Educational Evaluation and Policy Analysis
Month 201X, Vol. XX, No. X, pp. 1–22
DOI: 10.3102/0162373720938041
Article reuse guidelines: sagepub.com/journals-permissions

Growth-Mindset Interventions at Scale: Experimental Evidence From Argentina

Alejandro J. Ganimian
New York University

This is one of the first evaluations of a "growth-mindset" intervention at scale in a developing country. I randomly assigned 202 public secondary schools in Salta, Argentina, to a treatment group in which Grade 12 students were asked to read about the malleability of intelligence, write a letter to a classmate, and post their letters in their classroom, or to a control group. The intervention was implemented as intended. Yet, I find no evidence that it affected students' propensity to find tasks less intimidating, school climate, school performance, achievement, or post-secondary plans. I rule out small effects and find little evidence of heterogeneity. This study suggests that the intervention may be more challenging to replicate and scale than anticipated.

Keywords: *evaluation, international education/studies, self-concept, social processes/development, student behavior/attitude, regression analyses, experimental research, policy analysis*

There is mounting evidence indicating that the expectations of children and youths and their parents about the payoff from schooling influence their educational investments. These beliefs affect whether (and for how long) children go to school, the type of schools that they choose (e.g., public or private and academic or vocational), whether they invest in complements to schooling (e.g., tuition), and how much effort they exert in school (see Banerjee et al., 2013).

Experiments in developing countries have found that low-income families often hold beliefs that lead them to underinvest in schooling, but that they adjust their behavior when provided with information. Most studies have explored the effects of information on returns to schooling (Avitabile & de Hoyos, 2014; Berniell, 2014; Bonilla et al., 2016; Nguyen, 2009) and school quality (Andrabi et al., 2017; Camargo et al., 2018; Loyalka et al., 2013).

Providing information on child ability may have a larger effect on human capital accumulation by affecting not only demand for schooling, but potentially also student motivation and effort. It is also more likely to impact equity by correcting parental biases (e.g., about boys and girls). The few studies that provided this type of information have used objective measures of ability (i.e., test scores; see Barrera-Osorio et al., 2020; Bobba & Frisancho, 2016; Dizon-Ross, 2019). The main advantage of this approach is that it conveys individual-level information, making the intensity of the treatment inversely proportional to the gap between expected and actual child ability, increasing its chances of affecting those who need it most. Its main drawback, however, is that it does not account for the fact that these ability measures are partly a function of past educational investments (which may themselves be based on incorrect beliefs), so that the information could reinforce the inefficient and inequitable investments it seeks to address.

An alternative is to inform children of their *potential*—rather than their current—ability. A team of psychologists in the United States has designed an intervention with this objective. It asks students to read a short passage that synthesizes research showing that exposure to stimulating environments and practice at challenging tasks can help develop one's intelligence, much like setting ambitious exercise goals and working out at the gym can grow one's muscles. The reading is followed by a brief exercise so that students can internalize this main message. The intervention is based on a large body of research indicating that individuals' beliefs about whether intelligence is fixed or malleable influence their effort, and in turn, their performance (for reviews of this literature, see Dweck & Leggett, 1988; Dweck et al., 2014; Dweck & Yeager, 2019). Variations of this "growth-mindset" intervention have improved self-beliefs (Aronson et al., 2002), school performance (Good et al., 2003; Paunesku et al., 2015), achievement (Blackwell et al., 2007), health (Yeager et al., 2014), and peer relations (Yeager et al., 2011, 2013).

This article presents one of the first studies of this intervention at scale in a developing country. I randomly assigned 202 public secondary schools in the Province of Salta, Argentina, to a "treatment" group, in which representatives from the ministry of education visited schools, invited Grade 12 students to read the passage described above, write a letter to a classmate about how to apply its lessons to their own lives, and put up their letters next to a poster on one of the classroom walls, or to a "control" group that did not implement the intervention. The intervention was conducted during a non-academic period in which students discuss school-related matters with their teacher, so this study assesses whether using this time for this activity had an effect on students' beliefs, effort, performance in school, and achievement. I can verify that the intervention was implemented as intended in 90% of the treatment group using either pictures taken by implementers (83%) or confirmations from principals (7%).

I report five sets of results. First, I find no evidence that the intervention led students to find challenging tasks less intimidating. I show that the intervention had a precisely estimated null effect on students' perceptions of the difficulty of schoolwork, their self-efficacy, and the usefulness of classroom tests. In fact, I find that it may have had a negative effect on female students (increasing their perception of schoolwork as difficult and decreasing their self-efficacy), students from low-income families (decreasing their self-efficacy), and those who had repeated a grade (lowering their propensity to see classroom assessments as useful).

Second, I find no evidence that the intervention increased student effort in school-related tasks (e.g., going to school, attending private tuition), personal development (e.g., reading books, learning languages, playing sports), or existing obligations (i.e., work at or outside the home). I can even rule out small effects in all of these outcomes and find no heterogeneous effects.

Third, I find no evidence that the intervention improved school climate, including relationships between peers, bullying, or student vandalism (i.e., stealing and damaging of school property). In fact, some evidence suggests that it might have had a negative effect on female students (decreasing their propensity to get along with peers).

Fourth, consistent with these null results, I find that the intervention had no impact on students' school performance (e.g., passing, repetition, and dropout rates), their achievement in the national assessment of math and language, or plans to pursue post-secondary education. I even find some indication that it had a negative effect on students' aspirations in schools with higher levels of achievement, resources, and supports for low-performing students.

Finally, I find that the intervention is relatively inexpensive at a cost of USD$2.82 per student, but that it is considerably costlier than suggested by prior studies in developing countries. The main reason for the discrepancy stems from including the cost of training implementers, which had been avoided in a prior study by directly shipping intervention packages to schools. Most school systems are unlikely to deliver an intervention without training for implementers, so my figures seem to be more representative of the actual cost of this intervention at scale.

This study makes several key contributions to research on the growth-mindset intervention. To put these contributions in context, I conducted a detailed review of prior randomized evaluations of this intervention in both developed and developing countries (see Appendix B in the online version of the journal).

My review indicates that this is the first study that can rule out small positive effects from the intervention on mechanisms (e.g., student beliefs and effort) and outcomes (e.g., achievement). Several impact evaluations had previously found that the intervention had null or mixed effects (see, for example, Burnette et al., 2018; Dommett et al., 2013; Gandhi et al., 2019; Sriram, 2014). Yet, none of them was designed to distinguish between precisely estimated null effects and statistically insignificant but imprecise results. This is a major contribution of the present study because it demonstrates that the intervention does not always have the effects seen in efficacy trials, even if its materials are standardized and it can be implemented in one brief session with relatively little adult supervision.

According to my review, this is also one of the first evaluations of the intervention at scale. Until recently, it had been assessed through efficacy trials with small convenience samples. This approach has been instrumental in ensuring the intervention was implemented faithfully, establishing its proof of concept, and carefully measuring its potential mechanisms of impact, but it has been less helpful in understanding its effectiveness at scale within the school system. In recent years, there have been two large-scale randomized evaluations of this intervention. Outes et al. (2020) evaluated it in 800 public secondary schools in three regions of Peru and Yeager et al. (2019) in 65 public secondary schools across the United States. The differences in the context, implementation, and measurement between these studies raise useful questions about the intervention that can inform future research and policy decisions.

Finally, my review indicates that the sampling, randomization, and data collection strategies in this study are uniquely positioned to assess the effectiveness of this intervention at scale. First, its sample included nearly all secondary public schools of a (sub-national) school system. This approach circumvents the problems of site selection bias present in most prior studies and allows me to understand the effect of the intervention where it is not necessarily welcomed. Second, its school-level randomization avoids the spillovers of student-level randomization and allows me to estimate the impact of the intervention when it is conducted by an entire school. Third, its reliance on administrative data collected by the school system (mainly, through the annual census of schools and national student assessment) has multiple benefits: it minimizes the risk of differential attrition in data collection across experimental groups, it reduces the risk of social-desirability bias that may emerge when implementers collect data on the measures that their own intervention is designed to influence, and it enables me to measure its effect on a wide array of outcomes that it had previously been found to affect.

The rest of this article is structured as follows. Section "Experiment" presents the context, study design, and intervention. Section "Data" describes the data. Section "Empirical strategy" discusses the empirical strategy. Section "Results" reports the results. Section "Discussion" discusses implications for research and policy.

Experiment

Context

Schooling in Argentina is compulsory and free from the age of 4 years until the end of secondary school. In 12 out the 24 provinces including Salta, primary education runs from Grades 1 to 7 and secondary education from Grades 8 to 12 (Dirección Nacional de Información y Evaluación de la Calidad Educativa, 2013). The Argentine school system serves 11.4 million students: 1.8 million in pre-school, 4.8 million in primary school, and 3.7 million in secondary school (DiEE, 2016). The school year runs from February to December.

