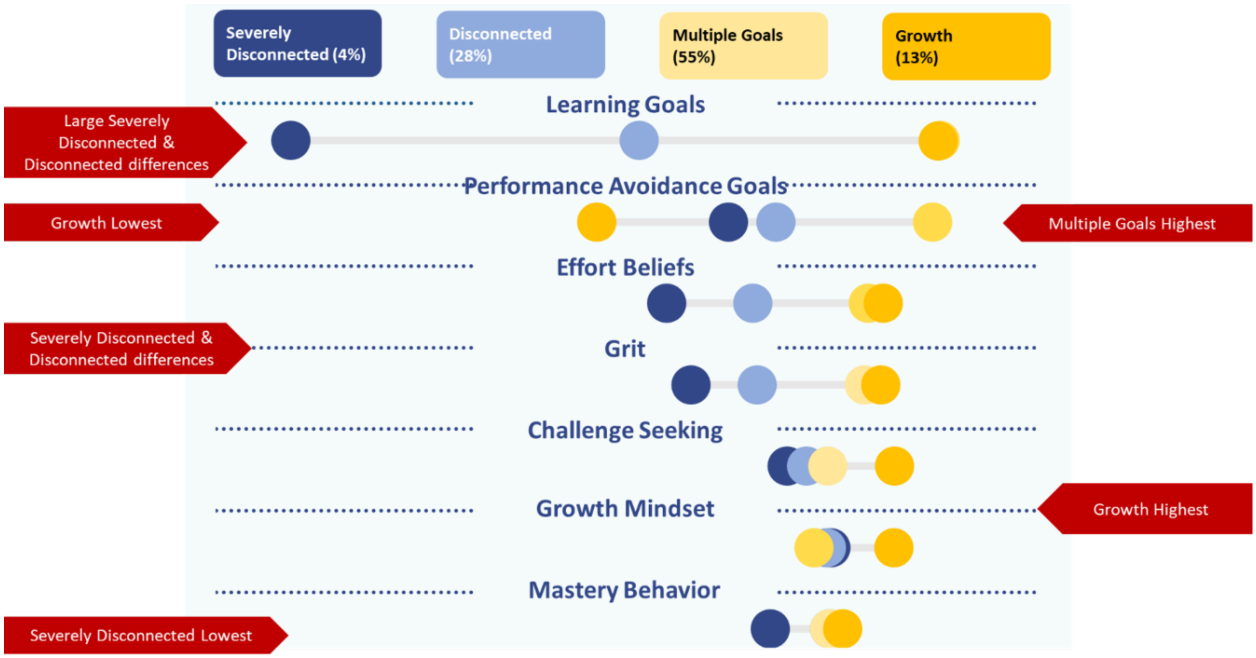

Figure 2. Motivation profiles across measures. The colored dots display the means for each measure by profile.

skills (i.e., low effort scores), in the degree to which they persevered in a task (i.e., low grit scores), and in their pursue of mastery (i.e., low learning goal scores).

Invariance of profiles by students' background. We found a small degree of profile non-invariance across students' gender, socioeconomic status, and academic achievement (see more details in the supplement). That is, student background characteristics did not substantively change the shapes in Figure 2, though the average scores on each measure did show minor differences depending on students' gender, socioeconomic status, and academic achievement (see Figure S3 in the online supplement). For example, female students consistently scored lower on mastery behavior compared to male students across all profiles. Yet, regardless of gender, growth students showed higher levels of mastery behavior than, for example, disconnected students.

Although the levels of non-invariance were small, including these covariates in the models is important to accurately predict students' most likely profile. Had we omitted these tests, we would have wrongly classified 17% of female students and 10% of male students in the growth profile, instead of the multiple goals profile. Therefore, we included gender, socioeconomic status, and academic achievement as covariates in our classification model (i.e., the model from which we classified students into a single, most likely profile as an observed categorical variable for Study 2).

Beyond issues with profile non-invariance, students' gender and socioeconomic status predicted students profile membership. Female students were more likely to show growth and multiple goals profiles than male students did (see Figure S4 in the online supplement). Additionally, low-achieving students at baseline (i.e., 1SD below the mean of baseline grades) were more likely to show either a multiple goals, disconnected, or severely disconnected profile (instead of a growth profile) compared to high-achieving students. Socioeconomic status did not predict profile membership after controlling for gender and academic achievement.

Discussion

We found four distinct motivation profiles. Each one reflects a unique approach to learning and could provide insights into why only some students engage with and benefit from universal school interventions. For example, students showing growth and multiple goals profiles displayed patterns of beliefs, goals, and behaviors that push them to develop skills, persist in the face of challenges, and believe they can improve. Previous studies have shown that students with similar profiles are better equipped to thrive at school (e.g., Roeser et al., 2002; Schwinger et al., 2016; Yu & McLelland, 2020), and thus, these students may have less of a need for a classroom environment that helps them align their beliefs, goals, and behaviors. In contrast, students showing disconnected and severely disconnected profiles showed motivational patterns that do not help them to navigate struggle, mistakes, and challenges at school. Given their misalignment of beliefs, goals, and behaviors, these two profiles may benefit the most from teaching and interventions that help them to successfully seek and persist through challenges that will lead to developing their skills.

Additionally, we found that the shapes of each profile were minimally affected by students' background. Yet had we not included background characteristics in Study 1, we would have wrongly classified multiple goals students into the growth profile. This result implies that background characteristics are connected to students' profiles, though a student's background and their profile are two distinct features of an individual.

Study 2

Given students vary in motivation profiles and these profiles reflect how students make sense of their school experiences and pursue their goals, then it is reasonable to assume that students will vary in how they engage with and learn from a universal school intervention. This is because a universal school intervention may not address the needs and struggles of every student in a classroom (i.e., variation in the person-environment fit across students, Eccles et al., 1993). For students whose needs are addressed by the intervention, we might expect them to engage with the content and activities, and thus, to learn important skills from the intervention. For students whose needs are not addressed by the intervention, we might expect them to disengage and not benefit from the intervention. In Study 2, we tested the hypothesis that the impact of All Can Succeed on students' national test scores and learning strategies depended on students' motivation profile assuming that the intervention would not address the needs of all students in each school.

Two versions of All Can Succeed were used to test this hypothesis. One version, ACS, consisted of 9th grade teachers or counselors teaching six lessons and the other version, ACS+, consisted of the same six lessons plus supplemental activities delivered by all of the 9th grade teachers in the school. We did not hypothesize whether ACS or ACS+ would be more beneficial across profiles, and thus, our hypotheses apply to both versions of the program.

One departure from the study registration is worth noting here. We originally hypothesized that fixed students would increase their use of learning strategies after participating in All Can Succeed because they would learn about how to improve their skills by seeking and persisting through challenges. Given that we did not find a fixed profile in Study 1, we focused on testing whether disconnected and severely disconnected students would increase their test scores and learning strategies after participating in All Can Succeed, as the same reasons why we expected fixed students to benefit applies to these two profiles.

Method

Participants. Study 2 participants were students who participated in Study 1 (see Figure 1). Overall, attrition was low and balanced in the intervention and control groups. All students with test score or learning strategy data (the outcomes) were included in the analysis, regardless of whether they had missing data on the predictors (see the supplement for more details about attrition and missing data estimation).

Procedure. Each school chose a guidance counselor or a teacher to deliver All Can Succeed (ACS) lessons to 9th grade students. In addition, schools allocated to All Can Succeed Plus (ACS+) invited other 9th grade teachers to deliver supplemental activities in their class. Schools received video tutorials explaining how to implement the program, which were shipped to schools with all the program's materials. The local research team was available (by phone or e-mail) to teachers and counselors who had questions about the implementation of the program. Students completed a post-intervention assessment at the end of 9^{th} grade, approximately year after the baseline assessment.

As often occurs in school intervention studies, not every school complied with their assignment. Among schools assigned to ACS, 63% delivered all six lessons, whereas 81% delivered at least one lesson. Among schools assigned to ACS+, only 37% delivered all six lessons and two supplemental activities, while 54% delivered at least one lesson and one supplemental activity. As described below, our models did not take into account compliance rates, and thus, our findings represent the effects of the intervention when their school is assigned to deliver All Can Succeed (instead of when students are actually exposed to it). The average lesson duration was 51 minutes, with 72% of schools reporting an average lesson time between 35 and 55 minutes.

All Can Succeed. All Can Succeed is a social and emotional learning classroom-based curriculum focusing on growth mindset and self-management. The goals of the program are to (1) reframe students' experiences of struggle at school, (2) promote the learning strategies students need to succeed in secondary school, and (3) raise students' educational and employment aspirations. These goals were motivated by the key transition Indonesian students experience at the end of 9th grade, when they have to decide to continue on a general education track or enroll in a vocational education track. Given the high dropout rates during this transition and the country's need for more youth to continue to higher education (Dilas, Mackie, Huang, & Trines, 2019), All Can Succeed was designed to promote the idea among students that they could aspire to a higher education degree and it offered tools for students who wanted to succeed academically.

The program builds on previous intervention studies in the US (e.g., Blackwell et al., 2007; Paunesku et al. 2015; Yeager et al. 2019), Peru (Outes-León, Sánchez, & Vakis 2020), and a prior pilot study in Indonesia (World Bank, 2019). The previous study in the country offered two 40-minute lessons focused on teaching students about the malleability of their intelligence and skills (i.e., a growth mindset). Building on these two lessons, the research team designed four more lessons that integrated teaching students about a growth mindset and self-management, the ability to regulate emotions and behaviors, motivate oneself, and work towards achieving personal and academic goals (Weissberg, Durlak, Domitrovich, & Gullotta. 2015). See supplement for more details about each lesson. The lessons took place during a regularly occurring weekly session in which school counselors discussed non-academic issues (e.g., occupational planning, social and emotional issues). Typically those sessions did not include a curriculum.

In the ASC+ condition, the six lessons were supplemented by activities and materials delivered by other teachers during their regular classes. These activities and materials were designed to create a learning environment in which students could transfer to and practice in other contexts the skills learned during the comic book-based lessons. For example, a “Sharing Board” activity aimed to normalize struggle among students (i.e., everyone struggles, dreams, learns, and can improve). At the beginning of the week, the homeroom teacher introduced the prompt of the week (i.e., a topic for which students could share experiences of struggle) and hung a poster noting the prompt on the wall. At any point throughout the week, the students could complete cards with the writing prompt and hang them on the poster. Then, at the end of the week, the homeroom teacher celebrated students who had participated in the activity. See the supplement for more details on the activities and materials included in ACS+.

Measures. Gender, socioeconomic status, and baseline academic achievement were used again as covariates at the student level.

Motivation profiles. Students' profile classifications from Study 1 (i.e., most-likely profile as an observed categorical variable) were used. At baseline, motivation profiles were balanced across the intervention and control groups (see Table S10 in the online supplement).

National test scores. Schools provided science, math, English, and Indonesian national test scores (at the end of 9th grade). See Table S8 in the online supplement for descriptive statistics on all Study 2 outcomes.

Learning strategies. Learning strategies were assessed with 14 items rated on a Likert-type scale (see Table 1). Students responded to these items by reading the prompt "During the past month, how often have you…" and choosing one of the following response categories: 1 = "I did not do this in the past month", 2 = "I did this once or twice in the past month", 3 = "I did this several times in the past month", 4 = "I did this once a week in the past month", or 5 = "I did this every day in the past month." These items were created to represent planned behavior as related to the intervention lessons. Although, these behaviors were not meant to capture an underlying or latent construct, we categorized items based on the general actions implied by each statement. Five items were categorized as planning strategies, five items as persevering strategies, and four items as help-seeking strategies.

Analysis Plan

Our primary focus was on the *Conditional Average Treatment Effects* (CATE). That is, the treatment effect for each motivation profile. To obtain these CATE, we estimated the marginal effects of the interactions between treatment assignment and profile membership (see the supplement for more details). Our estimates were *intent-to-treat* or the effects on students'

Table 1

Learning Strategies Organized by Category (Columns) and Self-management Lesson (Rows)

Lesson	Planning	Persevering	Help-seeking
3. Set your goals	Created a work plan for completing assignments.	Imagined achieving a long-term goal to help you stay motivated and focused on school work.	
4. Build good habits	Made a plan to form new habits. Planned to do something you like as a reward (not necessarily an object) for getting something done.	Took some deep breaths to calm down when stressed.	
5. Deal with distractions	Made a priority list (time planner). Decided to prioritize schoolwork over playing.	Checked in with a friend to help you stick to your plans. Stuck to your priority list.	
6. Learn from failure		Practiced on difficult problems.	Asked your teacher for help on something you don't understand. Asked your parent or guardian for help on something you don't understand. Asked your classmates or friends for help on something you don't understand. Asked questions during class to the teacher.

learning strategies and academic achievement when their school was randomized to deliver the treatment, instead of the effects on students who actually participated in the program. Therefore, our results represent how much Indonesian students would benefit if their school was offered the program.

We focused on CATE greater than or equal to .04 standard deviations. We chose this effect size as a threshold for meaningful effects because a recent review of large-scale, rigorous randomized controlled trials in education found an average impact on academic outcomes of .04 standard deviations, 95% CI [.03, .05] (Lortie-Forgues & Inglis, 2019). Although this effect size seems small to declare that an impact is *meaningful* compared to previous meta-analyses of socioemotional interventions (e.g., Tanner-Smith, Durlak, & Marx, 2018) or traditional guidelines (e.g., Cohen, 1988), the studies included in Lortie-Forgues and Inglis' analyses are much more comparable to our study than the studies pooled by previous meta-analyses and traditional guidelines. For example, our study and the ones analyzed by Lortie-Forgues and Inglis recruited a large representative sample of schools and students, which tends to make impacts smaller than observed in studies with a small convenience sample (Tipton & Hedges, 2017).

In the case of effects on learning strategies, we focused on average impacts greater or equal than 2% (i.e., student i is 2% or .02 more likely to use strategy j), which is equivalent to an impact of .04 standard deviations. Using a transformed threshold was necessary because we used an ordered-probit model, and thus, the results were in probability metric.

Statistical model. We fit multivariate (i.e., multiple outcomes) multilevel models, where the effects of All Can Succeed on outcome Y_k for student i in school j were moderated by their motivation profile. Multivariate models were chosen to reduce the estimation and interpretation issues associated with multiple comparisons (Berkey, Hoaglin, Antczak-Bouckoms, Mosteller, & Colditz, 1998; Gelman, Hill, & Yajima, 2012; Hox, 2010). At the student level, we controlled for students' gender, socioeconomic status, and baseline school grades. We used a linear model for test scores and an ordered-probit model for learning strategies.

The models used Bayesian estimation to test whether motivation profiles boosted or inhibited the effects of All Can Succeed. A key advantage of Bayesian estimation over frequentist estimation is that the model results are presented as *distributions*, instead of point estimates. In our case, for example, the effects of All Can Succeed on disconnected students' test scores are presented as a distribution of likely values. Having a distribution of reasonable answers to our question is useful because researchers are often most interested in knowing the probability that an effect is either positive (i.e., did the intervention increase students' grades?) or negative (i.e., did the intervention decrease students' grades?) (Deke & Finucane, 2019).

Once the posterior distributions for Conditional Average Treatment Effects were obtained, we estimated the probability that these effects were positive (i.e., a posterior probability, represented as $Pr(CATE > 0)$). For example, $Pr(CATE > 0) = .87$ means that the probability that an effect is positive is .87 based on our data, priors, and model. This type of inference is different from conclusions drawn from the more common frequentist *p*-values, which are focused on rejecting a null hypothesis (e.g., the intervention had zero effect on students' scores), as opposed to describing how much support there is for an answer. Note that we use 89% Credible Intervals to describe effects because the resulting intervals are more stable in representing uncertainty than the more common 95% intervals (Makowski, Ben-Shachar, Lüdecke, 2019; McElreath, 2018). See supplement for more details about the models and estimation.

Results

Hypothesis 2.1: Students with disconnected and severely disconnected profiles exposed to All Can Succeed will increase their national test scores. All Can Succeed increased severely disconnected students' national science, math, and Indonesian scores (see

Table 2

Summary of Conditional Average Treatment Effects on National Exam Scores by Profile (Mean and 89% CI)

Exam	Growth	Multiple Goals	Disconnected	Severely Disconnected
		ACS		
Science	0.01 [-0.06, 0.09]	0.02 [-0.05, 0.09]	-0.01 [-0.08, 0.06]	0.15 [0.06, 0.24]
Math	-0.03 [-0.1, 0.04]	0.00 [-0.07, 0.07]	-0.02 [-0.09, 0.05]	0.08 [-0.01, 0.17]
English	0.02 [-0.05, 0.09]	0.02 [-0.05, 0.08]	-0.02 [-0.09, 0.05]	0.01 [-0.08, 0.1]
Indonesian	0.00 [-0.07, 0.07]	0.03 [-0.03, 0.09]	-0.02 [-0.09, 0.05]	0.06 [-0.03, 0.15]
		ACS+		
Science	0.04 [-0.04, 0.11]	0.00 [-0.07, 0.07]	0.01 [-0.06, 0.08]	0.08 [-0.01, 0.17]
Math	0.01 [-0.07, 0.08]	0.01 [-0.07, 0.08]	-0.01 [-0.08, 0.07]	0.03 [-0.06, 0.12]
English	0.04 [-0.03, 0.11]	0.03 [-0.03, 0.1]	0.01 [-0.06, 0.07]	-0.04 [-0.13, 0.05]
Indonesian	0.01 [-0.06, 0.08]	0.02 [-0.05, 0.08]	0.01 [-0.05, 0.08]	0.00 [-0.09, 0.09]

Note. ACS = Six lessons taught by a teacher, ACS+ = Six lessons plus supplemental activities for all 9th grade teachers. Conditional Average Treatment Effects for male students at the 50th percentile of the baseline grades distribution who had an average number of household assets (see supplement for treatment effects for other demographic groups). Effects are presented in standard deviation metric.

