CREB Working Paper No. 01-21 (Unedited copy)

The Role of Non-Cognitive Skills in Improving Academic Performance: Evidence from a field experiment in Pakistan

Maryiam Haroon
Farah Said
Mahniya Zafar

Centre for Research in Economics and Business
Lahore School of Economics

Centre for Research in Economics and Business (CREB)

Naved Hamid
Director CREB

CREB Advisory Board

Lahore School of Economics

Intersection Main Boulevard Phase VI, DHA and Burki Road
Lahore 53200, Pakistan
Tel: 042-36561230; 042-36560936
Email: creb@lahoreschool.edu.pk

Approval Date: August 2021

Lahore School of Economics
Intersection of Main Boulevard, Phase VI, DHA, and Burki Road
Lahore 53200, Pakistan
Tel.: +92 42 3656 1230
www.creb.org.pk
creb@lahoreschool.edu.pk

Price: PRs100

Preface

The Centre for Research in Economics and Business (CREB) was established in 2007 to conduct policy-oriented research with a rigorous academic perspective on key development issues facing Pakistan. In addition, CREB (i) facilitates and coordinates research by faculty at the Lahore School of Economics, (ii) hosts visiting international scholars undertaking research on Pakistan, and (iii) administers the Lahore School's postgraduate program leading to the MPhil and PhD degrees.

An important goal of CREB is to promote public debate on policy issues through conferences, seminars, and publications. In this connection, CREB organizes the Lahore School's Annual Conference on the Management of the Pakistan Economy, the proceedings of which are published in a special issue of the Lahore Journal of Economics.

The CREB Working Paper Series was initiated in 2008 to bring to a wider audience the research being carried out at the Centre. It is hoped that these papers will promote discussion on the subject and contribute to a better understanding of economic and business processes and development issues in Pakistan. Comments and feedback on these papers are welcome.

Since the second half of 2018 we have had issues with our regular editing services, as a result of which there has been a growing backlog of working papers that had been approved by the editorial committee. To avoid further delays in dissemination of the ongoing research, we decided to publish approved but unedited working papers online. Working paper No 03-18, December 2018 was the first such paper.

The Role of Non-Cognitive Skills in Improving Academic Performance: Evidence from a field experiment in Pakistan.*

Maryiam Haroon,† Farah Said‡ & Mahniya Zafar§

July 29, 2021

Abstract

We conduct a randomized control trial to investigate the effect of a soft-skills intervention that highlights the role of effort and perseverance in achieving goals, the mindset, and academic performance of college students. We have a sample of 366 undergraduate students from women-only and public colleges in a major metropolitan city in Pakistan. We find that a brief discussion stressing sustained effort and a constructive interpretation of failures improves the likelihood of students showing greater willingness to set strategic, measurable, attainable, results-oriented, and time-bound (SMART) goals immediately after the intervention. We find no significant test scores two and nine months later. However, we find increased grit among treated students one year after the intervention, with indications of even larger effects for students from disadvantaged backgrounds. Our findings suggest that soft touch interventions may impact cognitive skills, but that can be insufficient for improving academic achievement.

*The authors thank Aiman Farrukh, Saniya Jilani, Fizza Rahman and Maria Syed for outstanding research assistance. Financial support from the Lahore School of Economics Research Development Fund for research is gratefully acknowledged. The trial is registered at the AEA RCT registry on 24 May 2019 (Available at (https://doi.org/10.1257/rct.4233-2.1) and received IRB approval from the Lahore School of Economics Research Ethics Review Committee on March 15 2019 (File no. RERC-032019-01)

†Lahore School of Economics. Corresponding author.