Argentina enroll a larger share of youths in secondary school than most Latin American countries; by the late 2000s, 75% of its youths had started secondary school at the appropriate age, compared to 59% in the average country in the region (Bassi et al., 2013). Yet, its graduation rate at this level lags behind those of its upper middle income neighbors: in 2016, it stood at 63%, compared to 65% in Brazil, 91% in Chile, and 77% in Colombia (Organisation for Economic Co-operation and Development [OECD], 2018). Furthermore, the relative standing of its students in the region has deteriorated: In 2012, Argentina was among the eight lowest performing school systems in all three subjects of the Program for International Student Assessment (PISA), while countries like Brazil, Chile, and Peru had improved and had either caught up with it or surpassed it (OECD, 2013). Many of its students fail to meet national standards. In 2017, 69% of Grade 12 students scored in the lowest two of the four levels of the national assessment in math (below basic, basic, satisfactory, and advanced) and 45% did so in reading (Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación [SEE-MEDN], 2018a).

The Province of Salta is the eighth largest subnational secondary school system In Argentina: In 2016, it served 125,207 students across 394 schools at that level (DiEE, 2016). It is also one of the lower performing systems: in 2017, 73% of Grade 12 students scored in the lowest two levels of the national test in math and 45% did so in reading (SEE-MEDN, 2018c).[1]

Sample

The sample for the study includes 202 public secondary schools in urban and semi-urban areas of the Province of Salta. I arrived at this sample as follows. First, of the 334 secondary schools in the province, I excluded all 94 private schools because I was interested in the potential of the intervention to impact public schools. Then, I dropped all 26 schools in rural areas because they are spread across the province, which would have limited the capacity of the local ministry of education to implement the interventions. (Note, however, that while rural schools account for 7.8% of all public schools in Salta, they only serve 1.2% of students in the province). Finally, I excluded 12 public schools in urban and semi-urban areas with fewer than 10 students in Grade 12 (the target grade of the intervention) to minimize sampling error from small schools.

The schools in the sample are different from out-of-sample schools, regardless of whether I compare them to all out-of-sample schools, public out-of-sample schools, or public and urban or semi-urban out-of-sample schools (see Table A.1 in Appendix A in the online Appendix). Specifically, in-sample schools are larger and have higher repetition rates than all three groups of out-of-sample schools. They also have slightly higher dropout rates across secondary school than the first two groups of out-of-sample schools and slightly lower dropout rates in Grade 12 than the last group.

In-sample schools had lower results on the 2016 national student assessment when compared to all out-of-sample schools, but they performed on par with public out-of-sample schools, except in math (see Table A.2 in the online Appendix). The mixed results in comparisons with public and urban or semi-urban out-of-sample schools may be related to the small number of schools in this group.

Randomization

I randomly assigned the 202 public secondary schools in the sample to (a) a "treatment" group that was offered an intervention (described in the next section) or (b) a "control" group that was not offered the intervention. I stratified the randomization by geographic location (i.e., whether schools were urban or semi-urban) and the school type (i.e., whether schools were "common" or "technical") to increase statistical power. This procedure resulted in 102 treatment and 100 control schools.

Control and treatment schools were comparable on all indicators of school performance tracked by the school system (see Table A.3 in the online Appendix). I find no statistical differences on any indicators in Grade 12, the target grade for the intervention (Panel B), but when I consider all students enrolled in secondary education at these schools (i.e., Grades 8 to 12), treatment schools appear to be smaller and have slightly lower repetition rates (Panel A). I test whether these differences matter by accounting for school-level averages of these indicators in my impact estimation.

Intervention

The growth-mindset intervention administered in Salta was a single-session adaptation of a multi-session version evaluated in the United States (Blackwell et al., 2007).[2] In September of 2017, schools assigned to the treatment group were visited by a representative from the Ministry of Education, Science, and Technology (MECyT) of Salta (locally known as an *Asistente Técnico Territorial* or ATT).[3] The ATT then visited each Grade 12 classroom at the school and proceeded as follows: First, he or she explained the purpose of the activity and sought informed consent from all students (students who chose not to participate were allowed to complete schoolwork in silence). Then, the ATT asked all students who agreed to participate to read a passage on how persisting through difficult challenges can develop the brain and write a letter to a classmate of their choice on the three main lessons from the reading and how they might help him or her.[4] Next, the ATT put up a poster in the classroom with all the letters around it to remind students of the activity for the rest of the school year.[5] Finally, the ATT took a picture of the poster and letters and shared it with the MECyT to verify that the intervention was implemented.

The intervention was scheduled to take place during a non-academic period called *tutorías*, which allows students to bring a wide array of concerns to a designated teacher (*tutor*). It is part of the official curriculum of Salta and of most provinces in Argentina (MECyT, 2012). *Tutorías* cover issues such as student–teacher relations, student body government, or bullying. This study assesses whether using this period for this activity has a positive effect on students. Importantly, *tutores* were not required to be in the classroom during the intervention. The MECyT kindly agreed to purposefully time the delivery of the intervention 2 months before the national assessment because prior studies had found effects of a similar intervention, also administered in a single session during *tutorías*, over this time frame (see Outes et al., 2020).

The reading consists of three parts. The first part seeks to convey the message that, when individuals practice and learn, their brain grows in a similar fashion to muscles after exercise. It explains that the brain is made up of neurons, that connections between neurons allow for problem solving, and that when individuals learn something, these connections multiply. The second part describes research on humans and animals that supports the initial message. It also shows photos of neural connectivity for animals with and without access to stimulating environments and for humans at birth and age of 6 years to illustrate the point from the prior section. The third part contends that, if intelligence can grow through practice at challenging tasks, it makes little or no sense to categorize individuals using labels such as "dumb" or "smart." Then, it concludes by encouraging the reader to engage in practice, even when it seems hard. The reading had been developed for Grade 7 students in New York City. I conducted a pilot in August 2016 with 15 out-of-sample Grade 12 students to check that they could understand the text. I did not make context-specific adaptations, as the developers of the intervention have done in new settings (e.g., Bettinger et al., 2018), to prevent any adjustments I introduced from dampening the effect of an otherwise seemingly effective intervention.

Table 1 shows the theory of change of the intervention, which outlines the hypothesized causal chain linking the intervention to its expected effect. The need that the intervention aims to address is that many students believe that intelligence is static, which leads them to want to look smart and thus engage in a series of counterproductive behaviors that ultimate confirm their deterministic worldview (Dweck & Leggett, 1988). The prevalence of this belief and its association with student achievement have been documented in a variety of settings, including many similar to the one that I study (Chaia et al., 2017; Claro et al., 2016).

I had hypothesized that the intervention would have five main effects. First, students would feel less intimidated by challenging tasks. Students could start perceiving challenging tasks as less difficult (because they anticipated the cognitive gains to be derived from attempting them), they could feel more capable of tackling these tasks (because they believed that, if they persisted, they would eventually solve them), or they could perceive the tasks as a formative experience (as part of the learning process).[6] If students felt less intimidated by challenges, they would exert more effort.[7] This increase in effort could manifest itself in schoolwork, but it could also emerge in other aspects of students' lives, such as their personal development and even existing obligations. Third, the change in mindset could lead students to improve their relationships with peers (by decreasing the threat that they had previously felt from the success of others).[8] Fourth, these changes would lead to improved school performance and achievement, in turn raising students' aspirations to pursue post-secondary education.[9] And ultimately, these improvements could lead students to want to pursue post-secondary education.[10]

TABLE 1

Theory of Change of the Growth-Mindset Intervention in Salta

(1)	(2)	(3)	(4)	(5)
Need	Inputs/activities	Outputs	Outcomes	Impact
• *Students believe intelligence is fixed*, which leads them to want to look smart and thus to avoid challenges, give up in the face of obstacles, see effort as pointless, ignore useful negative feedback, and feel threatened by the success of others	• *Ministry representatives receive training* on existing evidence on the growth-mindset intervention and on how to deliver it • *Ministry representatives visit schools* and secure permission from the principal to implement the intervention • *Ministry representatives deliver the intervention in Grade 12 classrooms* after securing consent from students	• *Students read a passage* on the malleability of intelligence • *Students write a letter to a classmate* on the three main lessons from the reading • *Students post their letters on the classroom*, next to a poster reminding them of the key messages of the intervention for the rest of the school year	• *Students are less intimidated by challenging tasks* because they see them as less difficult, they feel more capable of doing hard work, or they see them as part of a learning process • *Students exert more effort* in school-related tasks, tasks related to personal development, or existing obligations • *Students get along better with peers*, either by improving their existing relationships or at least engaging in fewer acts of hostility of vandalism toward each other	• *Students' performance in school improves* as a result from a shift in their approach to challenges, greater effort, or better relationships with peers • *Student learning increases* as a result from doing better in school • *Students plan to pursue post-secondary education* as a result from their improvement in school performance and learning
Assumptions	• Availability of non-academic period at the school (i.e., *tutorías*) • No opposition from principals, teachers, or students	• Students can read and comprehend the text • Students can write a letter • Students do not mind posting their letters on the classroom	• Teachers do not foster a fixed mindset • Outside factors do not hamper student effort • Other students do not encourage bullying, violence, or vandalism	• School work is attainable for students • Student assessments measure what students learn at school • Students do not need to work to support their families

Source: Author's elaboration.