Table 2). These increases are equivalent to moving the average severely disconnected student from the 50th to the 56th percentile in science, 53rd percentile in math, and 52nd percentile in Indonesian. In addition, severely disconnected students exposed to ACS+ increased their science scores. This increase is equivalent to moving the average severely disconnected student from the 50th to the 53rd percentile in science scores. No other increases in scores due to the program were found for disconnected and severely disconnected students.

Unexpected impacts. ACS+ slightly increased growth students' science and English scores, while decreasing severely disconnected students English scores. For growth students, these increases are equivalent to moving the average growth student from the 50th to the 51st percentiles in science and in English. For severely disconnected students, their decrease is equivalent to moving the average severely disconnected student from the 50th to the 49th percentile in English.

Figure 3 illustrates, for science scores, how the CATE posterior distribution may overlap with zero and still be informative about the degree to which the program benefited students. For multiple goals and disconnected students, the distribution is almost equally split in negative and positive values. That is, there is so much uncertainty that we cannot infer whether the effect is either positive of negative. However, for growth and severely disconnected students, most of the posterior distribution is positive. Therefore, we can infer that the effect of ACS+ on science scores for these students is positive, though the effect for growth students was most likely below .10 standard deviations.

Hypothesis 2.2: Students with disconnected and severely disconnected profiles exposed to All Can Succeed will increase their use of learning strategies.

Overall, disconnected and severely disconnected students whose schools were assigned to ACS increased their use of four out of 14 strategies. Similarly, disconnected and severely

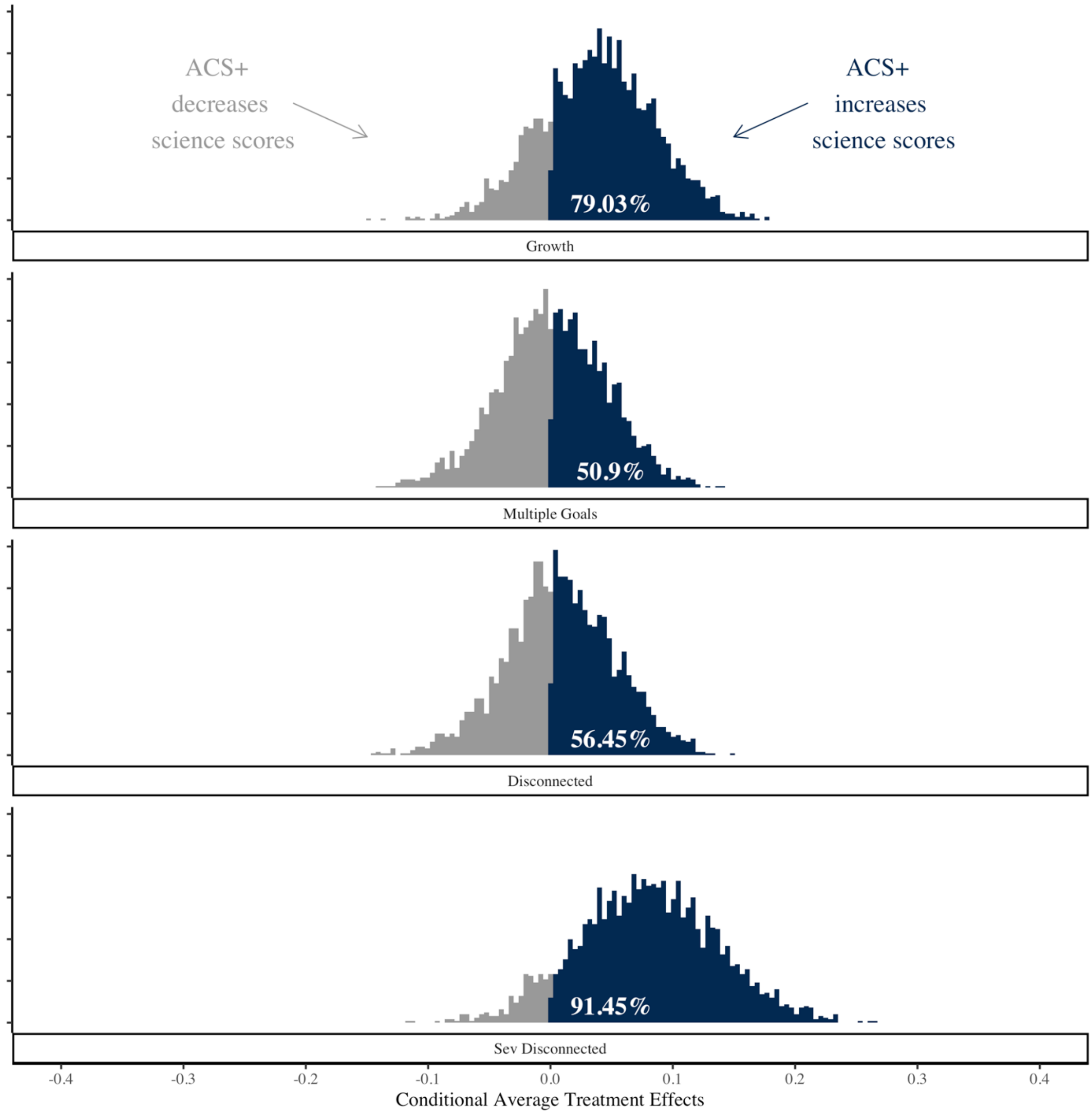

Figure 3. Percentages show how likely it is that the effect of the teacher intervention has a positive effect on a student's science scores, depending on their motivation profile. Blue bars are posterior draws (i.e., potential answers to our question) over zero.

disconnected students whose schools were assigned to ACS+ increased their use of three and six out of 14 strategies, respectively. Figures 4 and 5 illustrate these effects by plotting the posterior mean and 89% CI for each strategy. Although some of the lower bounds in the red error bars are below zero, most of the posterior distributions are concentrated around or above 2%, suggesting that All Can Succeed had positive impacts on learning strategies highlighted in red.

Across both versions of the program, disconnected students increased their use of time planners and were more likely to stick to their priorities than their counterparts in the control group. Across both versions of the program, severely disconnected students were more likely to stick to their priorities, reward themselves after completing a task, and plan to form new habits compared to their peers in the control group.

Unexpected impacts. Contrary to our expectations, both versions of the program made severely disconnected students less likely to check in with friends as a strategy to stick to their plans. Also unexpectedly, growth students increased their use of seven out of 14 strategies (see Tables S11-13 in the online supplement for more details). For example, growth students were more likely to check in with friends as a strategy to stick to their plans, ask friends for help with something they did not understand, and use time planners compared to growth students in the control group. Similarly, multiple goals students increased their use of five out of 14 strategies. For instance, multiple goals students were more likely to practice solving difficult problems, ask parents for help with something they did not understand, and stick to their priorities.

Discussion

In Study 2, we found that severely disconnected students showed small improvements in their national science and math scores after participating in ACS (i.e., six lessons taught by a teacher) and their science scores after participating in ACS+ (i.e., six lessons plus supplemental activities for other teachers). In addition, disconnected and severely disconnected students showed small improvements in their use of learning strategies that helped them deal with

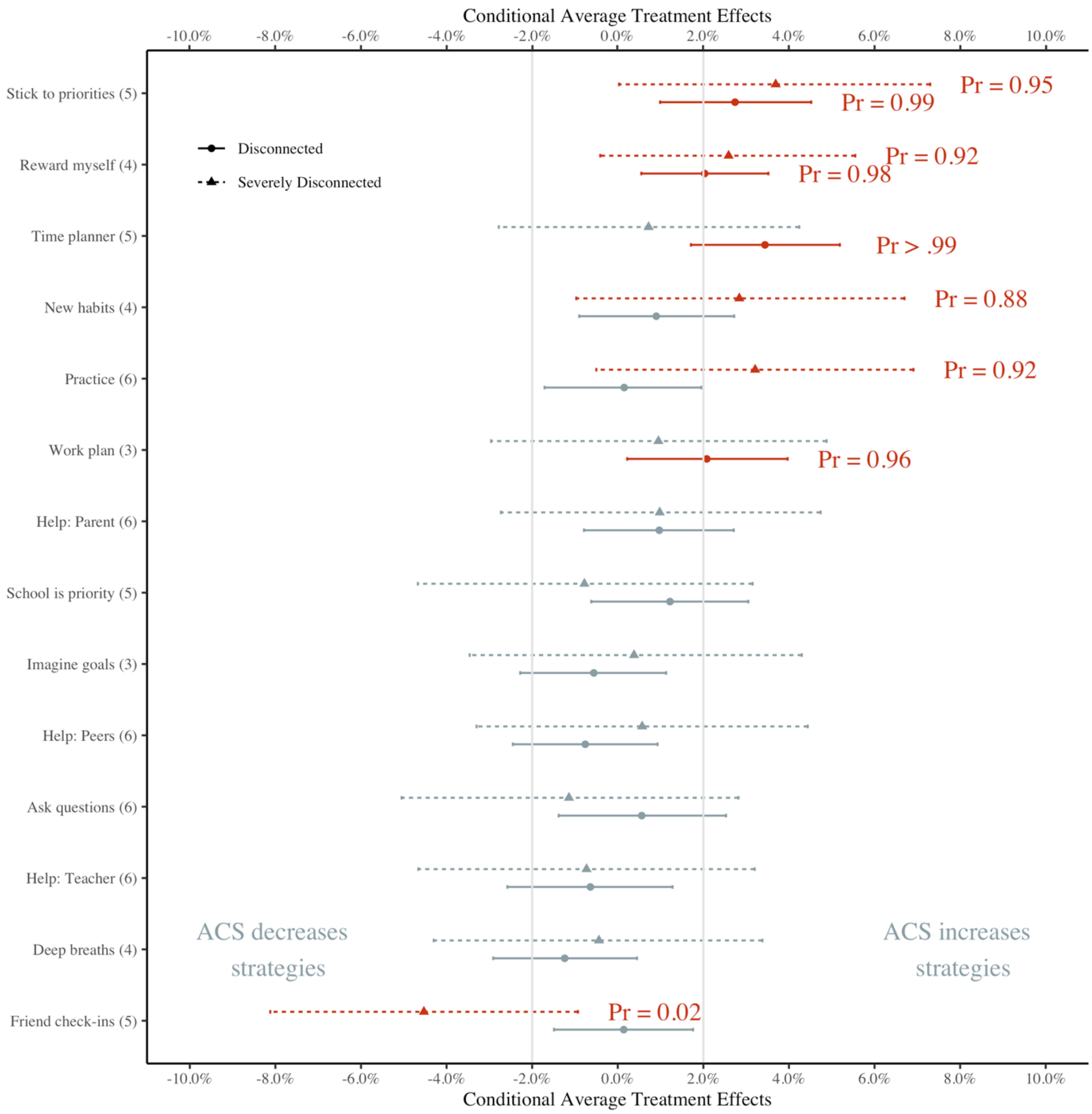

Figure 4. Conditional Average Treatment Effects of ACS on learning strategies among female students with median baseline grades who had average household assets. Red dots and error bars represent strategies for which the posterior mean was equal or greater than |2%|. A 2% change is equivalent to a .04 standard deviation change. *Pr* represents the probability that the posterior distribution of an effect is positive. Numbers in parentheses represent self-management lessons; (3) = “Set your goals”, (4) = “Build good habits”, (5) = “Deal with distractions”, and (6) = “Learn from failure”.

distractions and focus on schoolwork. Furthermore, growth and multiple goals students also benefited from the program. Growth students showed small increases in science and English scores and learning strategies that helped them deal with distractions and focus on schoolwork, whereas multiple goals students only showed increases in their use of learning strategies.

General Discussion and Conclusion

The present research sought to (1) identify distinct achievement motivation profiles among Indonesian adolescents and (2) test whether the effects of All Can Succeed on students' learning strategies and national test scores varied by these profiles. Our goals were motivated by the need in the field to identify the individual differences that explain who benefits from social and emotional school programs. In Study 1, we learned that adolescents show distinct motivation profiles. In Study 2, we learned that these motivation profiles help explain who benefited (and in what ways) from a universal social and emotional school intervention. The majority of students did not increase their academic achievement when their schools were offered the intervention. Yet, all students increased their use of learning strategies, with each profile showing a unique pattern of strategies impacted by the intervention.

Deepening the Understanding of the Growth Mindset Meaning System

Mindset theory has long argued that a person's beliefs about the malleability of intelligence are connected to how they view effort investment, how much they seek challenges, and how they frame struggle (Dweck, 1999, 2017; Hong et al, 1999). That is, a student's mindset offers a glimpse into a system of beliefs, goals, and behaviors with which individuals navigate their learning environment. Here, we found that this meaning system can take many forms or profiles. The most common profile was multiple goals (55% of students). Students with this profile have a coherent motivational system in which they are motivated to learn, believe in their ability to learn, seek out learning opportunities, and are willing to persist through challenges.

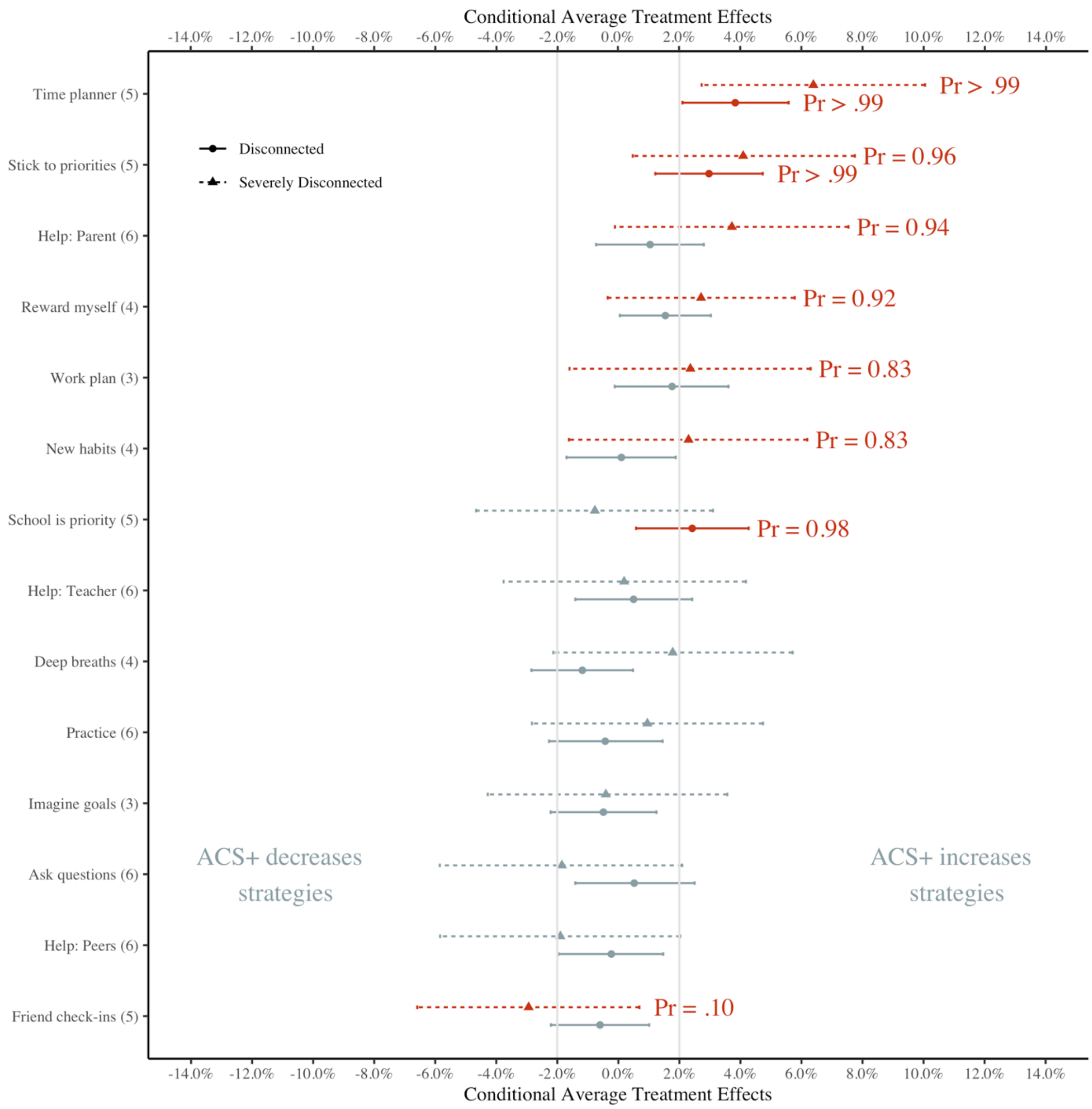

Figure 5. Conditional Average Treatment Effects of ACS+ on learning strategies among female students with median baseline grades who had average household assets. Red dots and error bars represent strategies for which the posterior mean was equal or greater than |2%|. A 2% change is equivalent to a .04 standard deviation change. *Pr* represents the probability that the posterior distribution of an effect is positive. Numbers in parentheses represent self-management lessons; (3) = "Set your goals", (4) = "Build good habits", (5) = "Deal with distractions", and (6) = "Learn from failure".

This profile, at first glance, appears fairly ideal. However, there is one goal that could hold the multiple goal students back: their performance-avoidance goals. These students avoid failing in front of others or taking risks that may lead to failing in public.

In contrast, growth students (12.97%) are not held back by performance avoidance goals. These students are motivated to learn, believe in their ability to learn, seek out learning opportunities, are willing to persist through challenge, and are not afraid of failing in public. This finding could provide insight into how we view the growth mindset meaning system. Growth mindset is typically conceptualized, and measured as, the belief that one's abilities can grow (Dweck, 1999). However, it is possible that greater insight into growth mindset might be gained from the understanding of motivational profiles that include not only growth mindset beliefs, but also related aspects of the mindset meaning system. For example, researchers typically focus on the mindset meaning system as an outcome of holding a growth mindset. However, it is possible that new insights could be gained through the study of the system as a whole, where capturing a *true* growth mindset depends on a host of different measures.