Prior theoretical and empirical work also suggested that the effects of the intervention could differ based on student- and school-level characteristics. First, the effects on the outcomes above could be larger for students who are more likely to be the subject of stereotypes, including students who are female, from low-income families, and/or who struggle at school.[11] Second, based on recent experiments, the effect of the intervention could vary across schools.[12] Yet, it is not clear which school characteristics would predict treatment heterogeneity. I hypothesized that the effect of the intervention may differ by schools' prior achievement, instructional resources, and supports for low-performing students.[13]

Costs

Part of the increasing enthusiasm for growth-mindset interventions stems from the fact that they can be administered in one session, with little supervision, and are relatively inexpensive. This is certainly the case in developed countries like the United States, where it can be delivered online (see, for example, Gandhi et al., 2019; Paunesku et al., 2015; Yeager et al., 2016, 2019). The potential for deploying these interventions at a low cost is arguably even larger in higher education, where students already regularly interact with instructors online (Oreopoulos et al., 2017, 2018; Oreopoulos & Petronijevic, 2018).

Yet, there is little information about the costs of this intervention in developing countries, where schools lack computers and Internet and the intervention must be conducted on paper. In Peru, where the ministry of education simply shipped the intervention packets to schools, Outes et al. (2020) estimated intervention costs to be only USD$0.20 per student. However, this setup is unlikely to lead schools to implement the intervention in other settings, where teachers are less willing or able to follow instructions without any training or support.

I calculated the costs of administering the intervention by training ministry staff, a model that is more likely to be accepted by education authorities and teachers in developing countries. Specifically, I did so using the ingredients method explained in detail in Dhaliwal et al. (2012). According to those calculations, the total cost of the intervention in Salta was USD$15,632. These include implementation and materials costs (1 hour of salary for the ministry staff in charge of delivering the intervention per classroom and printing costs for the instructions for implementers, instructions for students, and posters for classrooms), which accounted for 72% of the total, and training costs (two hours of salary for the ministry staff who participated in the training session), which accounted for 28% of the total. Considering that it reached an estimated 5,535 students, it cost about USD$2.82 per student, and the marginal cost of adding a classroom of 25 students to the intervention was USD$135. Therefore, the intervention is inexpensive compared to other education interventions (see, for example, Education Endowment Foundation, 2018), but its cost is higher than previously suggested.

Data

As Table 2 shows, I collected data on (a) implementation fidelity (in 2017); (b) students' beliefs, effort, school climate, and plans after secondary school, from surveys in the national assessment (in 2017); (c) schools' resources and supports, from principal surveys in the national assessment (in 2016); (d) students' performance in school, from the census of schools (in 2016 and 2017); and (e) students' achievement from the national assessment (in 2016 and 2017).

Implementation Fidelity

The MECyT of Salta provided me with the pictures submitted by each ATT at each school as proof for implementing the intervention and with the actual pictures.[14] I use these data to confirm that the intervention was implemented as intended in the vast majority of treatment schools and to estimate the effect of receiving the intervention. To my knowledge, this is the first study of a growth-mindset intervention at scale that can verify its implementation.[15]

Students' Beliefs, Effort, School Climate, and Plans for the Future

The MECyT also provided me with the responses of all Grade 12 students in Salta to a survey administered as part of the national assessment, roughly 2 months after the intervention. There are three aspects of this survey worth highlighting. First, it includes multiple questions on behaviors that ought to be affected by the intervention, allowing me to examine each step of its hypothesized causal chain, instead of relying on proxies.[16] Second, the survey was conducted independently from the intervention, which minimizes both the possibility of non-random attrition due to the intervention and of social-desirability bias. Third, the survey is census-based, which means that it seeks to cover all Grade 12 students.[17]

Schools' Resources and Supports

The MECyT also shared the responses of all secondary school principals in Salta to a survey administered at the same time as the national assessment.[18] I use responses to questions on school resources (which enquire about basic conditions, such as whether the school has electricity, and about educational resources, such as whether the school has a library) and school supports for low-performing students (e.g., whether the school develops a personalized plan for students who lag behind) from the 2016 survey to construct two indexes that I interact with the treatment indicator variable to explore heterogeneous effects by school characteristics.

Students' Performance in School

The MECyT also granted me access to all data collected on internal efficiency (e.g., passing, repetition, and dropout rates) through the annual census of schools in Argentina. Importantly, these data are available for the year prior to the intervention, which I use to compare in- and out-of-sample schools and to check balance across experimental groups (sections "Randomization" and "Intervention"), and for the year of the intervention, which I use to estimate impact. The data are reported for secondary schools and for Grade 12 students, allowing me to test for impacts at both levels.

Student Achievement

Finally, the MECyT provided the scores of all Grade 12 students to the national assessment. This assessment evaluates what students know and can do based on the national curriculum. It is administered on an annual basis, but it covers different grades and subjects on each year. In the year prior to the intervention, the Grade 12 test covered math, reading, and natural and social sciences, which I use to compare in- and out-of-sample schools and to check balance (sections "Randomization" and "Intervention"). In the year of the intervention, it only focused on math and reading, which I use to estimate impact. The national ministry of education scaled all scores using a two-parameter logistic Item Response Theory (IRT) model (Yen & Fitzpatrick, 2006), which means that all effects in this article are with respect to the overall national distribution. This feature sets this study apart from most prior evaluations of growth-mindset interventions, which use assessments designed by researchers and administered over a convenience sample.

I estimate the effect of the offer of the intervention (i.e., the intent-to-treat [ITT] effect) by fitting the following model:

$$Y_{is}^{t} = \alpha_{r(s)} + \gamma \overline{Y}_{s}^{t-1} + \beta T_{s} + \varepsilon_{is}^{t}$$

where Y_{is}^{t} is an outcome for student i in school s and year t, $r(s)$ is the randomization stratum of school s and $\alpha_{r(s)}$ is a stratum fixed effect, $\overline{Y}_{s}^{t-1}$ is the school-level average of the same outcome for year $t-1$, and T_s is an indicator variable for random assignment to treatment. (The census of schools and national assessment are repeated cross-sections of Grade 12 students, so I do not observe each student's prior-year outcome). The parameter of interest is β, which captures the causal effect of the intervention. I use cluster-robust standard errors to account for within-school correlations across students in outcomes and include false discovery rate q-values to account for multiple hypothesis testing, using the Simes procedure in the qqvalue program in Stata (Newson, 2009). I also test the sensitivity of my estimates to the inclusion of $\overline{Y}_{s}^{t-1}$. I fit variations of this model that interact the treatment dummy with student characteristics (indicator variables for female students, students from low-income families, and students who had previously repeated a grade) and school characteristics (indexes of prior-year achievement, resources,

TABLE 2

Timeline of the Study

(1) Month	(2) Event	(3) School participation rates: Control schools	(4) School participation rates: Treatment schools
Panel A. 2016			
February	School year starts		
November	National assessment of Grade 12 students (tests of math, reading, natural and social sciences, and principal survey)	96%	93%
December	School year ends		
Panel B. 2017			
February	School year starts		
April	MECyT shares data from national census of schools (2016 school year)	100%	100%
August	MECyT holds training for ATTs	—	100%
September	ATTs deliver the intervention	—	100%
November	National assessment of Grade 12 students (tests of math and reading and student survey)	99%	95%
December	School year ends		
Panel C. 2018			
February	School year starts		
April	MECyT shares data from national census of schools (2017 school year)	100%	100%
December	School year ends		

Note. The table shows the timeline for the interventions and rounds of data collection for the study, including the month in which each event occurred (Column 1), a brief description of the event (Column 2), and the percentage of schools that participated in each event by experimental group (Columns 3 and 4). MECyT refers to the Ministry of Education, Science, and Technology of Salta. ATTs refer to the *Asistentes Técnicos Territoriales* (ATTs); the MECyT staffers who delivered the intervention.

and supports) to estimate the heterogeneous effects of the intervention on these subgroups.[19]

Results

Implementation Fidelity

The intervention was implemented as intended in the vast majority of treatment schools. In 85 of the 102 schools in this group (83%), the MECyT received pictures from the ATTs verifying that at least one Grade 12 section had read the passage, wrote letters, and put them up next to the poster in their classroom. Furthermore, the MECyT received more pictures from schools with more students: The median treatment school had one picture for every 26 students, which is close to the average class size for Grade 12. In seven treatment schools, the ATTs did not send a picture, but a representative of the MECyT called the school and confirmed that the intervention was implemented with the principal. Therefore, the MECyT has verification that the intervention was implemented in 89 treatment schools (90.2% of schools in this group).

The intervention was not implemented in 10 of the 102 schools assigned to receive it (9.8%). In three cases, the principals refused to implement it; in four cases, the ATTs could not find a time that was convenient for them and for the school; and in three cases, the schools were located in areas that were difficult to access and the ATTs could not visit them in time.

ATTs did not track the number of students who did not grant consent for the study, so I do not know the actual share of students in each classroom who participated in the intervention. However, I estimate this share using two different strategies to offer a range of plausible values for students' participation rate in the intervention.

First, I estimate this share by (a) identifying the maximum number of eligible students at each school from the enrollment figures for Grade 12, the target grade for the intervention (using the school performance data described in section "Students' performance in school"); (b) adjusting this number based on the average number of absences self-reported by Grade 12 students (using the student achievement data described in section "Student achievement")[20]; and (c) dividing the result by the number of student letters from the implementation pictures in each school (using the implementation fidelity data described in section "Implementation fidelity").[21] This approach indicates that 58% of students completed the activity. This estimate, however, is extremely conservative because it does not consider that students tend to under-report absences, that some students in the enrollment registers may have already dropped out when the intervention was implemented (2 months before the end of the school year), and that school principals in Argentina face incentives to over-report student enrollment to keep the number of sections (and thus, the number of teachers they are allowed to hire) constant.

Then, I estimate this share by (a) identifying the likely number of eligible students based on the actual students who took the national assessment in Grade 12 (using the data from section "Student achievement"), which was administered 2 months after the intervention and thus offers a more realistic proxy for the actual number of students at the time of the study,[22] and (b) dividing the likely number of eligible students by the number of student letters, as above. This approach indicates that 65% of students completed the activity. This estimate, however, is probably still conservative given that ATTs were not instructed to include all letters from students in their implementation–verification pictures, and accordingly, many of these pictures display the edges of other letters, indicating that some letters were out of the picture frame.

Importantly, both of my estimates of student participation rates are above the 56% response rate in the largest evaluation of a growth-mindset intervention in the United States (see Gopalan & Tipton, 2018). Scaling up my ITT results by my estimates of the student participation rates would make it harder for me to rule out policy-relevant positive effect sizes. Yet, given that only 22 of the 102 treatment schools had pictures that were of high enough resolution to allow me to count the number of student letters, and that even among those schools, pictures did not include all the letters completed in a classroom, it is not possible to know whether such an adjustment would be preferred or even warranted.

Students' Beliefs

In spite of having been implemented with fidelity, the intervention had no effect on students' propensity to find challenging tasks less intimidating. I address this question in three ways, based on my theory of change of the intervention (see discussion in section "Intervention").

First, I explore whether treatment students perceived school-related tasks as less challenging. I identified several questions in the survey administered as part of the national assessment that asked students about the extent to which they found a set of school-related tasks difficult (e.g., paying attention in class) using a scale that ranged from 1 ("very simple") to 4 ("very difficult"). I coded responses dichotomously, using a 1 for "very difficult" or "difficult" and 0 otherwise and analyzed whether treatment students were less likely to find these tasks challenging. The intervention had a precisely estimated zero effect on all outcomes, ruling out even small effects of 4 percentage points (pp.) or more (Table 3).[23]

Then, I examine whether the intervention improved students' beliefs about their self-efficacy. I identified questions in the survey that asked students to indicate whether they understand and do well in math and language using a scale that ranged from 1 ("always") to 4 ("never"). I coded responses dichotomously, using a 1 for "always" or "most of the time" and 0 otherwise. The intervention had a null effect on all outcomes, ruling out effects larger than 5 pp. (Table 4).

Finally, I consider whether treatment students were more likely to see tests as formative. I used

TABLE 3

ITT Effect on Students' Perceptions of Difficulty of Schoolwork (2017)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Understanding texts	Writing texts	Speaking in public	Learning new concepts	Paying attention in class	Working in groups	Participating in class	Solving problems
Treatment	0.0049	−0.0056	0.0117	0.0003	0.0089	0.0089	0.0027	−0.0057
	(0.0112)	(0.0122)	(0.0153)	(0.0127)	(0.0110)	(0.0102)	(0.0136)	(0.0141)
Observations	9,372	9,372	9,372	9,372	9,372	9,372	9,372	9,372
R^2	0.002	0.001	0.001	0.002	0.005	0.001	0.006	0.003
Control mean	0.200	0.254	0.398	0.307	0.196	0.193	0.309	0.438
FDR q-value	0.916	0.916	0.916	0.984	0.916	0.916	0.962	0.916

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' perceived difficulty of tasks related to schoolwork. Students were asked to indicate how difficult they found the activities listed above using a scale ranging from 1 ("very simple") to 4 ("very difficult"). The dependent variables in this table are dummies that equal one for students who indicated the task was difficult or very difficult and zero otherwise. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

TABLE 4

ITT Effect on Students' Self-Efficacy, by Subject (2017)

	Math		Language	
	(1)	(2)	(3)	(4)
	I understand it quickly	I do well in it	I understand it quickly	I do well in it
Treatment	0.00004	0.01094	−0.01531	−0.00638
	(0.01630)	(0.01581)	(0.01928)	(0.01869)
Observations	9,372	9,372	9,372	9,372
R^2	0.001	0.002	0.003	0.004
Control mean	0.348	0.424	0.563	0.581
FDR q-value	0.998	0.978	0.978	0.978

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' perceived performance on math and language. Students were asked to indicate how often they agreed with the statements listed above using a scale ranging from 1 ("always") to 4 ("never"). The dependent variables in this table are dummies that equal one for students who indicated they agreed always or most of the times and zero otherwise. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

questions that explicitly asked students about the extent to which assessments served formative purposes, which employed the same scale as above. Once again, the intervention had a precisely estimated null effect on all outcomes, ruling out effects larger than 3 pp. (Table 5).

These null effects, however, mask heterogeneous effects for groups of disadvantaged students. I find some evidence that the intervention may have negatively impacted the beliefs of female students, students from low-income families, and those who had previously repeated a grade. I created indexes of students' perceptions of the difficulty of schoolwork, self-efficacy, and perceptions of the usefulness of classroom assessments by conducting principal component analyses of variables in Tables 3 through 5 and taking the first principal component of each analysis. Then, I estimated the effect of the intervention on the indexes (not on the individual variables) for each group to reduce the probability of false positives due to multiple hypothesis testing. The intervention seems to have *increased* the perceived difficulty of schoolwork among girls, *decreased* the self-efficacy of girls and students from low-income families, and *decreased*

TABLE 5

ITT Effect on Students' Perceptions of Classroom Assessments (2017)

	(1)	(2)	(3)
	Tests help me improve	Tests help me identify errors	Tests check if I understood what I was taught
Treatment	−0.0041	−0.0011	0.0029
	(0.0107)	(0.0079)	(0.0130)
Observations	9,112	9,372	9,372
R^2	0.008	0.000	0.001
Control mean	0.875	0.150	0.507
FDR q-value	0.894	0.894	0.894

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' perceived usefulness of classroom assessments. Students were asked to indicate how often they agreed with the statements listed above using a scale ranging from 1 ("always") to 4 ("never"). The dependent variables in this table are dummies that equal one for students who indicated they agreed always or most of the times and zero otherwise. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

the perceived usefulness of assessments among students who had repeated a grade (see Table A.5 in the online Appendix). Surprisingly, even when the coefficient on the interaction term is not statistically significant, its sign is typically the opposite of what I predicted in the theory of change of the intervention.

I do not find any evidence of heterogeneous effects on students' beliefs by school characteristics. I interact the treatment dummy with each school's prior-year average score on the national assessment (across all subjects), an index of school resources, and an index of school supports (see section "Empirical Strategy") and find no statistically significant interaction effects along these dimensions. Yet, most interactions are imprecisely estimated, so it is possible that they exist but I lack sufficient statistical power to detect them (see Table A.6 in the online Appendix).