Future studies could focus on understanding the extent to which a growth mindset belief is integrated into adolescents' self-view, aligned with their values and goals, and is a core part of how they see themselves. A research program of this nature would enrich our understanding of how much these beliefs have an effect on adolescents' approach to learning, perceptions of their classroom environment, and behaviors in the classroom. In turn, the insights gained from these studies would help inform the conceptualization of growth mindset and lead to better measurement of the belief system.

An important limitation of the current study is the length and reliability of the scales used to identify the profiles. The scales were created using between three and six items each and they showed lower reliability estimates than estimates obtained in previous studies (e.g., CITES for each measure?). The main issue with using short scales is that we risk capturing a small facet of broad constructs (i.e., three items may not capture the full meaning of, for example, effort beliefs). The main issue with low reliability estimates is that we risk adding uncertainty in the process of capturing each construct. As a result, it is possible that we would have observed different profile shapes, had we used longer and more reliable measures. More research is needed to understand what Indonesian adolescents think when they answer these items and to identify the constructs that best represent Indonesian adolescents' motivation.

Motivation Profiles Uncover "Hidden" Insights

Identifying motivation profiles allowed us to uncover findings that are hidden by more traditional variable-centered approaches. For example, across all students, growth mindset and effort beliefs seemed not to be associated (i.e., their mean association is approximately zero; see Table S2 in the online supplement). However, the average association between variables hides the fact that, for some students, these two beliefs are connected and cooperate to frame students' experiences of challenge at school (e.g., growth students). In contrast, for other students, growth mindset and effort beliefs are not aligned and may not help students to thrive at school (disconnected and severely disconnected students). These two contrasting experiences would have remained hidden had we used a variable-centered approach and, as a result, we could have concluded that Indonesian students do not view effort investment as an important part in improving their own skills.

Another example of uncovering insights that would have remained hidden by variable-centered approaches is what we learned about who benefited from All Can Succeed. Although the previous analyses revealed that the average student did not increase their test scores when their school was assigned to the program (Johnson et al., 2020), our findings suggest that increasing test scores depended on students' motivation profiles. Severely disconnected and growth students showed small increases in their scores. Yet, since these two groups were smaller in size than the other two profiles, their increases are hidden when looking at whether the average student benefited from the program.

These new insights into the effects of the program suggest that future research could explore why multiple goals and disconnected students increased their use of learning strategies, but not their test scores. Perhaps, these students engaged with and learned from a few lessons in the intervention (i.e., the lessons focused on self-management), though such engagement and learning may have not been enough to boost their performance on tests. To better understand how much multiple goals and disconnected students engaged with All Can Succeed, future research could explore their thoughts and feelings during the program (i.e., what they thought and felt after each lesson), how they translated their learning into changes to their classroom participation and skill development, and whether they experienced unexpected benefits in outcomes the study did not assess. Answering these questions would provide valuable insights to tailor the intervention to afford more opportunities for multiple goals and disconnected students to engage with and learn from All Can Succeed.

Students' Background is Not Enough to Explain Who Benefits

The impacts on severely disconnected students' test scores resemble findings in previous growth mindset school intervention trials. Specifically, Yeager and colleagues (2019) found that a direct-to-student intervention (i.e., students learn about a growth mindset through an activity delivered online) increased lower-achieving students' school grades by approximately .11 standard deviations. Then, what is the added value of identifying motivation profiles if the findings are similar in magnitude to when researchers use baseline achievement to explain who benefits from social and emotional school interventions? The added value is that we learn more about who the students are who benefit from the intervention. Not all lower-achieving students had a severely disconnected profile. Specifically, 73% of severely disconnected, 62% of disconnected, 52% of multiple goals, and 37% of growth students were lower-achieving within their schools at baseline (see Tables S14-15 in the online supplement for more details). Moreover, when only lower-achieving students are included in the model (similar to the analyses in Yeager et al., 2019), we did not observe positive impacts for all profiles (see Table S16 in the online supplement). Put differently, showing lower school grades does not equate to showing a less adaptive motivation profile or to automatically benefiting from the program.

Given that students' motivation is arguably a more proximal reason to explain who engages with and learns from a program, we can use these findings to understand why students' background is often useful to detect variation in intervention impacts. For example, we could hypothesize that social and emotional school interventions with a focus on growth mindset may help severely disconnected students to "connect" their beliefs, goals, and behaviors in more adaptive ways. In the long term, such connections may change how these students face challenges and perform at school. Testing these expectations would involve asking questions such as: To what extent do severely disconnected students connect their beliefs, goals, and behaviors after participating in this type of intervention? Do disconnected students experience a similar process of aligning their beliefs, goals, and behaviors? If so, does such alignment change their performance during academic challenges?

The Program-Profile Fit Helps to Understand Why Not Every Student Benefited in the Same Way

Universal school programs target every student in the classroom, yet it is reasonable to expect that not all students engage with and learn from these programs (Greenberg & Abenovali, 2017). Presumably, universal programs create a learning environment where the needs of some students are met, while the needs of other students are not. The idea that characteristics of the learning environment interact with characteristics of individuals to boost (or inhibit) adolescents' thriving is not new to educational and developmental scientists (e.g., Eccles et al., 1993). However, we can still learn much about the degrees of fit (or misfit) created between students with different motivation profiles and school interventions. Understanding the program-profile fit would expand what we know about how universal programs can create unique experiences for diverse students.

Two examples of program-profile fit are worth mentioning here. First, growth students did not increase their test scores when their schools were assigned to ACS (i.e., six lessons only). Yet, they did show small increases in science and English scores when exposed to ACS+. The goal of ACS+ was to create an environment where the core ideas and skills of the program could flourish (Walton & Yeager, 2020), by creating opportunities for students to internalize such ideas and practices skills beyond the six lessons. Perhaps, this type of "fertile soil" met the needs of growth students in a way that ACS did not. That is, the fit between growth students and ACS+ may have boosted their academic achievement, even though these students were already more likely to be higher achievers at baseline than disconnected and severely disconnected students were.

Second, each profile benefited from All Can Succeed by increasing a unique combination of learning strategies. For instance, while growth students sought more help from peers to stick to their priorities when their schools were assigned to All Can Succeed, multiple goals and severely disconnected students sought less help from their peers to stick to their priorities. Presumably, this is because the program created an environment that invited growth students to publicly admit that they had much to learn. At the same time, the intervention may have created an environment where multiple goals and severely disconnected students felt discouraged to reveal they had much to learn in front of their peers. Yet, the same environment encouraged these students to increase their use of strategies to manage their time and organize their schoolwork. In this example, the degree to which All Can Succeed created fertile soil for the development of strategies depends on what each profile may have needed the most to thrive at school.

An important caveat to our interpretations is that the learning strategies we measured may not accurately represent the learning strategies used by students on a daily basis or the strategies that should be associated with better test scores. The purpose of these strategies was to capture changes in specific behaviors promoted by All Can Succeed lessons, instead of capturing learning strategies as a broad construct. Even though the strategies were based on previous research, more research is needed to understand the extent to which these specific strategies represent students' daily behaviors and their consequences.

In addition, students only showed small increases on their use of learning strategies. To better understand the size of this effects, it is useful to put them in context. At baseline, students reported high rates of strategy use (see Table S9 in the online supplement), suggesting that most students were applying planning, persevering, and help-seeking strategies often even before All Can Succeed was delivered. Hence, small increases in their weekly use of strategies may have translated into meaningful shifts in how students approached their learning. For example, small changes in their weekly use of time planners may translate into tangible gains in how they prepare for challenging tasks. However, more research is needed to learn about the practical implications of small shifts in weekly use of learning strategies among secondary students.

Are Low-Frequency Profiles and Their Associated Impacts Important for The Literature?

We acknowledge that others may see the findings on low-frequency profiles (e.g., severely disconnected) as less important than the findings associated with high-frequency profiles (e.g., multiple goals). However, low-frequency profiles are key to move the field forward. Taking low-frequency profiles seriously is necessary if we accept the idea that some profiles may be rare in the population. In turn, rare profiles reflect configurations of beliefs, goals, and behaviors that may expand our theories.

For example, in Yu & McLellan's (2020) study, only 9% of students showed a *disengaged* profile, whose shape was similar to our severely disconnected profile. This profile is important for mindset theory in particular because the theory is often described in terms of two ends of the mindset continuum (fixed versus growth) and the beliefs, goals, and behaviors that align with each extreme (Yu & McLellan, 2020). Thus, the misalignment of beliefs, goals, and behaviors shown by a severely disconnected student allows us to expand the theory to describe people beyond the fixed and growth extremes. That is, holding more of a growth mindset is associated with better academic achievement when this belief is aligned with learning goals and mastery behavior, not when they are disconnected. Therefore, we can make different predictions for students who show alignment in their beliefs, goals, and behaviors from those who do not.

Second, finding impacts among low-frequency profiles can be revealing for education leaders and policy makers when the sample is representative. For example, despite severely disconnected students representing only four percent of the sample (approximately 2,000 students), the representative nature of our sample allows us to generalize the findings to some extent to the school districts of the majority of the Java and Sumatra, where most of the population resides. From the perspective of education leaders and policy makers, shifting the behaviors and performance of 4% of students in a district, region, or a country could have important cascading effects on enrollment and success in post-secondary education.

Third, our inferences would not have improved by "merging" low-frequency profiles with other similar profiles to obtain larger subgroups. In our case, pooling disconnected and severely disconnected students into the same group would have resulted in inaccurate conclusions. In spite their similarities, disconnected students did not increase their test scores after participating in the program. Perhaps, because these two motivation profiles are more different than their shapes suggest. That is, there could be *qualitative* differences in their experiences at school that were not captured by our measures. As a result, these qualitatively distinct experiences could have made students perceive the core ideas and skills in the program in a different light, which did not help disconnected students to turn higher levels of planning and perseverance into higher test scores. To better understand the implications of holding a disconnected or a severely disconnected profile, future research could explore the qualitative differences between these profiles (e.g., studying how each profile reacts to challenges, daily struggles, and skill development).

Conclusion

Although multiple calls have been made in the literature to increase the use of person-centered approaches to understand treatment effect heterogeneity (e.g., Caldwell, Bradley, & Coffman, 2009; Lanza & Cooper, 2016), only a few studies have made the attempt (e.g., Low et al., 2016). We hope our findings contribute to fill this gap in the field and that others consider replicating and expanding our findings. We believe that learning more about motivation profiles can inform a wide range of research projects. For instance, an important question to be addressed is how stable these profiles are and how students transition to different profiles over time. Information on stability and change can improve our investigation of causal questions (i.e., Can we change students' profiles through intervention?), as well as correlational questions that describe students' daily experiences at school (i.e., What features of the learning environment are associated with transitions to a different profile over time?).

Looking ahead, we believe an important area of exploration is the comparison of motivation profiles between adolescents living in WEIRD and non-WEIRD nations. Above, we argued that the profiles found here conceptually replicated profiles found in previous studies, in spite of the methodological differences across studies. However, much more needs to be explored about how the same motivation profile may encapsulate different educational experiences for adolescents in opposite sides of the world. Therefore, learning more about adolescents in non-WEIRD countries may help us better understand how each motivation profile helps youth to thrive (or not) at school given their context.

References

August, G. J., Gewirtz, A., & August, G. J. (2018). Moving Toward a Precision-Based, Personalized Framework for Prevention Science : Introduction to the Special Issue. *Prevention Science*, *20*(1), 1-9. https://doi.org/10.1007/s11121-018-0955-9

Bae, C. L., & DeBusk-Lane, M. (2018). Motivation belief profiles in science: Links to classroom goal structures and achievement. *Learning and Individual Differences, 67*, 91–104. https://doi.org/10.1016/j.lindif.2018.08.003

Bakk, Z., & Vermunt, J. K. (2016). Robustness of stepwise latent class modeling with continuous distal outcomes. *Structural Equation Modeling, 23*(1), 20–31. https://doi.org/10.1080/10705511.2014.955104

Berkey, C. S., Hoaglin, D. C., Antczak-Bouckoms, A., Mosteller, F., & Colditz, G. A. (1998). Meta-analysis of multiple outcomes by regression with random effects. *Statistics in Medicine, 17*(22), 2537–2550. https://doi.org/10.1002/(SICI)1097-0258(19981130)17:22<2537::AID-SIM953>3.0.CO;2-C

Blackwell, L., 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

Bloom, H. S., & Michalopoulos, C. (2013). When is the Story in the Subgroups?: Strategies for Interpreting and Reporting Intervention Effects for Subgroups. *Prevention Science, 14*(2), 179–188. https://doi.org/10.1007/s11121-010-0198-x

Bloom, H. S., Raudenbush, S. W., Weiss, M. J., & Porter, K. (2017). Using Multisite Experiments to Study Cross-Site Variation in Treatment Effects: A Hybrid Approach

With Fixed Intercepts and a Random Treatment Coefficient. *Journal of Research on Educational Effectiveness, 10*(4), 817–842. https://doi.org/10.1080/19345747.2016.1264518

Bray, B. C., Lanza, S. T., & Tan, X. (2015). Eliminating Bias in Classify-Analyze Approaches for Latent Class Analysis. *Structural Equation Modeling, 22*(1), 1–11. https://doi.org/10.1080/10705511.2014.935265

Caldwell, L. L., Bradley, S., & Coffman, D. (2009). A person-centered approach to individualizing a school-based universal preventive intervention. *American Journal of Drug and Alcohol Abuse, 35*(4), 214–219. https://doi.org/10.1080/00952990903005932

Chen, J. A., & Tutwiler, M. S. (2017). Implicit theories of ability and self-efficacy: Testing alternative social cognitive models to science motivation. *Zeitschrift Fur Psychologie / Journal of Psychology, 225*(2), 127–136. https://doi.org/10.1027/2151-2604/a000289

Chittum, J. R., & Jones, B. D. (2017). Identifying pre-high school students' science class motivation profiles to increase their science identification and persistence. *Journal of Educational Psychology*, *109*(8), 1163–1187. https://doi.org/10.1037/edu0000176

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collie, R. J., Martin, A. J., Nassar, N., & Roberts, C. L. (2019). Social and Emotional Behavioral Profiles in Kindergarten: A Population-Based Latent Profile Analysis of Links to Socio-Educational Characteristics and Later Achievement. *Journal of Educational Psychology*, *111*(1), 170-187. https://doi.org/10.1037/edu0000262

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*(12), 1250–1262. https://doi.org/10.1016/j.jbusres.2008.01.013

Deke, J., & Finucane, M. (2019). Moving Beyond Statistical Significance: The BASIE (BAyeSian Interpretation of Estimates) Framework for Interpreting Findings from Impact Evaluations (OPRE Report 2019 35). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research, 38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Dilas, D. B., Mackie, C., Huang, Y., & Trines, S. (2019). *Education in Indonesia*. Retrieved from https://wenr.wes.org/2019/03/education-in-indonesia-2

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*(2), 166–174. https://doi.org/10.1080/00223890802634290

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development, 82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Dweck, C. S. (1999). *Self-theories: Their Role in Motivation, Personality and Development*. Philadelphia: Psychology Press.

Dweck, C. S. (2017). From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development. *Psychological Review, 124*(6), 689–719. https://doi.org/10.1037/rev0000082

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Iver, D. Mac. (1993). Development During Adolescence. The Impact of Stage-Environment Fit on Young Adolescents' Experiences in Schools and in Families. *American Psychologist, 48*(2), 90–101. https://doi.org/10.1037/0003-066x.48.2.90

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*(1), 5–12. https://doi.org/10.1037/0022-3514.54.1.5

Gelman, A., Hill, J., & Yajima, M. (2012). Why We (Usually) Don't Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness, 5*(2), 189–211. https://doi.org/10.1080/19345747.2011.618213

Gillet, N., Morin, A. J. S., & Reeve, J. (2017). Stability, change, and implications of students' motivation profiles: A latent transition analysis. *Contemporary Educational Psychology, 51*(August), 222–239. https://doi.org/10.1016/j.cedpsych.2017.08.006

Greenberg, M. T., & Abenavoli, R. (2017). Universal Interventions: Fully Exploring Their Impacts and Potential to Produce Population-Level Impacts. *Journal of Research on Educational Effectiveness, 10*(1), 40-67. https://doi.org/10.1080/19345747.2016.1246632

Grund, A. (2013). Motivational profiles in study-leisure conflicts: Quality and quantity of motivation matter. *Learning and Individual Differences*, *26*, 201–211. https://doi.org/10.1016/j.lindif.2013.01.009

Gunderson, E. A., Park, D., Maloney, E. A., Beilock, S. L., & Levine, S. C. (2018). Reciprocal relations among motivational frameworks, math anxiety, and math achievement in early elementary school. *Journal of Cognition and Development, 19*(1), 21–46. https://doi.org/10.1080/15248372.2017.1421538

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

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling*, *25*(4), 621–638. https://doi.org/10.1080/10705511.2017.1402334

Haydel, A. M., & Roeser, R. W. (2002). Affective and Conative Factors as Aptitude Resources in High School Affective and Conative Factors as Aptitude Resources in High School Science Achievement. *Educational Assessment, 8*(2), 163–189. https://doi.org/10.1207/S15326977EA0802

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hodis, F. A., Hattie, J. A. C., & Hodis, G. M. (2017). Investigating student motivation at the confluence of multiple effectiveness strivings: A study of promotion, prevention, locomotion, assessment, and their interrelationships. *Personality and Individual Differences*, *109*, 181–191. https://doi.org/10.1016/j.paid.2017.01.009

Hong, Y., Chiu, C., Dweck, C. S., M-S Lin, D., & 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

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York, NY: Routledge.

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? Psychological Bulletin, 136(3), 422–449. doi: 10.1037/a0018947

Jiang, D., Santos, R., Josephson, W., Mayer, T., & Boyd, L. (2018). A Comparison of Variable- and Person-Oriented Approaches in Evaluating a Universal Preventive Intervention. *Prevention Science, 19*(6), 738–747. https://doi.org/10.1007/s11121-018-0881-x

Johnson, H. C., Pinzón Hernández, D. A., Trzesniewski, K. H., Indrakesuma, T., Vakis, R., Perova, E., Mueller, N., De Martino, S., & Catalán Molina, D. (2020). *Can teaching growth mindset and self-management at school shift student outcomes and teacher mindsets?* Washington, D.C : World Bank Group.

Korpershoek, H., Kuyper, H., & van der Werf, G. (2015). Differences in students' school motivation: A latent class modelling approach. *Social Psychology of Education*, *18*(1), 137–163. https://doi.org/10.1007/s11218-014-9274-6

Lanza, S. T., & Cooper, B. R. (2016). Latent Class Analysis for Developmental Research. *Child Development Perspectives, 10*(1), 59–64. https://doi.org/10.1111/cdep.12163

Lau, S., & Roeser, R. W. (2008). Cognitive abilities and motivational processes in science achievement and engagement: A person-centered analysis. *Learning and Individual Differences, 18*(4), 497–504. https://doi.org/10.1016/j.lindif.2007.11.002

Lazarides, R., Dicke, A. L., Rubach, C., & Eccles, J. S. (2020). Profiles of motivational beliefs in math: Exploring their development, relations to student-perceived classroom characteristics, and impact on future career aspirations and choices. *Journal of Educational Psychology, 112*(1), 70–92. https://doi.org/10.1037/edu0000368

Lerner, R. M. (2015). Promoting positive human development and social justice: Integrating theory, research and application in contemporary developmental science. *International Journal of Psychology, 50*(3), 165–173. https://doi.org/10.1002/ijop.12162

Lerner, R. M., Agans, J. P., DeSouza, L. M., & Hershberg, R. M. (2014). Developmental Science in 2025: A Predictive Review. *Research in Human Development, 11*(4), 255–272. https://doi.org/10.1080/15427609.2014.967046

Liem, G. A. D., Martin, A. J., Porter, A. L., & Colmar, S. (2012). Sociocultural antecedents of academic motivation and achievement: Role of values and achievement motives in achievement goals and academic performance. *Asian Journal of Social Psychology, 15*(1), 1–13. https://doi.org/10.1111/j.1467-839X.2011.01351.x

Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., & Hill, N. E. (2018). Multiple Pathways to Success: An Examination of Integrative Motivational Profiles Among Upper Elementary and College Students. *Journal of Educational Psychology, 110*(7), 1026-1048. https://doi.org/10.1037/edu0000245

Lortie-Forgues, H., & Inglis, M. (2019). Rigorous Large-Scale Educational RCTs Are Often Uninformative: Should We Be Concerned? *Educational Researcher, 48*(3), 158–166. https://doi.org/10.3102/0013189X19832850

Low, S., Smolkowski, K., & Cook, C. (2016). What Constitutes High-Quality Implementation of SEL Programs? A Latent Class Analysis of Second Step Implementation. *Prevention Science, 17*(8), 981–991. https://doi.org/10.1007/s11121-016-0670-3

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). BayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. doi: 10.21105/joss.01541

Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person- and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling, 16*(2), 191-225. https://doi.org/10.1080/10705510902751010

Masyn, K. E. (2017). Measurement Invariance and Differential Item Functioning in Latent Class Analysis With Stepwise Multiple Indicator Multiple Cause Modeling. *Structural Equation Modeling*, 24(2), 180–197. https://doi.org/10.1080/10705511.2016.1254049

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd Edition). Chapman; Hall/CRC.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., … Midgley, C. (2000). *Manual for the Patterns of Adaptive Learning Scales*. Ann Arbor. Retrieved from http://www.umich.edu/~pals/PALS 2000_V13Word97.pdf

Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide* (Eighth Edition). Los Angeles, CA: Muthén & Muthén.

Napolitano, C., Catalán Molina, D., Johnson, H. C., Oswald, F. L., Hernandez, D. A. P., Tiwari, A., … Trzesniewski, K. (2021, March 17). Are growth mindset, mastery orientation, and grit promising for promoting achievement in the Global South? Psychometric evaluations among Indonesian adolescents. https://doi.org/10.31234/osf.io/dx4et

Oga-Baldwin, W. L. Q., & Fryer, L. K. (2018). Schools can improve motivational quality: Profile transitions across early foreign language learning experiences. *Motivation and Emotion, 42*(4), 527–545. https://doi.org/10.1007/s11031-018-9681-7

Olivera-Aguilar, M., & Rikoon, S. H. (2018). Assessing Measurement Invariance in Multiple-Group Latent Profile Analysis. *Structural Equation Modeling, 25*(3), 439–452. https://doi.org/10.1080/10705511.2017.1408015

Outes-León, I., Sánchez, A., & Vakis, R. (2020). The Power of Believing You Can Get Smarter: The Impact of a Growth mindset intervention on Academic Achievement in Peru. *World Bank Policy Research Working Paper*, 9141. Washington, D.C.: World Bank.

Page, L. C., Feller, A., Grindal, T., Miratrix, L., & Somers, M.-A. (2015). Principal Stratification: A Tool for Understanding Variation in Program Effects Across Endogenous Subgroups. *American Journal of Evaluation, 36*(4), 514–531. https://doi.org/10.1177/1098214015594419

Parhiala, P., Torppa, M., Vasalampi, K., Eklund, K., Poikkeus, A. M., & Aro, T. (2018). Profiles of school motivation and emotional well-being among adolescents: Associations with math and reading performance. *Learning and Individual Differences, 61*(December 2017), 196–204. https://doi.org/10.1016/j.lindif.2017.12.003

Park, D., Gunderson, E. A., Tsukayama, E., Levine, S. C., & Beilock, S. L. (2016). Young children’s motivational frameworks and math achievement: Relation to teacher-reported instructional practices, but not teacher theory of intelligence. *Journal of Educational Psychology, 108*(3), 300–313. https://doi.org/10.1037/edu0000064

Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology, 32*(1), 8–47. https://doi.org/10.1016/j.cedpsych.2006.10.003

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

Porter, T., Catalán Molina, D., Blackwell, L., Roberts, S., Quirk, A., Duckworth, A. L., & Trzesniewski, K. H. (2020). Measuring Mastery Behaviours at Scale: The Persistence, Effort, Resilience, and Challenge- Seeking (PERC) Task. *Journal of Learning Analytics, 7*(1), 5–18. https://doi.org/10.18608/jla.2020.71.2

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Raudenbush, S. W., & Bloom, H. S. (2015). Learning about and from a distribution of program impacts using multisite trials. *American Journal of Evaluation, 36*(4), 479–499. https://doi.org/10.1177/1098214015600515

Roeser, R. W., Strobel, K. R., & Quihuis, G. (2002). Studying early adolescents' academic motivation, social-emotional functioning, and engagement in learning: Variable- and person-centered approaches. *Anxiety, Stress and Coping, 15*(4), 345–368. https://doi.org/10.1080/1061580021000056519

Schwinger, M., Steinmayr, R., & Spinath, B. (2016). Achievement Goal Profiles in Elementary School: Antecedents, Consequences, and Longitudinal Trajectories. *Contemporary Educational Psychology, 46*, 164–179. https://doi.org/10.1016/j.cedpsych.2016.05.006

Solmeyer, A. R., & Constance, N. (2015). Unpacking the "Black Box" of Social Programs and Policies: Introduction. *American Journal of Evaluation, 36*(4), 470–474. https://doi.org/10.1177/1098214015600786

Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically Based Mean Effect Size Distributions for Universal Prevention Programs Targeting School-Aged Youth: A Review of Meta-Analyses. *Prevention Science*, *19*(8), 1091–1101. https://doi.org/10.1007/s11121-018-0942-1

Tipton, E. (2020). Beyond generalization of the ATE: Designing randomized trials to understand treatment effect heterogeneity. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 1–18. https://doi.org/10.1111/rssa.12629

Tipton, E., & Hedges, L. V. (2017). The Role of the Sample in Estimating and Explaining Treatment Effect Heterogeneity. *Journal of Research on Educational Effectiveness, 10*(4), 903–906. https://doi.org/10.1080/19345747.2017.1364563

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. https://doi.org/10.1177/0963721420904453

Weissberg, R. P., J. A. Durlak, C. E. Domitrovich, and T. P. Gullotta. 2015. "Social and emotional learning: Past, present, and future." in J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, and T. P. Gullotta (Eds.). *Handbook for social and emotional learning: Research and practice*. New York, NY, US: Guilford.

World Bank (2019). Instilling a Growth Mindset in Indonesia. *eMBeD brief*. Washington, DC: World Bank. Retrieved from https://bit.ly/3bvscEA

World Bank (2021). *World Development Indicators.* Retrieved from https://data.worldbank.org/indicator/SP.POP.TOTL

Wormington, S. V., Corpus, J. H., & Anderson, K. G. (2012). A person-centered investigation of academic motivation and its correlates in high school. *Learning and Individual Differences, 22*(4), 429–438. https://doi.org/10.1016/j.lindif.2012.03.004

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., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature, 573*, 364–369. https://doi.org/10.1038/s41586-019-1466-y

Yu, J., & McLellan, R. (2020). Same mindset, different goals and motivational frameworks: Profiles of mindset-based meaning systems. *Contemporary Educational Psychology, 62*, 101901 https://doi.org/10.31234/osf.io/khd8c

Zvoch, K. (2012). How Does Fidelity of Implementation Matter? Using Multilevel Models to Detect Relationships Between Participant Outcomes and the Delivery and Receipt of Treatment. *American Journal of Evaluation, 33*(4), 547–565. https://doi.org/10.1177/1098214012452715

Supplemental Materials

Study 1

Method. To identify the motivation profiles, we used six survey measures and one behavioral task, all of which were adapted, translated, and back translated (the students took the survey in Indonesian). Below, student background measures are described first. Then, we describe the profile measures: Growth mindset, effort beliefs, grit, challenge-seeking preferences, learning goals, performance-avoidance goals, and mastery behavior. See tables S1 and S2 for descriptive statistics of and associations among these measures.

Students' background: Gender, socioeconomic status, and school grades. Gender and several indicators of socioeconomic status were self-reported by students during survey administration. Gender was reported with a dichotomous item (0 = Male, 1 = Female). Students' socioeconomic status was assessed through dichotomous questions about household assets (e.g., a car, a refrigerator, and cellphones with access to internet), experiencing hunger in the previous 30 days, and missing school because of work. Principal components analysis was used to combine assets into a standardized score.

Baseline grades were obtained from schools during data collection. Grades range from 0 to 100 and the passing grade is 75 for most subjects. We averaged math, science, and English (a required second language class). In addition, we used Indonesian grades as a proxy for reading comprehension.

Growth mindset. Growth mindset captures the degree to which students believe that their intelligence can change. Students low in this measure believe that their intelligence is a fixed trait, whereas students high in growth mindset believe that their intelligence is a malleable characteristic. Students' mindset has been found to be associated with academic achievement, beliefs about effort, learning goals, attributions about success and failure, and challenge-seeking (Blackwell, Trzesniewski, & Dweck, 2007; Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013; Yeager et al., 2016).

Growth mindset was measured by four items rated on a Likert-type scale, ranging from (1) Not true at all to (5) Completely true. An example item is "You cannot change how smart you are." These items were administered targeting the fixed end of the mindset continuum—i.e., when students believe their intelligence is a fixed trait—to reduce acquiescence bias. We, then, reverse-coded item responses so that higher scores represent the growth end of the mindset continuum (α = .65). These items were adapted from the Theory of Intelligence survey (Dweck, 1999).

Effort beliefs. Effort beliefs assesses the extent to which students believe that exerting effort will lead to improved ability. Students high on this measure see effort investment as a necessary part of learning and doing well in school. Effort beliefs have been found to be associated with a growth mindset, helpless attributions, and study strategies (Blackwell et al., 2007).

Effort beliefs were assessed by six items rated on a Likert-type scale, ranging from (1) Not true at all to (5) Completely true (α = .57). An example item is "Doing well in school requires hard work and effort." These items were adapted from Blackwell et al. (2007).

Grit. Grit captures the degree to which students show perseverance and passion for long-term outcomes or goals (Duckworth, Peterson, Matthews, & Kelly, 2007). Students high on grit tend to work through challenges, failure, and adversity to accomplish goals, even over a span of several years. Moreover, students high on grit continue pursuing their goals even when faced with disappointment and boredom. Grit has been found to be associated with academic achievement, intrinsic and extrinsic motivation, and a sense of purpose in schoolwork (Duckworth et al., 2007; Duckworth & Quinn, 2009; Yeager et al., 2014).

Grit was measured with four items rated on a Likert-type scale, ranging from (1) Almost never to (5) Almost always (α = .59). An example item is "I finish whatever I start." These items were adapted from the Short Grit Scale (Duckworth & Quinn, 2009). These items capture the perseverance of effort dimension of grit. However, for simplicity, we refer to the measure as grit throughout the manuscript.

Challenge-seeking preferences. Challenge-seeking captures whether students prefer to engage in challenging or easy school-related activities. Students high on challenge-seeking welcome difficult tasks (Dweck, 1999, 2017). Whether students seek challenging tasks has been found to be associated with their mindset and achievement goals (Jagacinski, Kumar, & Kokkinou, 2008; Lee & Kim, 2014; Yeager et al., 2016).

Challenge-seeking was assessed with three dichotomous items for which students displayed a preference. An example item is "If you had to choose between having easy or difficult class work, which one would you choose? Easy class work or Difficult class work." This measure was adapted from items used in past growth mindset research (Dweck, 1999). We do not report alpha as a measure of internal consistency because this measure was not adapted to capture a latent construct. Instead, it was adapted to summarize specific challenge-seeking preferences. In other words, this is a formative measure that does not rely on correlations among items (Coltman, Devinney, Midgley, & Venaik, 2008; Diamantopoulos & Winklhofer, 2001).

Achievement goals. Achievement goals capture the targets that students have in mind when thinking about their learning process. Here, the focus is on two different types of goals: learning and performance-avoidance goals (Elliott & Dweck, 1988; Sommet & Elliot, 2017).

When students hold learning goals, they are driven by understanding the material and developing their competence. When students hold performance-avoidance goals, they are driven by a fear of looking incompetent. Compared to students who hold learning goals, students holding performance-avoidance goals are more likely to hold a fixed mindset, to show worse performance outcomes, and to display less interest in academic activities (Burnette et al., 2013; Hulleman, Schrager, Bodmann, & Harackiewicz, 2010).

Learning goals were assessed with three items rated on a Likert-type scale, ranging from (1) Not true at all (5) Completely true (α = .58). An example item is "It's important to me that I completely understand my class work." Performance-avoidance goals were assessed with four items rated on a Likert-type scale, ranging from (1) Not true at all (5) Completely true (α = .66). An example item is "I would only answer a question in class if I knew I was right." This measure was adapted from two subscales of the Patterns of Adaptive Learning Scale (Midgley et al., 2000).

Mastery behavior. As a behavioral measure of motivation, we used the Persistence Effort Resilience Challenge-seeking (PERC) Task (Porter et al., 2020). In this task, students complete a series of puzzles of different levels of difficulty. The puzzles are divided into different sets and student responses to each set are used to assess their levels of persistence, effort, resilience, and challenge-seeking. We will use a mastery behavior composite score that summarizes students' behavioral levels of persistence, effort, resilience, and challenge-seeking. This measure is also a formative scale, and thus, we do not report alpha as an index of the reliability of this measure.

Analysis plan. To identify motivation profiles, we used 10-fold cross-validation. This strategy consists of partitioning a sample into 10 random subsamples that are first used to train (or calibrate) ideas about the hypothesized number of profiles and, then, to test the validity of those hypotheses with a different sample. The cross-validation process would follow ten iterative steps. In the first step, nine of the ten subsamples (90% of the participants) are merged to create a *training* dataset. In a series of nested models, the training data are used to test the hypothesized number of profiles. Then, the same number of models are fitted to the remaining 10% of participants or the *validation sample*. For the validation sample, the parameters in each model are fixed to the values found in the training data. In our case, the means for each construct will be fixed in validation sample models. Last, the $-2 * Loglikelihood$ (-2LL) obtained in the validation sample is saved.

This process is repeated nine more times; each time, a different partition serves as validation sample and the others are merged into the training sample. By the end of this process, the -2LL values can be compared across models to discern the number of profiles that is more likely to be replicated given different sample configurations. This decision is based on a) minimizing large differences in the average -2LL across models (i.e. the mean -2LL in one model would not overlap with the lower bound of the previous model's error bar) and b) balancing model parsimony with improvements in fit (Grimm et al., 2017).

Grimm and colleagues provided criteria for selecting the number of folds, but they were explicit in stating that good rules of thumb have not been developed yet. One of their suggestions is to work with a large number of folds to identify low prevalence profiles, which is what we expect to find. We base this expectation on previous studies on motivation profiles that have consistently found solutions that contain at least one profile that represents a small percentage of the sample (e.g. 5% of approximately 6,000 students in Olivera-Aguilar et al., 2017; 3% of approximately 550 students in Schwinger et al., 2016; and 9% of approximately 500 students in

Gillet et al., 2017. In motivation research, low prevalence profiles are often meaningful because they represent a type of student who needs attention form their teachers or school.