Student Effort

I also estimate the impact of the intervention on three sets of indicators of student effort. I begin by focusing on school-related tasks, the domain in which I most expected to see changes. I examine whether treatment students were more likely to attend school or private tuition.[24] The survey in the national assessment asks how often students missed school during the year using a scale from 1 ("never") to 4 ("more than 24 times"). I coded responses that constituted "chronic absenteeism" (15 absences or more, see Gottfried, 2014) as 1 and 0 otherwise. The question on tuition was a yes/no question, so I coded answers dichotomously. Surprisingly, treatment students were 3.7 pp. *more* likely than their control peers to miss school (Table 6). However, this difference is only marginally statistically significant, and as the q-value indicates, likely to have emerged due to multiple hypothesis testing. I find no effect on the intervention on students' propensity to attend tuition.

I also examine whether treatment students worked harder on their personal development (e.g., read books outside of school, take art lessons, learn a foreign language, or play sports). All of these were yes/no questions, so I coded them dichotomously. Again, I find a precisely estimated zero effect on all outcomes, allowing me to rule out effects larger than 4 pp.

Finally, I consider whether the intervention increased student effort on existing obligations (e.g., work at or outside of home). Both were yes/no questions and were coded dichotomously. Once again, I find precisely estimated null effects on all outcomes.

I find no evidence of heterogeneous effects on any group of variables measuring student effort. I created indexes of student effort on

TABLE 6

ITT Effect on Student Effort (2017)

	School-related tasks		Personal development				Existing obligations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Chronically absent to school	Attends private tuition	Reads books outside of school	Takes art lessons	Learns a foreign language	Plays sports	Works at home	Works outside of home
Treatment	0.0374*	−0.0070	0.0065	−0.0046	0.0090	0.0084	−0.0080	0.0084
	(0.0206)	(0.0229)	(0.0137)	(0.0110)	(0.0091)	(0.0136)	(0.0172)	(0.0149)
Observations	9,372	9,002	8,236	8,243	8,156	8,558	8,969	8,946
R^2	0.017	0.013	0.005	0.003	0.003	0.009	0.023	0.002
Control mean	0.288	0.233	0.343	0.166	0.083	0.617	0.459	0.264
FDR q-value	0.572	0.761	0.761	0.761	0.761	0.761	0.761	0.761

Note. This table shows the intent-to-treat (ITT) effect of the intervention on student effort. Students were asked to indicate how many schooldays they had missed during the year using a scale ranging from 1 (“never”) to 4 (“more than 24 days”). The dependent variable on absenteeism is a dummy that equal one for students who reported to have missed 15 or more days and zero otherwise. All questions on personal development and existing obligations were yes/no questions and were coded dichotomously. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

school-related tasks, personal development, and existing obligations using the first principal component from separate principal component analyses. Then, I estimated the effect of the intervention on these indexes for the same groups as above. The coefficients on the interactions are around zero and statistically insignificant (see Table A.7 in the online Appendix). I find no evidence of heterogeneity by school characteristics; in fact, most interaction effects are estimated around zero, allowing me to rule out small-to-moderate effects (see Table A.8 in the online Appendix).

School Climate

Next, I estimate the effect of the intervention on three measures of school climate. I first focus on a question that asked students whether they get along with their peers using a scale from 1 (“no, I do not get along with anyone”) to 5 (“yes, I get along with everyone”). I coded responses as 1 if students indicated that they got along with some, most, or all their peers and 0 if they reported that they did not get along with anyone or with only a few peers. The intervention reduced students’ propensity to get along with peers by 1.5 pp., but as the q-value indicates, this effect is likely to have emerged due to multiple hypothesis testing (Table 7).

I also estimate the effect of the intervention on the student-reported prevalence of bullying. The survey in the national assessment asks students how often peers at their school engage in bullying on a number of groups using a scale that ranges from 1 (“always”) to 4 (“never”). I coded responses as 1 if students indicated bullying occurred “often” or “always” and 0 otherwise. Consistent with the results above, I find that the intervention increased the prevalence of bullying against female students by 1.5 pp. and had no effects on other types of bullying. Again, however, this effect seems to have emerged due to multiple hypothesis testing.

Finally, I consider whether the intervention had any effect on student-reported vandalism, which was measured using the same scale and which I coded in the same manner as above. I did not find that the intervention affected the incidence of theft or damages to school property.

I find little evidence of heterogeneous effects on any of the variables measuring school climate. I used the first indicator variable in Table 7 by itself and created indexes of bullying and vandalism using the first principal component from separate principal component analyses. I estimated the effect of the intervention on these

TABLE 7

ITT Effect on School Climate (2017)

		Bullying				Vandalism	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gets along with peers	Bullying on students with good grades	Bullying on students who repeat grades	Bullying on students because of their personal or family characteristic	Bullying on female students	Stealing	Damaging school property
Treatment	−0.0149**	0.0074	−0.0016	0.0126	0.0153**	0.0164	0.0106
	(0.0072)	(0.0133)	(0.0119)	(0.0146)	(0.0074)	(0.0167)	(0.0182)
Observations	9,150	9,372	9,372	9,372	9,372	9,372	9,372
R^2	0.002	0.001	0.002	0.004	0.002	0.006	0.011
Control mean	0.904	0.192	0.197	0.252	0.079	0.160	0.314
FDR q-value	0.141	0.672	0.894	0.672	0.141	0.672	0.672

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' perceived school climate. Students were asked to indicate how frequently other students at their school engaged in the activities listed above using a scale ranging from 1 ("always") to 4 ("never"). The dependent variables in this table are dummies that equal one for students who indicated that the activities occurred always or many times and zero otherwise. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

variables for the same groups as above. The intervention had a *negative*, but marginally statistically significant, effect on the propensity of female students to get along with peers, but all other interaction terms were consistently estimated around zero and statistically insignificant (see Table A.9 in the online Appendix). I do not find any evidence of heterogeneous effects on school climate by school characteristics (see Table A.10 in the online Appendix).

Students' Performance in School

I also estimate the effect of the intervention on students' performance in school, as measured by the number of enrolled students, and the percentage of students who passed, failed, or repeated the grade, or who dropped out of school. I do not find evidence that the intervention had a positive effect on these outcomes, but my estimates are more imprecise than those for other outcomes because these data are collected at the school level. (This is also why I cannot estimate heterogeneous effects on these outcomes by students' characteristics). The results are similar when I account for schools' performance in the year before the intervention (Table 8).

Student Achievement

Then, I estimate the effect of the intervention on student achievement, as measured by the results of the national assessment of math and reading in Grade 12. I find no evidence that the intervention improved test scores in either subject, before or after accounting for the schools' performance in the year prior to the intervention (Table 9). I can rule out effects larger than .07 standard deviations in both subjects. In fact, the distribution of student achievement looks nearly identical across the control and treatment groups, 2 months after the intervention (see Figure A.1 in the online Appendix). Furthermore, I find no evidence of heterogeneous effects by students' sex, socioeconomic status, or prior repetition (see Table A.11 in the online Appendix).

Interestingly, all interactions between the treatment and school-level characteristics are negative, suggesting that schools with higher levels of achievement, resources, and supports benefit less from the intervention. The only statistically significant interaction, however, is the one between the treatment and school resources for math (see Table A.12 in the online Appendix).

TABLE 8

ITT Effect on Students' Performance in School (2017)

	Number of students enrolled		Percentage of students who passed the grade		Percentage of students who failed the grade		Percentage of students who dropped out of school		Percentage of students who repeated the grade	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	−10.35	−10.46	−2.28	−2.21	1.93	1.86	−0.29	−0.27	0.01	0.01
	(6.58)	(6.59)	(2.64)	(2.63)	(2.48)	(2.47)	(0.95)	(0.95)	(0.51)	(0.51)
Prior-year school index		1.51		−1.43*		1.27		−0.32		0.01
		(2.08)		(0.83)		(0.78)		(0.30)		(0.16)
Observations	189	189	195	195	195	195	199	199	195	195
R^2	0.331	0.332	0.255	0.267	0.292	0.302	0.033	0.039	0.007	0.007
Control mean	68.88		72.15		25.18		3.39		2.46	
FDR q-value	0.588	0.588	0.753	0.753	0.753	0.753	0.970	0.970	0.985	0.985

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' performance in school. This information is collected at the school level through the national census of schools. The prior-year school index is the first principal component from a principal component analysis that included the enrollment, passing, failure, repetition, and dropout rates for Grades 8 to 12 in all schools in the sample for the 2016 school year. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

TABLE 9

ITT Effect on Student Achievement (2017)

	Math (IRT-scaled score)		Reading (IRT-scaled score)	
	(1)	(2)	(3)	(4)
Treatment	−0.020	0.015	−0.051	−0.008
	(0.049)	(0.035)	(0.063)	(0.043)
Prior-year school index		0.641***		0.801***
		(0.107)		(0.117)
Observations	8,814	8,814	8,865	8,865
R^2	0.025	0.076	0.018	0.067
Control mean	−0.259		−0.055	
FDR q-value	0.853	0.853	0.853	0.853

Note. This table shows the intent-to-treat (ITT) effect of the intervention on student achievement. All scores have been scaled using a two-parameter logistic Item Response Theory (IRT) model with respect to the national distribution to have a mean of zero and a standard deviation of one. Prior-year school achievement refers to the first principal component from a principal component analysis that included school-level average scores in assessments of math, reading, natural, and social sciences in Grade 12 during the 2016 school year. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.

Students' Plans After Secondary Education

Finally, I estimate the effect of the intervention on students' post-secondary education plans. Students were asked whether they planned to study, work, or do both, so I coded each option dichotomously. I do not find any indication that the intervention affected the plans of the average student (Table 10) or of the sub-groups of students mentioned above (see Table A.13 in the online Appendix).

I find some evidence of heterogeneity by school characteristics. First, in schools with

TABLE 10

ITT Effect on Plans After Secondary Education (2017)

	(1)	(2)	(3)
	Plans to work	Plans to study	Plans to do both
Treatment	−0.001	−0.004	0.013
	(0.006)	(0.020)	(0.016)
Observations	9,377	9,377	9,377
R^2	0.004	0.006	0.001
Control mean	0.043	0.453	0.360
FDR q-value	0.903	0.903	0.903

Note. This table shows the intent-to-treat (ITT) effect of the intervention on students' post-secondary plans. The dependent variables in the regressions are indicator variables for students who indicated that they plan to work, study, or do both after secondary education. All estimations include randomization strata fixed effects. FDR = False Discovery Rate.
*Significant at 10%. **Significant at 5%. ***Significant at 1%.

higher levels of achievement, the intervention increased the share of students who plan to work and study after secondary school. Second, in schools with more resources, the intervention increased the share of students who plan to work. Third, in schools with more supports, the intervention reduced the share of students who plan to study and increased the share of students who plan to work and study by a similar magnitude (see Table A.14 in the online Appendix).

Discussion

Implications for Research

The present study highlights the importance of evaluating promising educational interventions at scale to understand their effectiveness when they are implemented within a school system. The null effects that I found differ considerably from the encouraging results of efficacy trials and they are more consistent with the results from two recent large-scale impact evaluations. Outes et al. (2020) evaluated the intervention in 800 secondary schools in Peru. They found that it raised achievement in math (by .05 standard deviations), but not in reading comprehension. Effects were driven by one region; results for the other two were not statistically significant. Yeager et al. (2019) evaluated the intervention in 65 secondary schools in the United States. They found that it had no effects on the grades of or courses taken by the average student, but low performers improved their grades and high performers took more challenging classes.

This study, when read alongside the two other effectiveness trials, also suggests that the intervention only improves achievement when it changes students' beliefs about intelligence. In Salta, I found that it had no effect on beliefs (see section "Students' beliefs"), so it is perhaps not surprising that it had no impact on effort, climate, performance, or achievement (see sections "Student effort" to "Students' plans after secondary education"). In Peru, Outes et al. (2020) found that the intervention only had a positive impact on math achievement in Ancash, where it also improved students' self-beliefs in math. They found no such effects on beliefs or achievement in Junín or Lima, the two other regions. In the United States, Yeager et al. (2019) found that the intervention only improved grades among low performers, who not only changed their mindsets but also had margin for improvement. These studies draw attention to the importance of piloting the intervention to ensure that it changes students' mindsets before evaluating its impacts on school performance or achievement.

The studies in Salta, Peru, and the United States also raise important questions about how context may moderate the effects of the intervention. Context may matter for at least four reasons. First, systems, schools, and classrooms may differ in their capacity to implement the intervention. Outes et al. (2020) found that Ancash, the region of Peru that most benefited from the intervention, had implemented it with greater fidelity than the other two regions. Yet, the Salta study shows that the intervention can fail even when it is implemented correctly. Second, systems, schools, and students may differ in their margin for potential improvements.[25] As the authors of the Peru study note, Ancash is also the most rural of the three regions, so it is also possible that it benefited the most because it started from a lower level of performance. This interpretation is consistent with the results of the U.S. study for low-performing students. Third, students may also differ in their baseline beliefs about the malleability of intelligence. It is possible that students in Salta, Junín, and Lima did not change their beliefs or raise their effort because they did not hold a fixed mindset before they participated in the intervention. This would be consistent with the results of some previous studies in the United States, where the intervention has been impactful among students with fixed mindsets (e.g., Yeager et al., 2014), but Yeager et al. (2019) do not find treatment heterogeneity by students' baseline mindsets. Finally, schools and teachers may differ in their capacity to help students increase their effort. Yeager et al. (2019) see this as the reason why the intervention has larger effects in schools whose students exhibit challenge-seeking behaviors. Yet, the Salta study finds no heterogeneity across schools with different levels of resources or supports for low-performing students.

The differences in the results of these studies also raise questions about how the intervention may change mindsets. The focus has been on the reading that students are asked to complete. Yet, there are at least two important differences in how the intervention was delivered across Salta, Peru, and the United States that may play a more important role than previously anticipated. One difference is whether students are required to check their understanding of the reading (before they are asked to write a letter to a classmate on the main lessons from the passage). In Peru, students were asked to answer review questions and discuss the reading in groups. In the United States, students were asked to summarize the findings of the reading in their own words. This step may be especially important in developing countries, where reading skills are low, and it may partly explain why the intervention had no effects in Salta, where it was omitted.[26] Another difference is whether the activity is led by the students' teachers, as it was in Peru. This could potentially both educate teachers and influence their interactions with students (for a broader discussion of this possibility, see Raudenbush, 1984; Yeager & Walton, 2011). The relative importance of this aspect, however, is unclear, as the intervention in the United States was effective even if it was delivered online and teachers did not know which students received it.

These studies also offer several lessons for the design of future evaluations of this intervention. First, they highlight the importance of not only evaluating the intervention at scale but also of having sufficient statistical power to detect heterogeneous treatment effects across sites. Second, they illustrate the usefulness of measuring students' pre-intervention mindsets and their post-intervention understanding of the reading to make sense of potential null results. This objective may be achieved either by combining lab and field experiments or by embedding the former in the latter to keep data collection costs manageable. Third, these studies make clear how essential it is to collect information on students' backgrounds and schools' resources and practices to examine heterogeneous effects along these dimensions.

Implications for Policy

The present study draws attention to the importance of context, intervention design, and implementation in taking education initiatives to scale in developing countries. The case of Salta suggests that delivering the growth-mindset intervention using materials and following processes that have yielded positive effects in other settings will not necessarily lead to similar results (see Yeager & Walton, 2011). Furthermore, the costs of implementing it are not trivial and should be compared against those of initiatives with evidence of effectiveness in these settings (see Ganimian & Murnane, 2016).

This study also offers governments interested in implementing the intervention guidance on some of the aspects that they should consider when deciding whether and how to do so. First, they should try to understand whether potential beneficiaries hold a fixed mindset and whether the extent to which they hold such beliefs is related to their academic performance. They should also consider whether schools will seek to implement the intervention with little training or support (as in Peru) or with both (as in Salta).[27] This decision will play an important role in determining the costs of implementation. Finally, if possible, governments should consider using data already collected by their school system to evaluate the impact of the first iteration of the intervention through a randomized rollout. This will reduce costs, avoid bias in responses, and minimize differential participation and allow the government to understand whether the intervention works for their school system.

Conclusion

I present experimental evidence on a growth-mindset intervention implemented at scale in public secondary schools in Salta, Argentina, and find it had no effects, either on intermediate outcomes (e.g., students' beliefs, effort, or school climate) or the ultimate outcomes of interest (e.g., students' performance in school, achievement, and post-graduation plans). Nearly all results are precisely estimated and allow me to rule out even small effects. I find little evidence of heterogeneous effects by students' sex, socioeconomic status, and prior grade repetition, or by schools' educational resources and support for low-performing students.

This study and my review of the literature seek to raise important questions about the effectiveness of growth-mindset interventions when implemented at scale in developing countries. It does not seek to call into question the efficacy of variations of this intervention when implemented by its developers among small convenience samples of schools and students, let alone the decades of work that developed the theory on which these interventions are based. It simply proposes a way forward for identifying the conditions that would maximize impact.