Grimm and colleagues compared the findings across 5-, 10-, and 100-fold strategies. Although they did not define what a "large number of folds" means, they did demonstrate that 10-fold strategy showed similar results to a 100-fold strategy. In our case, a 10-fold strategy means that 10% of the participants would be used to validate the results of our *training* models—in other words, we would use around 5,000 students to validate results 10 times during the process. This large number of folds will provide more estimates of bias—e.g. -2LL—, and thus, will provide more precision than using only 5 folds. Moreover, the large number of participants in the validation samples will allow us to argue that low prevalence profiles are meaningful groups to pay attention to. For example, if we found a profile that consisted of 3% of our validation sample, our inferences would apply to over 150 students in each subsample or 1,500 students in the overall sample.

Departures from the registered analysis plan. We made two decisions that were not outlined in the registered analysis plan. Both decisions were related to the models that tested whether motivation profiles were invariant across students' gender, socioeconomic status, and academic achievement. First, Masyn (2017) outlined a likelihood ratio test to compare nested models based on the maximum likelihood estimator. Since we used restricted maximum likelihood to identify motivation profiles, we included the scaling correction factors to adjust the likelihood ratio test (Satorra, 2000).

The likelihood ratio test then becomes:

$$TR_d = -2(L_0 - L_1)/cd$$

where L_0 is the loglikelihood of the nested model and L_1 is the loglikelihood of the comparison model. The nested model is the more restrictive model with more degrees of freedom than the comparison model. Last, cd is:

$$cd = (p_0 * c_0 - p_1 * c_1)/(p_0 - p_1)$$

where cd is the difference test scaling correction, p_0 is the number of parameters in the nested model, c_0 is the scaling correction factor in the nested model, p_1 is the number of parameters in the comparison model, and c_1 is the scaling correction factor in the comparison model.

A second departure from the registered analysis plan was related to making decisions about which items showed profile non-invariance. The procedure outlined by Masyn (2017) is organized in steps in which decisions are made on the degree to which profiles are non-invariant across levels of a covariate. Step 2 is focused on identifying items that are a source of noninvariance and this decision is based on the likelihood ratio test. As described in the *Measurement invariance results* section, Step 2 resulted in all items showing evidence of noninvariance. However, reaching the conclusion that all items are noninvariant without further exploration would have resulted in an unreasonably complicated classification model for Study 2. Therefore, instead of focusing on the likelihood ratio test, we focused on the effect sizes of potentially noninvariant items to move on to the next steps outlined by Masyn (2017).

Results. Tables S1 and S2 present descriptive statistics for each measure and correlations among them.

Cross-validation results. The 10-fold cross-validation process showed that two solutions fit the data and replicated well: four and five profiles. The result of the cross-validation process is a set of fit indices (i.e., $-2 * Loglikelihood;\ -2LL$) for each of the ten validation samples from the one- to the five-profile solutions (Grimm, Mazza, & Davoudzadeh, 2017). To decide which solution is more likely to be replicated given different sample configurations, we proposed two criteria: a) minimizing large differences in the average -2LL across models (i.e., the mean -2LL in one model would not overlap with the lower bound of the previous model's error bar) and b) balancing model parsimony with improvements in fit. Although the first criterion would suggest retaining the five-profile solution (see Figure S1), we decided to retain the four-profile solution for a few reasons. First, the four-profile solution is more parsimonious. That is, the improvement in fit does not seem to translate into meaningful improvements in the classification of individual students (see Figure S2). Second, the five-profile solution identified a profile that only represents around 1% of the validation samples and does not show a novel shape compared to the other profiles.

Measurement invariance results. We found a small degree of profile noninvariance across students' gender, socioeconomic status, and academic achievement (see a summary of results in Table S3). That is, student background characteristics did not substantively change the shapes in Figure 1, but they did influence whether students in the same profile responded in similar ways to our measures. Specifically, female and male students responded differently to our measures of grit, effort, and mastery behavior. In addition, students with different socioeconomic background (e.g., low vs high number of household assets) responded differently to our measure of mastery behavior. Last, students with different baseline grades responded differently to our measures of growth mindset, challenge-seeking preferences, and mastery behavior. In spite of these differences, the shapes of the profiles remained similar between students of different backgrounds (see Figure S3 as an example of how profile shapes varied when covariates were included in the model). Below, we describe how we arrived at these results.

Steps in measurement invariance testing. Tables S4-7 show the likelihood ratio tests that we used to arrive at the conclusions above. Each Table summarizes model comparisons along the steps outlined by Masyn (2017). Table S4 shows Step 1, which aims to identify whether any evidence of non-invariance can be detected. The results showed that a model with extreme evidence of non-invariance fit better than a model that assumed complete profile invariance. Therefore, we moved to Step 2, which aims to identify items that are potentially noninvariant (Table S5).

Step 2 resulted in all items showing evidence of non-invariance. However, reaching the conclusion that all items are noninvariant without further exploration would have resulted in an unreasonably complicated classification model for Study 2. Therefore, instead of focusing on the likelihood ratio test, we focused on the effect sizes (coefficients larger than |.10|) of potentially non-invariant items to move on to the next steps outlined by Masyn (2017). For gender, we selected mastery behavior, effort beliefs, and grit for further exploration of non-invariance. For assets, we selected mastery behavior for further exploration of non-invariance. For the two school grade covariates, selected mastery behavior, challenge-seeking, and growth mindset for further exploration of non-invariance.

Table S6 shows Step 3, which compares a model that includes nonuniform non-invariance for some items (selected in Step 2) to the two more extreme models tested in Step 1. This model that included nonuniform non-invariance for some items fit better than the model that assumed complete profile invariance, but worse than the model that assumed extreme nonuniform non-invariance. We then moved forward to Step 4, which tests whether the items selected in Step 2 showed uniform non-invariance (as opposed to nonuniform). As suggested by Table S7, we concluded that assets and gender had uniform profile non-invariance over the items selected in Step 2, whereas school grades had nonuniform profile non-invariance over the items selected in Step 2.

Classification model. To balance potential issues with non-invariance and issues with model complexity, we retained a model that included gender, socioeconomic status, and academic achievement as predictors of motivation profiles and the noninvariant measures (see more details in the Mplus output at https://osf.io/cg4ue/?view_only=6122ff7b0ad648e2b8401a7db2cde64f). Based on this model, we classified students into their most likely profile to use in Study 2. Accounting for students' background in the classification model resulted in a more accurate classification of students. For example, 17% of female students would have been wrongly classified into the growth profile had we not accounted for non-invariance, when they should be classified as multiple goals. Similarly, 10% of male students would have been wrongly classified as growth, when they should be classified as multiple goals.

Associations between students' background and motivation profile membership. Our main findings were that (1) covariates were least influential in the classification of students as multiple goals (compared to the growth profile) and (2) students' household assets were the least influential covariate across profiles compared to students' gender and grades (see Figure S4). Female students were less likely to show a disconnected or severely disconnected profile, instead of a growth profile, compared to male students. In addition, higher school grades were associated with higher probability of showing a growth profile compared to a multiple goals, disconnected, or severely disconnected profiles.

Study 2

Method. Study 2 participants were the same as Study 1 participants. At endline, a portion of students' test scores and learning strategies were missing. Most of the missing data was due to cluster-level attrition (i.e., 18% of schools did not report test scores for students). However, schools in every condition experienced similar levels of attrition, resulting in 6% differential attrition between ACS and control schools and 3% differential attrition between ACS+ and control schools. Following current standards for school intervention studies (What Works Clearinghouse, 2020), we assumed that missing test score data induced low levels of bias in our results. In terms of learning strategies, cluster-level missing data was even less of a problem with overall and differential attrition rates below 1%. To handle missing data on predictors, we relied on Bayesian estimation, which in Mplus is comparable to full information maximum likelihood estimation. Specifically, assuming data on predictors are missing at random, the algorithm uses information from every student to estimate regression parameters (Asparouhov & Muthén, 2010).

Randomization and implementation. The [blinded for peer review], in collaboration with [blinded for peer review] and the Indonesian Ministry of Education and Culture, designed an efficacy study in which schools were randomized to one of three conditions: (a) business-as-usual control (n = 699), (b) All Can Succeed (ACS) (n = 699), or (c) All Can Succeed Plus (ACS+) (n = 699). This probabilistic and representative sample of Indonesian schools ensured equivalent numbers of schools in each condition by island and school district. At baseline, motivation profiles showed balanced frequencies within each assignment condition (see Table S10).

Each school chose a counselor or a teacher to deliver All Can Succeed lessons to 9th grade students. In addition, schools allocated to ACS+ invited other 9th grade teachers to deliver supplemental activities in their class. As often occurs in school intervention studies, not every school complied with their assignment. Among schools assigned to ACS, 63% delivered all six lessons, whereas 81% delivered at least one lesson. Among schools assigned to ACS+, only 37% delivered all six lessons and two supplemental activities, while 54% delivered at least one lesson and one supplemental activity. The average lesson duration was 51 minutes, with 72% of schools reporting an average lesson time between 35 and 55 minutes.

All Can Succeed. All Can Succeed is a social and emotional learning classroom-based curriculum focusing on growth mindset and self-management. The goals of the program are to (1) reframe students' experiences of struggle at school, (2) promote the learning strategies students need to succeed in secondary school, and (3) raise students' educational and employment aspirations. These goals were motivated by the key transition Indonesian students experience at the end of 9th grade, when they have to decide to continue on a general education track or enroll in a vocational education track. Given the high dropout rates during this transition and the country needs for more youth to continue to higher education (Dilas, Mackie, Huang, & Trines, 2019), All Can Succeed was designed to promote the idea among students that they could aspire to a higher education degree and it offered tools for students who wanted to succeed academically.

The program builds on previous intervention studies in the US (e.g., Blackwell et al., 2007); Paunesku et al., 2015); Yeager et al., 2019), Peru (Outes-León, Sánchez, & Vakis, 2020), and a prior pilot study in Indonesia (World Bank, 2019). The previous study in the country offered two 40-minute lessons focused on teaching students about the malleability of their intelligence and skills. Building on these two lessons, the research team designed four more lessons that integrated teaching students about a growth mindset and self-management, the ability to regulate emotions and behaviors, motivate oneself, and work towards achieving personal and academic goals (Weissberg, Durlak, Domitrovich, & Gullota, 2015).

In All Can Succeed, guidance counselors or teachers taught six lessons (45 minutes each) on growth mindset and self-management. Each lesson is focused on a comic book story that introduces a topic to be discussed by the whole class. After reading the story, the teacher discussed the story's core ideas with the class. Then, students wrote individual reflections tying the core ideas in the comic to personal experiences (e.g., experiences of struggle at school). Next, students shared their reflections with a classmate. After, teachers invited students to share what they learned through individual reflections and conversations with peers. Each lesson closed with a "group yell" that was designed to reiterate core ideas and engage students as a group in the process of improving their skills.

Lesson 1 focused on the idea that intelligence is not a fixed or inherent trait and that the brain is like a muscle that gets stronger with training. Lesson 2 focused on challenging student biases about how gender and socioeconomic status are linked to aspirations and motivation. Lesson 3 introduced students to a systematic approach to goal setting based on the Wish–Outcome–Obstacle–Plan (WOOP) framework (Oettingen & Gollwitzer, 2001). Lesson 4 focused on social and emotional skills to overcome common obstacles that could impede the achievement of goals, including building new, positive habits through a trigger–action–reward loop and using deep breathing to manage negative emotions. Lesson 5 continued the work on goal achievement by teaching students about prioritization and avoiding distraction. Finally, Lesson 6 discussed learning from failure, including ideas such as not being afraid to ask questions and perceiving failure as an opportunity to learn.

In the All Can Succeed Plus condition, the six lessons were supplemented by activities and materials delivered by other teachers in their own classes. These activities and materials were meant to create a learning environment in which students could transfer to and practice in other contexts the skills learned during the comic book-based lessons. First, there were guides and materials for two activities that could be integrated into classroom routines: “I see All Can Succeed (ACS) behavior” and “Sharing Board.” The goal of “I See ACS Behavior” was to encourage students to demonstrate and recognize examples of behaviors promoted during the week’s lesson. At the beginning of the week, the homeroom teacher introduced the targeted behaviors of the week and hung a poster on the classroom wall to remind students to participate. Teachers and students could fill out cards noting the names of students who they observed demonstrating the targeted behaviors and what they were doing. At the beginning of each day, the homeroom teacher read the names of both the students who had submitted and been mentioned in cards to recognize their participation and hung the cards on the poster. “Sharing Board” aimed to normalize struggle among students (i.e., everyone struggles, dreams, learns, and can improve). At the beginning of the week, the homeroom teacher introduced the prompt of the week (i.e., a topic for which students could share experiences of struggle) and hung a poster noting the prompt on the wall. At any point throughout the week, the students could complete cards with the writing prompt and hang them on the poster. At the end of the week, the homeroom teacher celebrated students who had participated in the activity.

In addition to the two classroom activities, teachers received feedback tip sheets to provide concrete examples of how they could provide process feedback. Process feedback is when teachers evaluate students' work by connecting an outcome to the process by which students arrived at such outcome. For example, when a teacher says "Your answer is not correct, perhaps, because you did not use the right strategy to solve the problem." Furthermore, teachers received materials to create a change team ("Tim Perubahan") within the school to support teachers in implementing the different activities. The materials included a detailed guide for how to create a change team and organize a training for teachers on how to use the activities. For the training, video tutorials for each aspect of the interventions explaining how to use the materials were provided on a USB drive shipped to the schools. Although the teachers and guidance counselors delivering the intervention did not receive formal training on the materials, detailed lesson plans and instructions were provided, and a video tutorial was shared with the schools through a dedicated website. Moreover, the teachers and guidance counselors could reach out to the project team by phone or e-mail with any questions related to implementation.

Measures: learning strategies. Learning strategies assess the rates with which students have planned the completion of their schoolwork, persevered in the pursue of long-term goals, and sought help when they struggle while learning in the past month. This measure was developed for the larger intervention study to capture the extent to which the curriculum impacts students' learning behaviors. The strategies captured in this measure were inspired by literature on self-regulation and motivation (Credé & Phillips, 2011; Duckworth & Gross, 2014; Duncan & Mckeachie, 2005; Oettingen & Gollwitzer, 2001) and aligned with each self-management lesson in the curriculum (see Table 1).

Learning strategies were assessed with 14 items rated on a Likert-type scale (see Table 1). Students responded to these items by reading the prompt "During the past month, how often have you…" and choosing one of the following response categories: 1 = "I did not do this in the past month”, 2 = “I did this one or twice in the past month”, 3 = “I did this several times in the past month”, 4 = “I did this once a week in the past month”, and 5 = “I did this every day in the past month.” Given these items were created to represent planned behavior as related to the intervention lessons, rather than to form a unifying concept, we categorized these items based on the actions implied by each statement. Five items were categorized as planning strategies, five items as persevering strategies, and four items as help-seeking strategies.

Departures from registered analysis plan. We made three decisions that departed or were not mentioned in the registered analysis plan. First, from the final model in Study 1, we extracted profiles as observed categorical variables. Originally, we planned to extract classification probabilities as weights to implement a multi-step process referred to as the BCH procedure (Bakk, Tekle, & Vermunt, 2013; Nylund-Gibson, Grimm, & Masyn, 2019). The goal of this procedure is to take into account the measurement error associated with profiles (i.e., not every student is perfectly represented by a single profile) when estimating the effects of profiles on distal outcomes. As currently implemented in Mplus 8 version 1.6, the BCH model cannot accommodate a TYPE=TWOLEVEL MIXTURE setting, which is what we needed to fit our registered models. Therefore, we opted for using profiles as observed categorical variables to have more flexibility in our model implementation. This decision implies that the association between profiles and outcomes may have been attenuated (Bray, Lanza, & Tan, 2015), which could would result in more conservative Conditional Average Treatment Effects.

Second, we used Bayesian estimation in Study 2. Originally, we planned to use frequentist estimation and interpret Conditional Average Treatment Effects with a *p*-value below .01 and effect sizes between .10 and .30. We switched to Bayesian estimation because we wanted to have a deeper understanding of uncertainty than the understanding provided by frequentist estimation (Gelman, 2015). Bayesian estimation and inference allowed us to explore how likely small effects were, as opposed to ignoring small and uncertain effects based on a p-value.

Third, we used an ordered-probit model to test for impacts on learning strategies. This decision was not mentioned in the registration and was made to better communicate the changes induced by All Can Succeed on learning strategies. A simpler avenue would have been to fit a linear model. In that case, the results would have been interpreted, for example, as "All Can Succeed increased the use of time planners for growth students by X units." One of the problems with this avenue is that it ignores that changing the use of a strategy (e.g., time planners) from *never* to *once a month* is qualitatively different from increasing its use from *once a week* to *every day*. Therefore, the results of a linear model may not accurately reflect how likely students are to change in different areas of the scale. An ordered-probit model solves this problem by providing the probabilities associated with changes in different areas of the scale.

Models. Figure S5 illustrates the main paths in our model. We used multivariate (i.e., multiple outcomes) multilevel models, where the effects of All Can Succeed on outcome Y_k for student i in school j were moderated by their motivation profile, where paths $\beta_1 - \beta_k$ represent the interactions between motivation profiles and both versions of All Can Succeed. At the student level, we controlled for students' gender, socioeconomic status, and baseline school grades. For learning strategies, we estimated the probability that students would use each strategy (i.e., an ordered-probit model), whereas a linear model was used to estimate national test scores.

Conditional average treatment effects. Our primary focus was on the interactions between motivation profiles (as dummy variables with growth as the reference category) and both versions of All Can Succeed. These interactions revealed whether each profile increased their use of learning strategies and their national test scores after being exposed to the program. We present results as the marginal effects on each outcome by profile or the Conditional Average Treatment Effects (CATE).