Acknowledgments

I gratefully acknowledge the funding provided by the Abdul Latif Jameel Poverty Action Lab's Post-Primary Education initiative and the Inter-American Development Bank for this study. I thank Andrea Bergamaschi, Analía Berruezo, Loreto Biehl, Gloria Crespo, Elena Duro, Miriam Goldszier, and Gabriela Guerrero for making this study possible. I also thank Hunt Allcott, Felipe Barrera-Osorio, Clancy Blair, Susana Claro, Sean Corcoran, Andy de Barros, Rajeev Dehejia, Bill Easterly, Raquel Fernández, Jill Gandhi, Isaac Mbiti, Karthik Muralidharan, Dick Murnane, Cybele Raver, Martin Rotemberg, David Yeager, Hiro Yoshikawa, Tyler Watts, Marty West, seminar participants at NYU, Paco Martorell, and two anonymous reviewers for comments that informed this draft. Aditi Bhowmick, Nicolás Buchbinder, and María Cortelezzi provided excellent research assistance. This study was registered with the AEA Trial Registry (RCT ID: AEARCTR-0002970). It was approved by the institutional review board of the Massachusetts Institute of Technology. All views expressed are my own and not of any institution with which I am affiliated.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author received funding from the Abdul Latif Jameel Poverty Action Lab's Post-Primary Education Initiative and the Inter-American Development Bank for this study.

Notes

1. The reason why Salta is one of the lower performing school systems in the country but has similar percentages of students in the lowest two levels as the national average is that two-thirds of students in Argentina go to school in the Province of Buenos Aires (a single-school system), which generally drives the national averages in the national assessment (SEE-MEDN, 2018b).
2. It should be noted, however, that several studies have found positive effects of similar interventions after two sessions (Good et al., 2003; Yeager et al., 2014) or one session (Mendoza-Denton et al., 2008; Paunesku et al., 2015; Yeager et al., 2013, Study 3; Yeager et al., 2011) including a 15-min session (Yeager et al., 2013, Study 2).
3. ATTs have teaching degrees and either serve or have served in the past as teachers. The MECyT trained all ATTs on how to deliver the intervention in August of 2017, using materials I had prepared.
4. The original English version of the reading can be accessed at: https://bit.ly/2IRAJI5. The Spanish translation used in Salta can be found at: https://bit.ly/2YfL1VS.
5. This component of the intervention was first used by Outes et al. (2020) in Peru. The original English version of the poster can be accessed at: https://bit.ly/2HWQfQJ. The Spanish translation used in Salta can be found at https://bit.ly/2Tl0HU9. The poster

was translated by Mind-set Works, the organization that had developed the original version.

6. Several studies have examined whether the growth-mindset intervention impacts students' perceived difficulty of school-related tasks (Burnette et al., 2018; Mendoza-Denton et al., 2008).

7. Prior studies have documented the effect of mindset interventions on motivation (Blackwell et al., 2007; Eccles et al., 1998), but few have included actual measures of effort.

8. This expectation was informed by the evidence on the effect of mindset interventions on stereotype threat (Aronson et al., 2002; Good et al., 2003) and hostile intent, aggression, and desire to seek revenge (Yeager et al., 2011, 2013).

9. Multiple evaluations of mindset interventions have found effects on school performance (Blackwell et al., 2007; Paunesku et al., 2015), but only a few have evaluated its effect on achievement on standardized tests (Good et al., 2003).

10. Several studies have found that mindset interventions can affect students' post-secondary education plans (Outes et al., 2020; Yeager et al., 2019).

11. Multiple studies have found that the intervention only works or works best for these sub-groups of students (Aronson et al., 2002; Broda et al., 2018; Good et al., 2003; Paunesku et al., 2015; Yeager et al., 2016, 2019).

12. The two largest field experiments in this literature document considerable treatment heterogeneity across schools (Outes et al., 2020; Yeager et al., 2019).

13. To my knowledge, only one study has examined treatment heterogeneity by school characteristics (Yeager et al., 2019).

14. Unfortunately, the photos are not of high enough quality to allow me to analyze the content of the letters (e.g., to gauge whether students understood or were persuaded by the reading).

15. In Peru, Outes et al. (2020) also asked schools that were randomly assigned to implement a similar intervention to submit pictures, but they received such pictures for less than half of treatment schools.

16. The original survey in Spanish can be accessed at https://bit.ly/2I0C39h.

17. Salta has traditionally had high participation rates in the national assessment. In 2016, 92% of all public secondary schools and 83% of all students in these schools participated in the assessment (SEE-MEDN, 2016). In 2017, 97% of public schools and 79% of students at this level participated (SEE-MEDN, 2018c).

18. The original survey in Spanish can be accessed at https://bit.ly/2WhogPp.

19. The index of prior-year achievement is the school-level average score in the 2016 national assessment, which covered math, reading, and natural and social sciences (see section "Student achievement"). The indexes of school resources and supports are the first principal components from principal component analyses of questions in the 2016 survey of principals on the resources and supports for low-performing students at the school, respectively.

20. I imputed the mean absence rate in the treatment group for four schools without absence data.

21. I imputed the mean number of letters for schools without clear pictures, under the assumption that the resolution of the pictures of student letters (which is largely determined by the quality of the camera of each implementer's smart phone) is orthogonal to actual implementation fidelity.

22. These assessments are not attached to any stakes and the National Education Law of 2006 expressly prohibits the dissemination of achievement data at the school, teacher, or student level (see Ganimian, 2015), so schools face no incentives to discourage lower achieving students from taking the exam.

23. Throughout the manuscript, when I state that I can rule out effects of a given magnitude, I am referring to the upper bound of the 95% confidence interval (see, e.g., Hoxby, 2000). For example, the upper bound of the first estimate in Table 3 is 2.6 pp., so effects above this magnitude are unlikely. When I make this claim and multiple related hypotheses are being tested, I use the largest upper bound that I observe in a family. For example, in Table 3, I state that I can rule out effects larger than 4 pp. because that is the largest upper bound I observe across all outcomes in that table.

24. In Argentina, the word "tuition" (*apoyo escolar*) refers not only to fee-charging private providers but also to programs offered by the government and non-profits for free. Therefore, cost is not as much of a barrier as the word may suggest from its use in other developing countries.

25. A variation of this argument is that Grade 12 students, who are about to graduate from secondary school, may have fewer reasons to change their beliefs and mindsets than Grade 9 students, who are transitioning into what is known as middle school in the United States and as lower secondary school in other countries.

26. I explored whether the effect of the intervention in Salta varied either by students' self-assessment of their capacity to understand texts or by their schools' prior-year reading levels, but did not find any evidence of heterogeneous effects on mechanisms or outcomes. These results are available upon request.

27. As mentioned in the section "Costs," very few developing countries have the requisite technological infrastructure to deliver this intervention online throughout the school system.

References

Andrabi, T., Das, J., & Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, *107*(6), 1535–1563. https://doi.org/10.2139/ssrn.2538624

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

Avitabile, C., & de Hoyos, R. (2014). *The heterogeneous effects of information about the returns to schooling on student learning: Evidence from a randomized control trial in Mexico*. World Bank Group.

Banerjee, A. V., Glewwe, P., Powers, S., & Wasserman, M. (2013). *Expanding access and increasing student learning in post-primary education in developing countries: A review of the evidence* [Unpublished manuscript]. Abdul Latif Jameel Poverty Action Lab (J-PAL).

Barrera-Osorio, F., Deming, D., González, K., & Lagos, F. (2020). *Providing performance information in education: An experimental evaluation in Colombia* [Unpublished manuscript]. Harvard Graduate School of Education (HGSE).

Bassi, M., Busso, M., & Munoz, J. S. (2013). *Is the glass half empty or half full? School enrollment, graduation, and dropout rates in Latin America* (IDB Working Paper Series No. IDB-WP-462). Inter-American Development Bank.

Berniell, L. (2014). *Condicionantes de la entrada de los jóvenes al mundo laboral: Educación, información y oportunidades laborales de calidad* (AAEP 2014). Banco de Desarrollo de América Latina (CAF) [*Factors that condition the entry of youths to the labor market: Education, information, and quality work opportunities* (AAEP 2014). Development Bank of Latin America (CAF)].

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

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246–263.

Bobba, M., & Frisancho, V. (2016). *Learning about oneself: The effects of signaling academic ability on school choice* [Unpublished manuscript]. Inter-American Development Bank.

Bonilla, L., Bottan, N. L., & Ham, A. (2016). *Information policies and higher education choices: Experimental evidence from Colombia* [Unpublished manuscript]. University of Illinois.

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

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

Camargo, B., Camelo, R., Firpo, S., & Ponczek, V. (2018). Information, market incentives, and student performance: Evidence from a regression discontinuity design in Brazil. *The Journal of Human Resources*, *53*(2), 414–444.

Chaia, A., Child, F., Dorn, E., Frank, M., Krawitz, M., & Mourshed, M. (2017). *Drivers of student performance: Latin America insights*. McKinsey & Company.

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

Dhaliwal, I., Duflo, E., Glennerster, R., & Tulloch, C. (2012). Comparative cost-effectiveness analysis to inform policy in developing countries: A general framework with applications for education. In P. Glewwe (Ed.), *Education policy in developing countries*, pp. 285–338. The University of Chicago Press and Abdul Latif Jameel Poverty Action Lab (J-PAL).

DiEE (2016). *Anuario estadístico 2016*. Buenos Aires, Argentina: Dirección de Información y Estadística Educativa (DiEE) [*Statistical annex 2016*. Buenos Aires, Argentina: Directorate of Education Information and Statistics (DiEE)].

DiNIECE (2013). *Redefiniciones normativas y desafíos de la educación secundaria en Argentina. Acuerdos federales en un sistema descentralizado*. (*La Educación en Debate, Documentos de la DiNIECE No. 10*). Buenos Aires, Argentina: Dirección Nacional de Información y Evaluación de la Calidad Educativa (DiNIECE) [*Normative redefinitions and challenges of secondary school in Argentina. Federal agreements in a decentralized system.* (*Educational Debates, DiNIECE Document No. 10*). Buenos Aires, Argentina: National Directorate of Information and Evaluation of Education Quality (DiNIECE)].

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, *109*(8), 2728–2765.

Dommett, E. J., Devonshire, I. M., Sewter, E., & Greenfield, S. A. (2013). The impact of participation in a neuroscience course on motivational measures and academic performance. *Trends in Neuroscience and Education*, *2*, 122–138.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273.

Dweck, C. S., Walton, G. M., & Cohen, G. L. (2014). *Academic tenacity: Mindsets and skills that promote long-term learning*. Bill & Melinda Gates Foundation.

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

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Handbook of child psychology, Vol. 3: Social, emotional, and personality development* (5th ed., pp. 1017–1095). Wiley.

Education Endowment Foundation. (2018). *Teaching and learning toolkit*. https://educationendowmentfoundation.org.uk/public/files/Toolkit/complete/EEF-Teaching-Learning-Toolkit-October-2018.pdf

Gandhi, J., Watts, T. W., Masucci, M. D., & Raver, C. C. (2019). *The effects of two mindset interventions on low-income students' academic and psychological outcomes* [Unpublished manuscript]. New York University.

Ganimian, A. J. (2015). *El termómetro educativo: Informe sobre el desempeño de Argentina en los Operativos Nacionales de Evaluación (ONE) 2005-2013*. Proyecto Educar 2050. [*Educational thermometer: Report on the performance of Argentina on the National Operatives of Educational Assessment (ONE) 2005-2013*. Buenos Aires, Argentina: Educate Project 2050.]

Ganimian, A. J., & Murnane, R. J. (2016). Improving educational outcomes in developing countries: Lessons from rigorous evaluations. *Review of Educational Research*, *86*(3), 719–755.

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

Gopalan, M., & Tipton, E. (2018). *Is the National Study of Learning Mindsets nationally representative?* [Unpublished manuscript]. Pennsylvania State University.

Gottfried, M. A. (2014). Chronic absenteeism and its effects on students' academic and socioemotional outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, *19*(2), 53–75.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics*, *115*(4), 1239–1285.

Loyalka, P., Liu, C., Song, Y., Yi, H., Huang, X., Wei, J., . . . Rozelle, S. (2013). Can information and counseling help students from poor rural areas go to high school? Evidence from China. *Journal of Comparative Economics*, *41*, 1012–1025. https://doi.org/10.1016/j.jce.2013.06.004

Mendoza-Denton, R., Kahn, K., & Chan, W. (2008). Can fixed views of ability boost performance in the context of favorable stereotypes? *Journal of Experimental Social Psychology*, *44*(4), 1187–1193.

Ministerio de Educación, Ciencia y Tecnología, Gobierno de la Provincia de Salta. (2012). *Diseño curricular para la educación secundaria*. Ciudad de Salta, Salta: Ministerio de Educación, Ciencia y Tecnología de Salta (MECyT) [*Curricular design for secondary education*. City of Salta, Salta: Ministry of Education, Science, and Technology of Salta (MECyT)].

Newson, R. (2009). *QQVALUE: Stata module to generate quasi-q-values by inverting multiple-test procedures*. Statistical Software Components. Department of Economics, Boston College.

Nguyen, T. (2009). *Information, role models and perceived returns to education: Experimental evidence from Madagascar* [Unpublished manuscript]. Massachusetts Institute of Technology.

Oreopoulos, P., Brown, R. S., & Lavecchia, A. M. (2017). Pathways to education: An integrated approach to helping at-risk high school students. *Journal of Political Economy*, *125*(4), 947–984.

Oreopoulos, P., Patterson, R. W., Petronijevic, U., & Pope, N. G. (2018). *Lack of study time is the problem, but what is the solution? Unsuccessful attempts to help traditional and online college students* (NBER Working Paper No. 25036). National Bureau of Economic Research (NBER).

Oreopoulos, P., & Petronijevic, U. (2018). Student coaching: How far can technology go? *Journal of Human Resources*, *53*(2), 299–329.

Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results: What students know and can do. Student performance in mathematics, reading and science* (Vol. I).

Organisation for Economic Co-operation and Development. (2018). *Education at a glance 2018: OECD indicators*.

Outes, I., Sánchez, A., & Vakis, R. (2020). *The power of believing you can get smarter: The impact of a growth-mindset intervention on academic achieve-*

ment in Peru (Policy Research Working Paper No. 9141). The World Bank.

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

Raudenbush, S. W. (1984). Magnitude of teacher expectancy effects on pupil IQ as a function of the credibility of expectancy induction: A synthesis of findings from 18 experiments. *Journal of Educational Psychology*, *76*(1), 85.

Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación. (2016). *Aprender 2016: Informe de resultados*. Buenos Aires, Argentina [*Learn 2016: Report of results*. Buenos Aires, Argentina: Secretariat of Educational Assessment, National Ministry of Education and Sports (SEE-MEDN)].

Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación. (2018a). *Aprender 2016: Informe de resultados de secundaria*. Buenos Aires, Argentina: Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación (SEE-MEDN) [*Learn 2016: Report of results for secondary school*. Buenos Aires, Argentina: Secretariat of Educational Assessment, National Ministry of Education and Sports (SEE-MEDN)].

Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación. (2018b). *Aprender 2017: Informe de resultados, Buenos Aires, 6to año de secundaria*. Buenos Aires, Argentina: Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación (SEE-MEDN) [*Learn 2017: Report of results, Buenos Aires, 6th year of secondary school*. Buenos Aires, Argentina: Secretariat of Educational Assessment, National Ministry of Education and Sports (SEE-MEDN)].

Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación. (2018c). *Aprender 2017: Informe de resultados, Salta, 5to año de secundaria*. Buenos Aires, Argentina: Secretaría de Evaluación Educativa, Ministerio de Educación y Deportes de la Nación (SEE-MEDN) [*Learn 2017: Report of results, Salta, 5th year of secondary school*. Buenos Aires, Argentina: Secretariat of Educational Assessment, National Ministry of Education and Sports (SEE-MEDN)].

Sriram, R. (2014). Rethinking intelligence: The role of mindset in promoting success for academically high-risk students. *Journal of College Student Retention*, *15*(4), 515–536.

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J., Crosnoe, R., Muller, C., . . . Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature Human Behaviour*, *573*, 364–369.

Yeager, D. S., Johnson, R., Spitzer, B., Trzesniewski, K., Powers, J., & Dweck, C. S. (2014). The far-reaching effects of believing people can change: Implicit theories of personality shape stress, health, and achievement during adolescence. *Journal of Personality and Social Psychology*, *106*, 867–884.

Yeager, D. S., Miu, A., Powers, J., & Dweck, C. S. (2013). Implicit theories of personality and attributions of hostile intent: A meta-analysis, an experiment, and a longitudinal intervention. *Child Development*, *84*, 1651–1667.

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Scheinder, 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.

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

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

Yen, W. M., & Fitzpatrick, A. R. (2006). Item response theory. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.). American Council on Education and Praeger Publishers.

Author

ALEJANDRO J. GANIMIAN is an assistant professor of applied psychology and economics in the Steinhardt School of Culture, Education, and Human Development at New York University. His research interests are the development and measurement of academic and socioemotional skills in developing countries, from pre-school to secondary school. Specifically, he studies how to improve school management and classroom instruction to address the needs of the shifting and diverse student populations in these settings.

Manuscript received June 20, 2019
First revision received March 23, 2020
Second revision received June 5, 2020
Accepted June 9, 2020