In the model for test scores, the CATE is:

$$CATE_p = \mathbb{E}(Y_{pt}|gender, grades, assets) - \mathbb{E}(Y_{pc}|gender, grades, assets)$$

(Eq. 1)

where Y_{pt} is the predicted test score when profile p is exposed to the intervention and Y_{pc} is the predicted test score when profile p is not exposed to the intervention. In the models for learning strategies, the CATE are:

$$CATE_p = Pr(Y_{pt} > 3|gender_k, grades_h, assets = 0, Y_{base} = 3) - Pr(Y_{pc} > 3|gender_k, grades_h, assets = 0, Y_{base} = 3)$$

(Eq. 2)

where Y_{pt} is the predicted probability of using strategy s at least once a week when profile p is exposed to the intervention and Y_{pc} is the predicted probability of using strategy s at least once a week when profile p is not exposed to the intervention.

The CATE for learning strategies are presented as marginal effects at representative values (see Agresti & Tarantola, 2018, for a description of alternative ways to interpret effects of ordered models). That is, we estimated the predicted probabilities of using learning strategies for meaningful values of our predictors. Specifically, we obtained predicted probabilities for male and female students who had an average number of household assets, used strategies several times a month (i.e., category 3 in the scale), and were at the 25th, 50th, and 75th percentiles in the baseline grades distribution. In other words, we estimated the predicted probabilities for six hypothetical students: Three male students (low, median, and high achievers) and three female students (low, median, and high achievers).

For ease of presentation, we included the Conditional Average Treatment Effects for female students with median baseline grades who had average household assets in the manuscript. See below for more details on how the effects varied across different combinations of achievement and gender.

Here, we focused on the intent-to-treat effects of All Can Succeed. That is, we focused on the effects on students' learning strategies and academic achievement when their school was randomized to deliver the treatment, instead of the effects on students who actually participated in the program. As a result, our results below are conservative estimates of how much Indonesian students would benefit if their school was offered the program.

Bayesian estimation. We used Bayesian estimation to test whether motivation profiles boosted or inhibited the effects of All Can Succeed. A key aspect of Bayesian estimation is that the model iterates thousands of times looking for answers to a question (e.g., What is the impact of All Can Succeed on students' academic achievement?). After the final iteration, the model has learned about thousands of potential answers, often referred to as the posterior distribution. In other words, imagine the process a person goes through when they move to a new city and they search for a new favorite coffeeshop. Before starting their search, they have ideas about what type of coffeeshop they like, which they use to visit several coffeeshops on a Saturday. By the end of Saturday, they have an idea of which coffeeshop is their new favorite. On Sunday, while discussing with a friend, this person changes their mind and chooses a different shop. The next day, while talking to their parents, they change their mind yet again. These three choices are potential answers to the question, "Which coffee shop could become my new favorite?"

Similarly, our Bayesian model used prior information and the data to iterate thousands of times looking for an answer. In the end, the model has learned about thousands of potential answers (i.e., the posterior distribution). See McElreath (2020) and Kruschke (2015) for more detailed descriptions of how Bayesian estimation generates a posterior distribution.

Model priors. We used weakly informative priors, $N(0,1)$, for the paths illustrated in Figure S5. In addition, we used weakly informative priors, $IG(.10, .10)$ for random intercepts (i.e., school-specific deviation from the overall outcome mean). Last, to facilitate the convergence of probit models, we used strong priors for the means of imputed baseline learning strategies ($N(2.5, .50)$) and the item thresholds (based on results from a listwise deletion model), as well as starting values for item thresholds.

Interpretation of the posterior distribution. Obtaining a distribution of effects facilitates interpretation of results because, often, researchers are most interested in knowing the probability that an effect is either positive (i.e., did the intervention increase students' grades?) or negative (i.e., did the intervention decrease students' grades?), instead of knowing whether effects meet a significance threshold (i.e., $p < .05$) (Deke & Finucane, 2019). Specifically, once the posterior distributions for Conditional Average Treatment Effects were obtained, we estimated the probability that these effects were positive (i.e., a posterior probability, represented as $Pr(CATE > 0)$). For example, $Pr(CATE > 0) = .87$ means that the probability that an effect is positive is .87 based on our data, priors, and model. This type of inference is different from conclusions drawn from the more common frequentist *p*-values, which are focused on rejecting a null hypothesis (e.g., the intervention had zero effect on students' scores), as opposed to describing how much support there is for an answer. Note that we used 89% Credible Intervals to describe effects because the resulting intervals are more stable in representing uncertainty than the more common 95% intervals (Makowski, Ben-Shachar, & Lüdecke, 2019; McElreath, 2020).

Results. Table S8 presents descriptive statistics for student outcomes, Table S9 presents baseline use rates for each learning strategy, and Table S10 presents the proportion of students in each experimental group by motivation profile.

Conditional Average Treatment Effects on Learning Strategies. Tables S11-13 show the Conditional Average Treatment Effects on learning strategies for all motivation profiles. For disconnected and severely disconnected students, the estimates below are the same as the ones presented in Figures 4 and 5 in the main manuscript.

Effects at other representative values of covariates. As described in the Analysis Plan section, these estimates are marginal effects at representative values (see Agresti & Tarantola, 2018). That is, we estimated the predicted probabilities of using learning strategies for meaningful values of our predictors. Specifically, we obtained predicted probabilities for male and female students who had an average number of household assets, used strategies several times a month (i.e., category 3 in the scale), and were at the 25th, 50th, and 75th percentiles in the baseline grades distribution. In other words, we estimated the predicted probabilities for six hypothetical students: Three male students (low, median, and high achievers) and three female students (low, median, and high achievers). In the manuscript, we presented the Conditional Average Treatment Effects for female students with median baseline grades who had average household assets. In Figure S6 below, we illustrate how the Conditional Average Treatment Effects showed only small differences across levels of baseline achievement and gender. Although Figure S6 illustrates only one strategy, the pattern of results is highly similar across strategies.

Additional Analyses. Below, we describe two sets of supplemental analyses. First, we describe analyses on lower-achieving students following the analyses by Yeager et al. (2019). Second, we describe univariate analyses of learning strategies to explore sources of model misfit.

Impacts among lower-achieving students. Yeager and colleagues (2019) reported a large randomized controlled trial of a direct-to-student growth mindset intervention. A main feature of the study design was that the authors hypothesized that the impacts of the program would be concentrated among lower-achieving students (i.e., at or below the average GPA of their school). Following this expectation, their analytic sample only included lower-achieving students. Among these students, the direct-to-student growth mindset intervention increased students' school grades by .11 standard deviations.

Given that Yeager and colleagues' findings were concentrated among lower-achieving students and lower-achieving students were more likely to show a severely disconnected profile in our sample, it was possible that our main findings were driven by students' baseline achievement, instead of the motivation profile. Table S14 illustrates the association between students' baseline achievement and profile membership. Clearly, having a more adaptive profile (i.e., growth or multiple goals) is associated with a lower probability of being considered a lower-achieving student. Fortunately, the proportion of lower-achieving students within each profile did not vary as much across conditions (see Table S15), which suggests the balance across groups was not threatened by splitting the sample.

To explore the possibility that our main results were driven by students' baseline achievement, we tested models that only included lower-achieving students as defined by Yeager and colleagues. As suggested by Table S16, the pattern of Conditional Average Treatment Effects remained similar to those reported in the manuscript. The main difference is that lower-achieving growth students decreased their math scores. Nonetheless, lower-achieving multiple goals and disconnected students (i.e., the majority of lower-achieving students) did not increase their math or science test scores. That is, had we only relied on students' baseline achievement to answer "Who benefits?", we would have missed the positive impacts shown among severely disconnected and growth students.

For completeness, we also tested models that included only higher-achieving students. Table S17 shows that, again, multiple goals and disconnected students did not increase their test scores due to the program. In addition, higher-achieving growth and disconnected students showed increases in math and science. Taken together, these follow-up analyses suggest that motivation profiles were more likely to answer "Who benefits from All Can Succeed?" than students' baseline achievement was.

Multivariate strategy models did not fit well. An important limitation in our interpretations of impacts on learning strategies is that our models did not fit the data well (see Table S18). In short, model fit evaluation addresses the following question: Based on the model results, can we simulate data that is consistent with the real data? This is called a posterior predictive check and aims to provide information on whether the model makes reasonable predictions (Asparouhov & Muthén, 2010). In our case, the multivariate multilevel models for learning strategies did not fit the data well, consistently making the prediction that "simulated students" used strategies less often than the actual students did.

To explore potential sources of model misfit, we tested univariate models (i.e., a single outcome). We did so because multivariate models have too many components that may influence posterior predictive checks, and thus, decomposing the model into less components can help reveal *where* misfit is coming from. As an illustration, the multivariate model for planning strategies included five strategies and showed poor model fit. However, when the only outcome in the model was “Created a work plan for completing assignments”, the model showed reasonable fit (95% CI for the difference between observed and replicated χ^2 values = [-20.77, 18.59], Posterior Predictive P-Value = .58). Therefore, we could have changed our approach to reporting results only from univariate models. However, we decided to focus on our registered multivariate models, considering that univariate models ignore the possibility that strategies are associated, and thus, may need to be modeled together to obtain more accurate treatment effect estimates. Obviously, more work is needed to understand how these strategies are clustered within individuals and what are the most effective ways of modeling them.

References

Agresti, A., & Tarantola, C. (2018). Simple ways to interpret effects in modeling ordinal categorical data. *Statistica Neerlandica*, *72*(3), 210–223. doi: 10.1111/stan.12130

Asparouhov, T., & Muthén, B. (2010). *Bayesian analysis using Mplus: Technical implementation (version 3)*. Retrieved from https://www.statmodel.com/download/Bayes3.pdf

Bakk, Z., Tekle, F. B., & Vermunt, J. K. (2013). Estimating the Association between Latent Class Membership and External Variables Using Bias-adjusted Three-step Approaches. *Sociological Methodology*, *43*(1), 272–311. https://doi.org/10.1177/0081175012470644

Bank, W. (2019). *Instilling a growth mindset in Indonesia*. World Bank, Washington, D.C.

Blackwell, L., 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

Bray, B. C., Lanza, S. T., & Tan, X. (2015). Eliminating Bias in Classify-Analyze Approaches for Latent Class Analysis. *Structural Equation Modeling*, *22*(1), 1–11. https://doi.org/10.1080/10705511.2014.935265

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, *139*(3), 655–701. https://doi.org/10.1037/a0029531

Clearinghouse, W. W. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation; Regional Assistance.

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, *61*(12), 1250–1262. https://doi.org/10.1016/j.jbusres.2008.01.013

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Deke, J., & Finucane, M. (2019). *Moving Beyond Statistical Significance: The BASIE (BAyeSian Interpretation of Estimates) Framework for Interpreting Findings from Impact Evaluations (OPRE Report 2019 35)*. Office of Planning, Research,; Evaluation, Administration for Children; Families, U.S. Department of Health; Human Services.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*, *38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Dilas, D. B., Mackie, C., Huang, Y., & Trines, S. (2019). *Education in Indonesia*. Retrieved from https://wenr.wes.org/2019/03/education-in-indonesia-2

Duckworth, A. L., & Gross, J. J. (2014). Self-Control and Grit: Related but Separable Determinants of Success. *Current Directions in Psychological Science*, *23*(5), 319–325. https://doi.org/10.1177/0963721414541462

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and Passion for Long-Term Goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166–174. https://doi.org/10.1080/00223890802634290

Duncan, T. G., & Mckeachie, W. J. (2005). The Making of the Motivated Strategies for Learning Questionnaire The Making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologist*, *40*(2), 117–128. https://doi.org/10.1207/s15326985ep4002

Dweck, C. S. (1999). *Self-theories: Their Role in Motivation, Personality and Development*. Philadelphia: Psychology Press.

Dweck, C. S. (2017). From needs to goals and representations: Foundations for a unified theory of motivation, personality, and development. *Psychological Review*, *124*(6), 689–719. https://doi.org/10.1037/rev0000082

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*(1), 5–12. https://doi.org/10.1037/0022-3514.54.1.5

Gelman, A. (2015). The Connection Between Varying Treatment Effects and the Crisis of Unreplicable Research. *Journal of Management*, *41*(2), 632–643. https://doi.org/10.1177/0149206314525208

Grimm, K. J., Mazza, G. L., & Davoudzadeh, P. (2017). Model Selection in Finite Mixture Models: A k-Fold Cross-Validation Approach. *Structural Equation Modeling*, *24*(2), 246–256. https://doi.org/10.1080/10705511.2016.1250638

Hulleman, C. S., Schrager, S. M., Bodmann, S. M., & Harackiewicz, J. M. (2010). A meta-analytic review of achievement goal measures: Different labels for the same constructs or different constructs with similar labels? *Psychological Bulletin*, *136*(3), 422–449. https://doi.org/10.1037/a0018947

Jagacinski, C. M., Kumar, S., & Kokkinou, I. (2008). Challenge seeking: The relationship of achievement goals to choice of task difficulty level in ego-involving and neutral conditions. *Motivation and Emotion*, *32*(4), 310–322. https://doi.org/10.1007/s11031-008-9103-3

Kruschke, J. K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (2nd Edition). Academic Press.

Lee, W., & Kim, S. I. (2014). Effects of achievement goals on challenge seeking and feedback processing: Behavioral and fMRI evidence. *PLoS ONE*, *9*(9). https://doi.org/10.1371/journal.pone.0107254

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). BayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Masyn, K. E. (2017). Measurement Invariance and Differential Item Functioning in Latent Class Analysis With Stepwise Multiple Indicator Multiple Cause Modeling. *Structural Equation Modeling*, *24*(2), 180–197. https://doi.org/10.1080/10705511.2016.1254049

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd Edition). Chapman; Hall/CRC.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., … Midgley, C. (2000). *Manual for the Patterns of Adaptive Learning Scales* (pp. 1–83). Ann Arbor: University of Michigan. Retrieved from University of Michigan website: http://www.umich.edu/$\sim$pals/PALS 2000_V13Word97.pdf

Nylund-Gibson, K., Grimm, R. P., & Masyn, K. E. (2019). Prediction from Latent Classes: A Demonstration of Different Approaches to Include Distal Outcomes in Mixture Models. *Structural Equation Modeling*, *26*(6), 967–985. https://doi.org/10.1080/10705511.2019.1590146

Oettingen, G., & Gollwitzer, P. M. (2001). Goal setting and goal striving. In A. Tesser (Ed.), *Individual processes* (pp. 329–347). Malden, MA: Blackwell Publishing Ltd.

Outes-León, I., Sánchez, A., & Vakis, R. (2020). The Power of Believing You Can Get Smarter: The Impact of a Growth mindset intervention on Academic Achievement in Peru. *World Bank Policy Research Working Paper*, 9141. Washington, D.C.: World Bank.

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

Porter, T., Catalan Molina, D., Blackwell, L., Roberts, S., Quirk, A., Duckworth, A. L., & Trzesniewski, K. (2020). Measuring Mastery Behaviours at Scale: The Persistence, Effort, Resilience, and Challenge- Seeking (PERC) Task. *Journal of Learning Analytics*, *7*(1), 5–18. https://doi.org/10.18608/jla.2020.71.2

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis. A festschrift for heinz neudecker* (pp. 233–247). London: Kluwer Academic Publishers.

Sommet, N., & Elliot, A. J. (2017). Achievement Goals, Reasons for Goal Pursuit, and Achievement Goal Complexes as Predictors of Beneficial Outcomes: Is the Influence of

Goals Reducible to Reasons? *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000199

Weissberg, R. P., Durlak, J. A., Domitrovich, C. E., & Gullota, T. P. (2015). Social and emotional learning: Past, present, and future. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullota (Eds.), *Handbook for social and emotional learning: Research and practice.* New York, NY: Guilford Press.

World Bank (2019). Instilling a Growth Mindset in Indonesia. *eMBeD brief*. Washington, DC: World Bank. Retrieved from https://bit.ly/3bvscEA

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*, 364–369. https://doi.org/10.1038/s41586-019-1466-y

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but Important: A Self-Transcendent Purpose for Learning Fosters Academic Self-Regulation. *Journal of Personality and Social Psychology*, *107*(4), 559–580. https://doi.org/10.1037/a0037637.supp

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., … 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/10.1007/s13398-014-0173-7.2

Supplemental Tables

Table S1

Descriptive Statistics for Profile Indicators and Covariates

Measure	*M*	*SD*	*Min*	*Max*	*ICC*
Mindset	0.00	1.00	-2.76	2.20	0.15
Effort	0.00	1.00	-3.91	1.75	0.14
Grit	0.00	1.00	-4.28	2.17	0.12
Challenge-Seeking	0.00	1.00	-1.64	1.87	0.13
Learning Goals	0.00	1.00	-5.83	1.13	0.11
Performance-Avoidance Goals	0.00	1.00	-2.57	1.93	0.12
Mastery Behavior	0.00	1.00	-1.95	2.68	0.46
School Grades	0.00	1.00	-11.12	3.67	0.61
Household Assets	0.00	1.60	-4.46	2.89	0.37
Female	0.54	0.50	0.00	1.00	0.01

Note. ICC = Intraclass correlation obtained from a null multilevel model (schools as clusters). All variables, except for Female, were standardized.

Table S2

Associations Between Profile Indicators and Covariates

Variable	1	2	3	4	5	6	7	8	9
1. Mindset									
2. Effort	0.00								
3. Grit	0.07*	0.44*							
4. Challenge-Seeking	0.16*	0.13*	0.13*						
5. Learning Goals	0.02*	0.38*	0.35*	0.10*					
6. Performance-Avoidance Goals	-0.17*	0.16*	0.13*	-0.08*	0.26*				
7. Mastery Behavior	0.12*	0.04*	0.07*	0.17*	0.04*	-0.01			
8. School Grades	0.18*	0.10*	0.12*	0.17*	0.12*	-0.05*	0.11*		
9. Household Assets	0.16*	-0.03*	0.06*	0.08*	0.01*	-0.01	0.11*	0.16*	
10. Female	0.04*	0.11*	0.15*	0.02*	0.18*	-0.03*	-0.06*	0.22*	-0.02*

Note. *$p < .01$.

Table S3

Summary of Conclusions on Profile Non-invariance Across Gender, Socioeconomic Status, and Academic Achievement

Covariate	Non-invariance	Scales
Household Assets	Uniform	PERC
Female	Uniform	PERC, effort, grit
School Grades: Math, Science, and English	Nonuniform	PERC, challenge, mindset
School Grades: Indonesian	Nonuniform	PERC, challenge, mindset

Note. Uniform = Covariate X has same association with scale Z across profiles, above and beyond its association with profile membership. Nonuniform = Covariate X has differential associations with scale Z across profiles, above and beyond its association with profile membership.

Table S4

Step 1 in Models to Detect Profile Non-invariance

Covariate	*cd*	*lrts*	*df*	*p*
Assets	1.36	1,443.07	28	< 0.01
Bahasa	1.45	1,110.92	28	< 0.01
Girl	1.35	867.07	28	< 0.01
Math, Science, English	1.79	756.12	28	< 0.01

Note. *cd* = Difference test scaling correction, *lrts* = Likelihood Ratio Test. Step 1 compares an invariant model to a model in which all scales show nonuniform non-invariance.

Table S5

Step 2 in Models to Detect Profile Non-invariance

Scale	Assets	Bahasa	Girl	Math, Science, English
Avoidance Goals	2.99	104.27*	61.24*	99.23*
Challenge-seeking	227.77*	401.63*	245.87*	390.75*
Effort	48.34*	82.22*	1244.91*	88.78*
Grit	126.30*	209.05*	645.10*	186.73*
Learning Goals	7.93	142.73*	35.25*	95.72*
Mastery Behavior	81.54*	37.77*	418.43*	50.26*
Mindset	1135.01*	486.89*	87.52*	385.35*

Note. Numbers represent the likelihood ratio test (with 4 degrees of freedom) and stars represent significant ($p < .01$) estimates. Step 2 compares an invariant model on a specific scale to a model with nonuniform non-invariance on the same scale.

Table S6

Step 3 in Models to Detect Profile Non-invariance

Covariate	M1.0 vs M3.0	M1.1 vs M3.0
Assets	53.02* (4)	1355.92* (24)
Bahasa	877.45* (12)	397.13* (16)
Girl	447.83* (12)	429.93* (16)
Math, Science, English	464.63* (12)	323.15* (16)

Note. Numbers represent the likelihood ratio test (degrees of freedom in parentheses) and stars represent significant ($p < .01$) model differences. Step 3 compares three models: an invariant model (M1.0), an all nonuniform non-invariance model (M1.1), and a model that includes nonuniform non-invariance for some scales based on Step 2 (M3.0)).

Table S7

Step 4 in Models to Detect Profile Non-invariance

Covariate	M3.0 vs M4.0	M3.0 vs M4.1	M3.0 vs M4.2
Assets	10.57 (3), $p = 0.01$		
Bahasa	66.36 (6), $p < 0.01$	52.98 (6), $p < 0.01$	24.84 (6), $p < 0.01$
Girl	7.60 (6), $p = 0.27$	4.93 (6), $p = 0.55$	5.09 (6), $p = 0.53$
Math, Science, English	116.30 (6), $p < 0.01$	75.28 (6), $p < 0.01$	80.11 (6), $p < 0.01$

Note. Likelihood ratio test (degrees of freedom in parentheses) and corresponding *p*-values. Step 4 compares a uniform non-invariance model (M4.0) to a model that includes nonuniform non-invariance for some scales based on Step 2 (M3.0). For assets, M3.0 included only one non-invariant path, and thus, there was only one possible comparison. For Bahasa, M4.0 included nonuniform non-invariance for mastery behavior, M4.1 included nonuniform non-invariance for challenge-seeking, and M4.2 included nonuniform non-invariance for growth mindset. For female, M4.0 included nonuniform non-invariance for mastery behavior, M4.1 included nonuniform non-invariance for effort beliefs, and M4.2 included nonuniform non-invariance for grit. For Math, Science, and English, M4.0 included nonuniform non-invariance for mastery behavior, M4.1 included nonuniform non-invariance for challenge-seeking, and M4.2 included nonuniform non-invariance for growth mindset.

Table S8

Descriptive Statistics for Student Outcomes

Measure	*M*	*SD*	*Min*	*Max*	*Missing*	*ICC*
Practice (6)	3.27	1.03	1.00	5.00	0.00	0.10
Reward myself (4)	2.60	1.28	1.00	5.00	0.00	0.10
New habits (4)	3.57	1.11	1.00	5.00	0.00	0.08
Friend check-ins (5)	2.93	1.24	1.00	5.00	0.00	0.08
Time planner (5)	2.98	1.32	1.00	5.00	0.00	0.09
Stick to priorities (5)	3.06	1.33	1.00	5.00	0.00	0.10
School is priority (5)	3.52	1.18	1.00	5.00	0.00	0.09
Work plan (3)	3.46	1.12	1.00	5.00	0.00	0.10
Imagine goals (3)	3.83	1.11	1.00	5.00	0.00	0.09
Deep breaths (4)	3.79	1.25	1.00	5.00	0.00	0.07
Help: Teacher (6)	3.60	1.16	1.00	5.00	0.00	0.11
Help: Parent (6)	3.14	1.30	1.00	5.00	0.00	0.09
Help: Peers (6)	3.55	1.13	1.00	5.00	0.00	0.07
Ask questions (6)	3.45	1.24	1.00	5.00	0.00	0.12
English Test	-0.01	0.99	-3.22	5.11	0.18	0.54
Indonesian Test	-0.01	0.99	-4.41	2.08	0.18	0.47
Math Test	-0.02	1.00	-2.70	2.91	0.18	0.62
Science Test	-0.01	1.00	-3.16	3.17	0.18	0.59

Note. Missing = Proportion of missing data. ICC = Intraclass correlation obtained from a null multilevel model (schools as clusters). Test scores were standardized. The numbers in parenthesis next to each learning strategy represents the lesson in which it was promoted: (3) =

“Set your goals”, (4) = “Build good habits”, (5) = “Deal with distractions”, and (6) = “Learn from failure.”

Table S9

Baseline Rates of Use of Learning Strategies in the Full Sample

Measure	*Pr(Weekly)*	*M*	*SD*	*Min*	*Max*
Practice (6)	0.43	3.27	1.10	1	5
Reward myself (4)	0.26	2.51	1.34	1	5
New habits (4)	0.54	3.54	1.19	1	5
Friend check-ins (5)	0.37	2.94	1.30	1	5
Time planner (5)	0.37	2.91	1.40	1	5
Stick to priorities (5)	0.42	3.08	1.42	1	5
School is priority (5)	0.55	3.55	1.25	1	5
Work plan (3)	0.52	3.47	1.20	1	5
Imagine goals (3)	0.63	3.80	1.19	1	5
Deep breaths (4)	0.61	3.70	1.35	1	5
Help: Teacher (6)	0.58	3.67	1.24	1	5
Help: Parent (6)	0.46	3.24	1.35	1	5
Help: Peers (6)	0.53	3.54	1.21	1	5
Ask questions (6)	0.50	3.43	1.32	1	5

Note. Pr(Weekly) = Average probability that students used a specific strategy at least once a week. The numbers in parenthesis next to each learning strategy represents the lesson in which it was promoted: (3) = “Set your goals”, (4) = “Build good habits”, (5) = “Deal with distractions”, and (6) = “Learn from failure.”

Table S10

Balance of Motivation Profiles Across Experimental Conditions

Profile	Condition	*n*	% of Profile
Disconnected (*n* = 13,696)	ACS	4,587	0.33
	ACS+	4,567	0.33
	Control	4,542	0.33
Growth (*n* = 6,254)	ACS	2,077	0.33
	ACS+	2,058	0.33
	Control	2,119	0.34
Multiple Goals (*n* = 27,635)	ACS	9,278	0.34
	ACS+	9,102	0.33
	Control	9,255	0.33
Severely Disconnected (*n* = 1,934)	ACS	648	0.34
	ACS+	596	0.31
	Control	690	0.36

Note. % Profile = Percent of students within a motivation profile.

Table S11

Summary of Conditional Average Treatment Effects (Percent Change) on Planning Strategies by Profile (Mean and 89% CI)

Strategy	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
Planned to do something you like as a reward (not necessarily an object) for getting something done	ACS	3.98 [2.08, 5.88]	1.41 [0.16, 2.69]	2.03 [0.55, 3.52]	2.59 [-0.41, 5.55]
	ACS+	2.01 [0.12, 3.87]	1.91 [0.67, 3.18]	1.54 [0.05, 3.03]	2.71 [-0.35, 5.78]
Made a plan to form new habits	ACS	1.39 [-1.03, 3.82]	-0.20 [-1.66, 1.28]	0.90 [-0.90, 2.72]	2.84 [-0.97, 6.70]
	ACS+	1.95 [-0.53, 4.39]	1.72 [0.24, 3.18]	0.10 [-1.70, 1.88]	2.30 [-1.62, 6.19]
Made a priority list (time planner)	ACS	5.00 [2.68, 7.32]	2.40 [0.94, 3.88]	3.44 [1.71, 5.19]	0.72 [-2.78, 4.24]
	ACS+	4.83 [2.51, 7.19]	3.92 [2.44, 5.42]	3.83 [2.10, 5.58]	6.39 [2.73, 10.05]
Decided to prioritize school work over playing	ACS	1.25 [-1.12, 3.63]	0.07 [-1.44, 1.59]	1.22 [-0.62, 3.05]	-0.78 [-4.67, 3.15]
	ACS+	4.06 [1.68, 6.42]	1.15 [-0.35, 2.68]	2.42 [0.58, 4.27]	-0.77 [-4.66, 3.10]
Created a work plan for completing assignments	ACS	2.19 [-0.28, 4.67]	0.01 [-1.56, 1.53]	2.08 [0.22, 3.97]	0.95 [-2.96, 4.88]

Strategy	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
	ACS+	3.75 [1.30, 6.21]	1.71 [0.15, 3.26]	1.76 [-0.12, 3.61]	2.36 [-1.60, 6.30]

Note. Conditional Average Treatment Effects for female students at the 50th percentile of the baseline grades distribution who had an average number of household assets. Effects are presented as percentage change, instead of probabilities.

Table S12

Summary of Conditional Average Treatment Effects (Percent Change) on Persevering Strategies by Profile (Mean and 89% CI)

Strategy	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
Practiced on difficult problems	ACS	-1.48	0.20	0.15	3.21
		[-4.06, 1.05]	[-1.39, 1.75]	[-1.71, 1.95]	[-0.50, 6.91]
	ACS+	1.02	0.33	-0.43	0.95
		[-1.55, 3.56]	[-1.26, 1.91]	[-2.27, 1.45]	[-2.83, 4.74]
Checked in with a friend to help you stick to your plans	ACS	0.72	0.58	0.14	-4.53
		[-1.4, 2.89]	[-0.77, 1.91]	[-1.49, 1.76]	[-8.12, -0.93]
	ACS+	2.16	2.01	-0.6	-2.94
		[0.02, 4.35]	[0.63, 3.36]	[-2.21, 1.01]	[-6.59, 0.69]
Stuck to your priority list	ACS	4.43	1.07	2.74	3.69
		[2.03, 6.79]	[-0.42, 2.57]	[0.99, 4.52]	[0.03, 7.30]
	ACS+	4.36	2.73	2.97	4.09
		[1.94, 6.70]	[1.21, 4.26]	[1.21, 4.73]	[0.47, 7.75]
Imagined achieving a long-term goal to help you stay motivated and focused on school work	ACS	-1.06	0.03	-0.56	0.38
		[-3.14, 1.06]	[-1.35, 1.38]	[-2.28, 1.13]	[-3.46, 4.30]
	ACS+	-0.13	1.05	-0.49	-0.41
		[-2.24, 2.01]	[-0.34, 2.42]	[-2.22, 1.25]	[-4.28, 3.57]
	ACS	-0.03	0.11	-1.24	-0.44
		[-2.13, 2.09]	[-1.24, 1.44]	[-2.91, 0.45]	[-4.3, 3.38]

Strategy	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
Took some deep breaths to calm down when stressed	ACS+	0.44 [-1.66, 2.55]	1.08 [-0.24, 2.40]	-1.18 [-2.85, 0.48]	1.78 [-2.13, 5.71]

Note. Conditional Average Treatment Effects for female students at the 50th percentile of the baseline grades distribution who had an average number of household assets. Effects are presented as percentage change, instead of probabilities.

Table S13

Summary of Conditional Average Treatment Effects (Percent Change) on Help-seeking Strategies by Profile (Mean and 89% CI)

Strategy	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
Asked your teacher for help on something you don't understand	ACS	-0.07 [-2.47, 2.37]	-1.30 [-2.92, 0.29]	-0.64 [-2.58, 1.28]	-0.73 [-4.66, 3.20]
	ACS+	1.15 [-1.21, 3.62]	-0.60 [-2.18, 1.00]	0.50 [-1.41, 2.42]	0.19 [-3.76, 4.18]
Asked your parent or guardian for help on something you don't understand	ACS	-1.00 [-3.26, 1.30]	-0.93 [-2.39, 0.54]	0.97 [-0.79, 2.71]	0.98 [-2.73, 4.74]
	ACS+	1.44 [-0.91, 3.76]	0.27 [-1.20, 1.73]	1.04 [-0.73, 2.80]	3.72 [-0.11, 7.54]
Asked your classmates or friends for help on something you don't understand	ACS	1.46 [-0.76, 3.73]	-1.54 [-2.89, -0.17]	-0.76 [-2.45, 0.93]	0.57 [-3.30, 4.44]
	ACS+	2.43 [0.25, 4.62]	-0.50 [-1.85, 0.84]	-0.23 [-1.96, 1.47]	-1.90 [-5.84, 2.03]
Asked questions during class to the teacher	ACS	-1.09 [-3.51, 1.44]	-0.88 [-2.53, 0.75]	0.56 [-1.38, 2.53]	-1.14 [-5.05, 2.82]
	ACS+	-0.91 [-3.32, 1.55]	0.81 [-0.85, 2.45]	0.52 [-1.41, 2.50]	-1.85 [-5.85, 2.09]

Note. Conditional Average Treatment Effects for female students at the 50th percentile of the baseline grades distribution who had an average number of household assets. Effects are presented as percentage change, instead of probabilities.

Table S14

Percent of Lower-Achieving Students by Profile

Motivation Profile	*n*	% Lower-Achieving
Severely Disconnected	1,934	73.42
Disconnected	13,696	62.05
Multiple Goals	27,635	52.40
Growth	6,254	36.84

Note. Students were classified as lower-achieving if they were at or below their school average baseline grades following Yeager et al. (2019).

Table S15

Percent of Lower-Achieving Students by Profile and Treatment Assignment

Motivation Profile	Condition	*n*	% Lower-Achieving
Severely Disconnected	ACS	648	73.61
	ACS+	596	74.83
	Control	690	72.03
Disconnected	ACS	4,587	62.46
	ACS+	4,567	62.05
	Control	4,542	61.65
Multiple Goals	ACS	9,278	52.39
	ACS+	9,102	51.65
	Control	9,255	53.15
Growth	ACS	2,077	37.27
	ACS+	2,058	39.46
	Control	2,119	33.88

Note. Students were classified as lower-achieving if they were at or below their school average baseline grades following Yeager et al. (2019).

Table S16

Conditional Average Treatment Effects Among Lower-Achieving Students by Profile

Subject	Intervention	Growth	Multiple Goals	Disconnected	Severely Disconnected
Science	ACS	0.03	0.02	-0.01	0.16
		[-0.05, 0.11]	[-0.05, 0.09]	[-0.08 , 0.06]	[0.07 , 0.26]
	ACS+	0.00	-0.01	0.01	0.04
		[-0.09, 0.08]	[-0.08, 0.06]	[-0.06, 0.08]	[-0.05 , 0.14]
Math	ACS	-0.07	0.01	-0.03	0.05
		[-0.15, 0.01]	[-0.06 , 0.08]	[-0.1, 0.04]	[-0.05, 0.14]
	ACS+	-0.08	0.00	-0.01	-0.01
		[-0.17, 0.00]	[-0.07 , 0.07]	[-0.08, 0.06]	[-0.11, 0.09]

Note. Students were classified as lower-achieving if they were at or below their school average baseline grades following Yeager et al. (2019).

Table S17

Conditional Average Treatment Effects Among Higher-Achieving Students by Profile

Subject	Treatment	Growth	Multiple Goals	Disconnected	Severely Disconnected
Science	ACS	0.01 [-0.07, 0.09]	0.02 [-0.06, 0.09]	-0.01 [-0.09, 0.06]	0.08 [-0.05, 0.22]
	ACS+	0.08 [0.00, 0.16]	0.01 [-0.06, 0.08]	0.01 [-0.07, 0.08]	0.13 [-0.01, 0.27]
Math	ACS	0.00 [-0.08, 0.09]	0.00 [-0.08, 0.07]	0.01 [-0.07, 0.09]	0.17 [0.03, 0.31]
	ACS+	0.07 [-0.01, 0.16]	0.01 [-0.07, 0.08]	0.01 [-0.07, 0.09]	0.10 [-0.03, 0.24]

Note. Students were classified as higher-achieving if they were above their school average baseline grades following Yeager et al. (2019).

Table S18

Model Fit of Multivariate Models of Learning Strategies

Model	95% CI χ^2 Difference
Help-seeking	[4239.941, 4383.098]
Persevering	[6951.099, 7088.690]
Planning	[7518.652, 7678.350]

Note. In each iteration of the model estimation, the difference between observed and replicated χ^2 values is calculated. The 95% CI of such difference is used to evaluate whether the range of potential differences is consistently on one side of zero or not.

Supplemental Figures

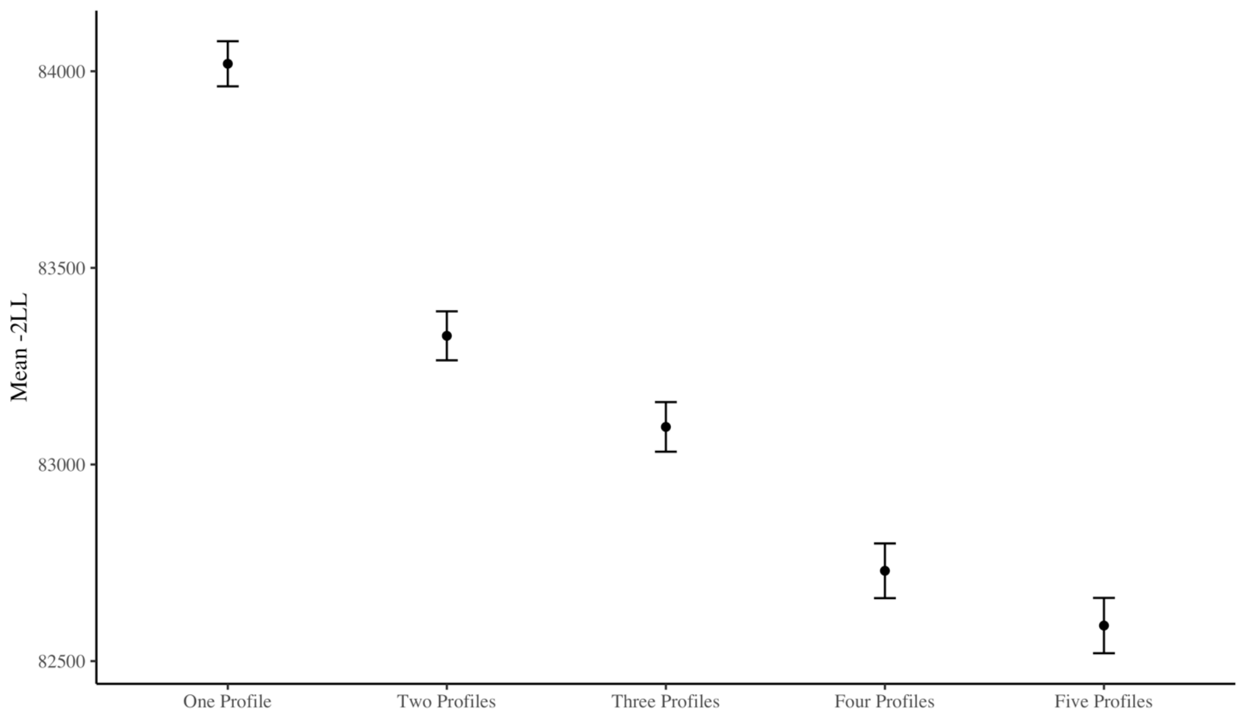

Figure S1. Average model fit values ($-2 * LogLikelihood$) across validation samples for each profile model. Smaller values indicate better model fit. Dots represent the average loglikelihood across validation samples, while error bars represent 1 standard error above and below the average.

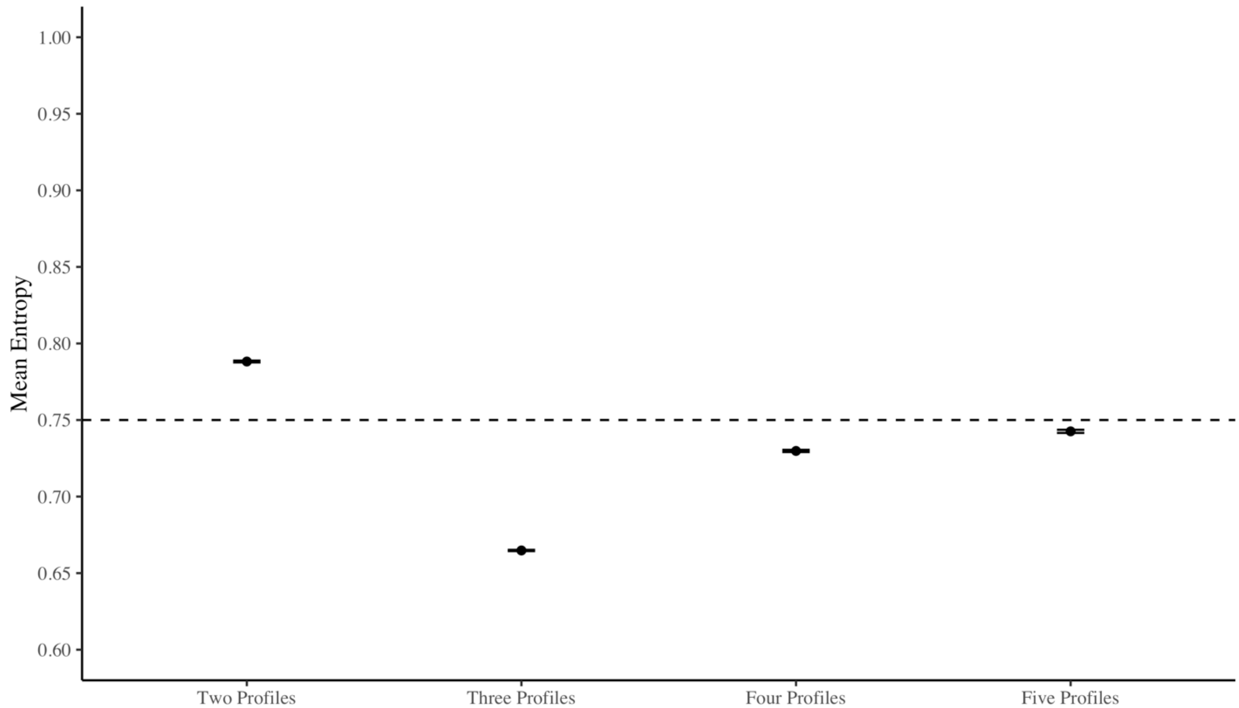

Figure S2. Average entropy values across validation samples for each profile model. Higher values indicate a model is able to more accurately classify individuals to a specific profile. Dots represent the average loglikelihood across validation samples, while error bars represent 1 standard error above and below the average.

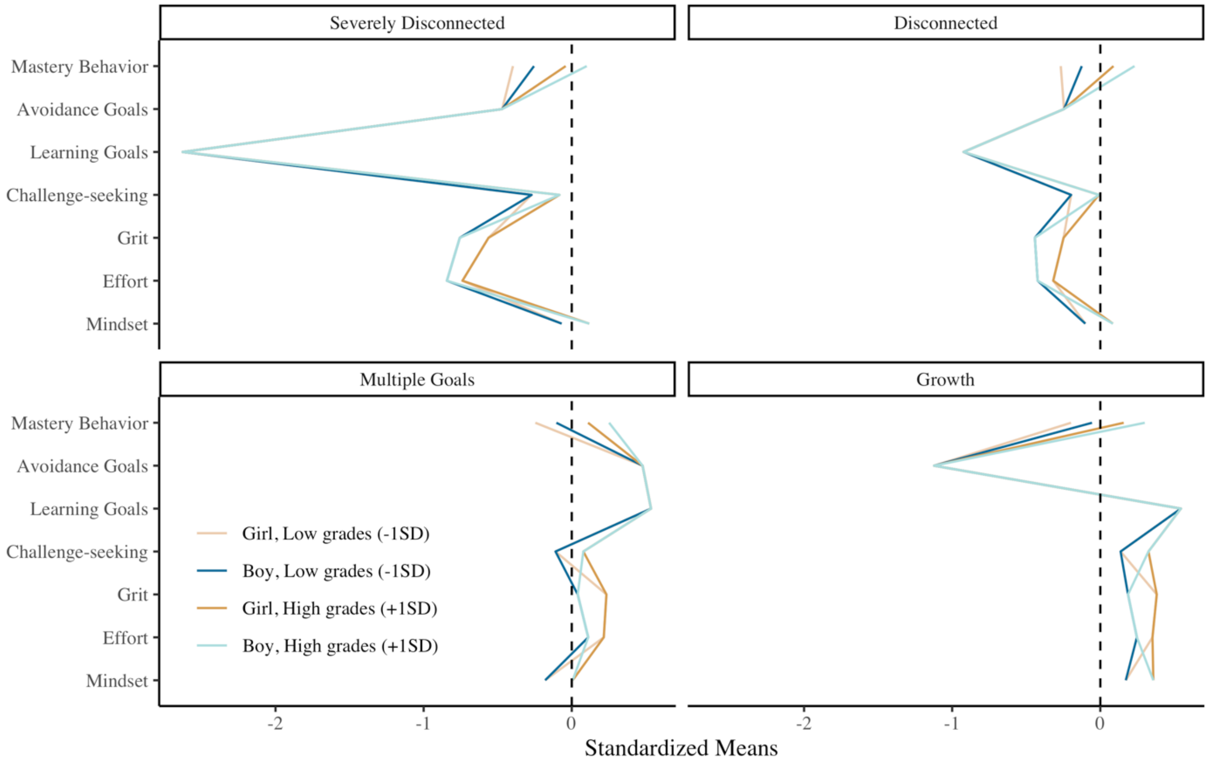

Figure S3. To test for invariance of profiles, we estimated the means for each profile when covariates are included in the model. Results showed some differences in average levels across gender and academic achievement (not across socioeconomic status), but the overall shapes of the profiles remained similar.

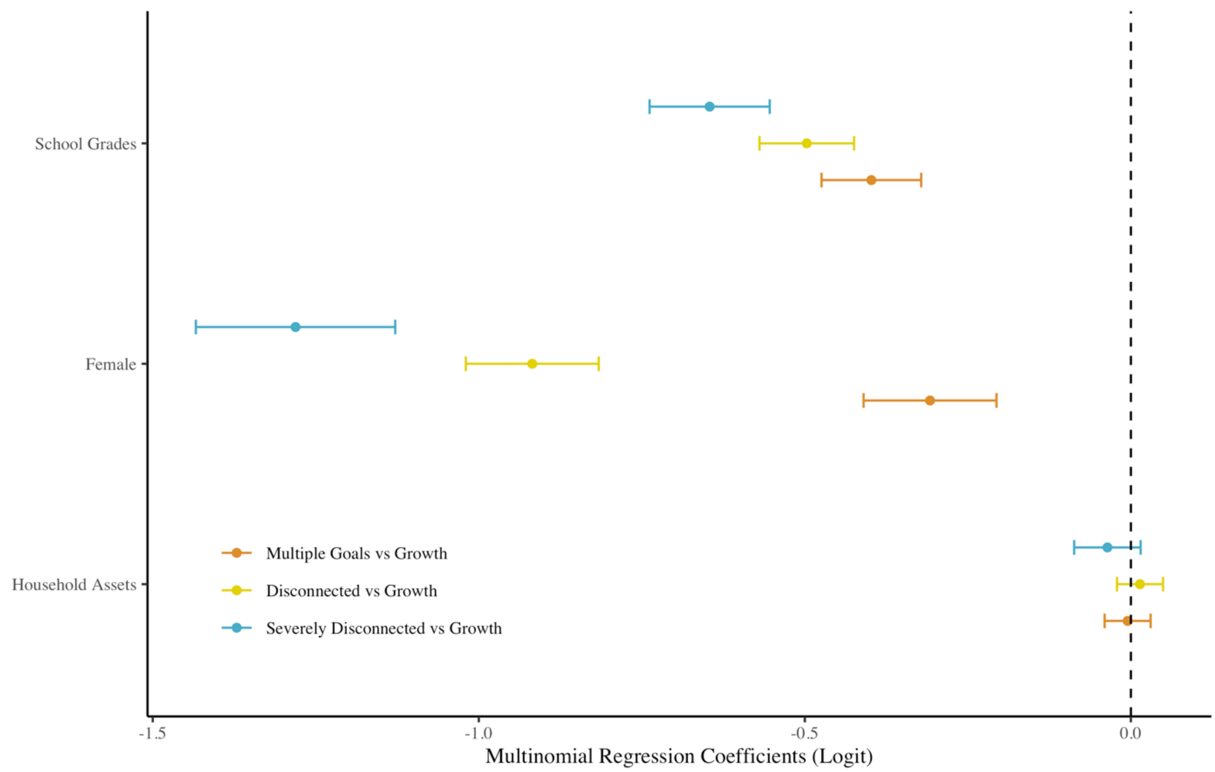

Figure S4. Association between covariates and profile membership in logit scale. Error bars represent 95% Confidence Intervals. The growth profile was the reference category in the multinomial regression, and thus, each coefficient represents how likely a student was to show a specific profile compared to showing a growth profile when their background changed by one unit. School grades and household assets were standardized.

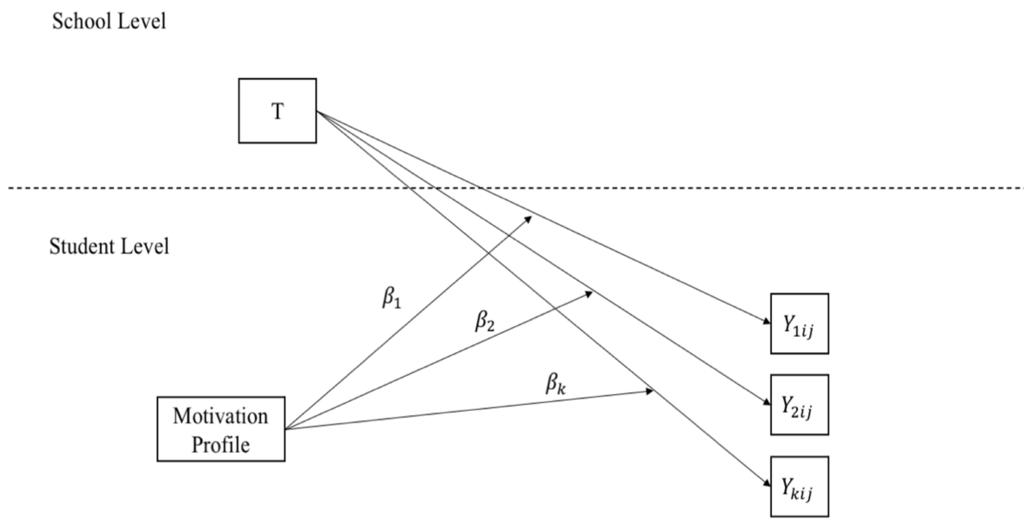

Figure S5. Main paths that define the Conditional Average Treatment Effects. Paths $\beta_1 - \beta_k$ represent the interactions between motivation profiles and both versions of All Can Succeed. $Y_{1ij} - Y_{kij}$ represent multiple outcomes for student i in school j. At the student level, we controlled for students' gender, assets, and baseline school grades. For learning strategies, we estimated the probability that students would use each strategy (i.e., an ordered-probit model), whereas a linear model was used to estimate national exam scores.

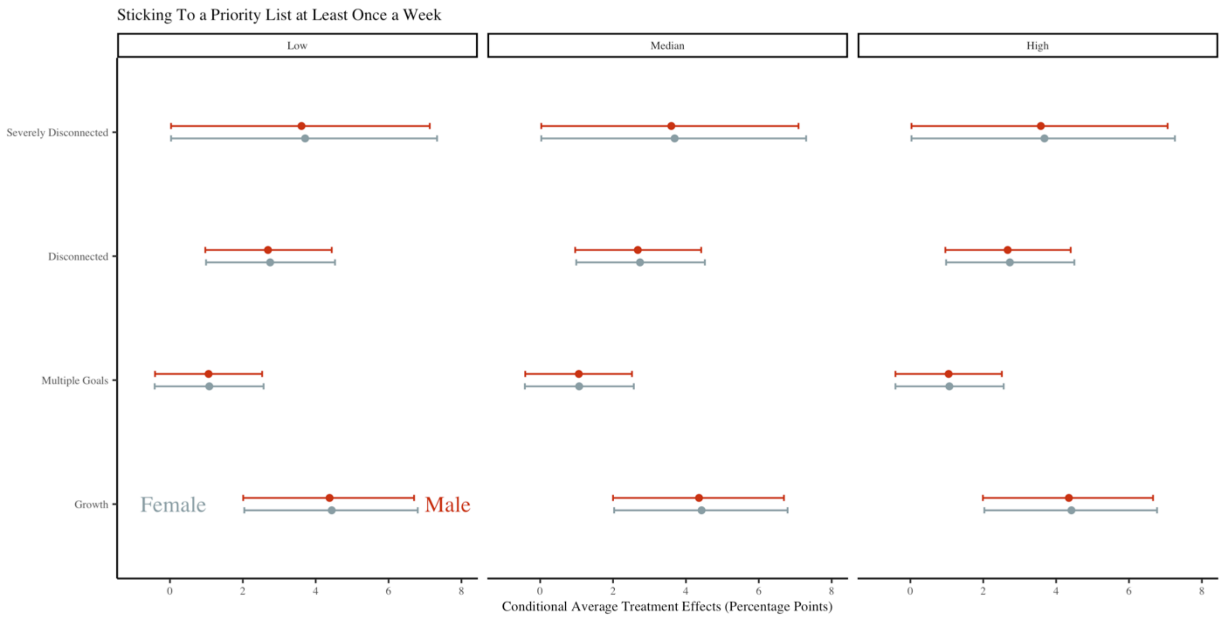

Figure S6. Illustration of variation in impact estimates across levels of baseline achievement (columns) and gender (colors). Dots represent posterior means and error bars represent 89% Confidence Intervals. The estimates in this figure correspond to treatment effects when schools were assigned to ACS. For disconnected and severely disconnected female students at median baseline achievement, the estimates above are the same as the ones presented in Figures 4 and 5 in the main manuscript.