‡Lahore School of Economics

§Lahore School of Economics

1 Introduction

Improving student learning and well-being has long been an important concern for development policy. An extensive literature on improving educational outcomes has focused on improving academic enrolment and documents positive yet largely modest effects of supply-side interventions and those that incentivise students and parents via financial or nutritional incentives.[1] More recently, a growing body of literature finds non-cognitive skills, such as grit, perseverance, and self-control, to be important predictors of academic performance and later-life, job market outcomes.[2] Grit, in particular, can be a crucial determinant of educational attainment and retention rate (Duckworth et al., 2007; Duckworth, 2009). Similarly, individuals who believe effort is imperative for success, and that skills and intelligence can be developed through persistent effort, are more likely to set ambitious goals and persevere in the event of a failure or setback (Dweck, 2006). Further, evidence suggests that non-cognitive skills among children are malleable and can be enhanced through motivation, provision of performance feedback and self-regulation exercises (Alan et al. (2019); Durlak et al. (2011); Kautz et al. (2014)).

In this study, we use a randomized control trial to investigate the effect of an intervention that aims to foster grit and improve learning outcomes. We borrow from the literature on supportive psychology and develop a soft-touch, easily scalable, low-cost intervention using visual aids and one-to-one discussions with treated students (Alan et al. (2019); Blackwell et al. (2007); Polley (2018)). The treatment intervention is randomised at the individual level. It is composed of discussions with the student on the role of effort in enhancing skills and achieving goals, the importance of a constructive interpretation of failures, and perseverance and sustained effort to achieve personal goals. We conduct a 'placebo' discussion with the control group, comprising scientific facts about the human brain. Students participate in a goal-setting activity at the end of treatment and placebo discussions, where they are introduced to the concept of "SMART" - Specific, Measurable, Ambitious, Realistic, Time-bound goals. In addition to testing the impact of treatment on year-end and exams conducted nine months after our intervention test scores, we also test if the treatment intervention affects decisions in the goal-setting exercise.

We find treated respondents display greater willingness to modify their goals into SMART goals immediately after the mindset discussions – by approximately 10%, after controlling for college and interviewer characteristics. The treatment leads to a significant increase in grit levels, i.e., one year after the intervention is first delivered, grit among treated students is 4.0 - 4.2 % higher than among the placebo group.

[1] See McEwan (2015); Glewwe and Kremer (2006); Glewwe and Miguel (2008) provide a detailed review.

[2] See for instance, Humphries and Fabian (2017); Alan and Ertac (2015); Kautz et al. (2014); Lindqvist and Vestman (2011); Roberts et al. (2007); Heckman et al. (2006); Bettinger et al. (2018); Alan et al. (2019); Paunesku et al. (2015); Heckman and Rubinstein (2001)

However, the treatment does not translate into greater commitment to academic goals set during the goal setting exercise and does not improve average test scores measured two and nine months after the intervention.

An analysis of results by student characteristics reveals important insights: first, the intervention has been more effective in improving grit among students that were disadvantaged or ranked low in performance at the baseline – grit increases among students who, at baseline, were less likely to rank academic career goals as important, for those who score lower on cognitive Raven's matrices and for less competitive students. Second, we also find that grit substantially improves for treated students in their first year of study, but it fails to translate into an improvement in performance. Treated freshmen students are, in fact, less likely to modify their results to SMART goals increase in these students' test scores. They also score lower on exams at 2 months, though this effect is short-lived and disappears at 9 months. Our results indicate an increase in grit, without subsequent improvement in plans to follow through, fails to improve performance. Finally, our results support the role of parental education and involvement. Treated students with educated mothers score higher on growth mindset and self-efficacy scales, and commitment to self-reported academic goals is higher among students who discuss their homework and assignments with their parents.[3]

We contribute to the literature in four key ways. First, our study adds to the growing literature that tests the effect of mindset interventions on non-cognitive skills, such as grit and self-efficacy, and whether they have the capacity for impacting academic performance.[4] Our intervention is directly comparable to the mindset intervention conducted with elementary school students in Istanbul (Alan et al., 2019) and with secondary school students in Peru (Outes-León et al., 2020), that in addition to improvement in psychological mindset, also show significant improvement in test scores.

Second, we also contribute to a small group of studies investigating the potential role of mindset interventions in developing world contexts (Ganimian, 2020; Outes-León et al., 2020). Students in our sample are women belonging to low-income households in Punjab, Pakistan.[5]. Literature suggests that school, teacher, and household inputs may be insufficient in resource-deprived and poor contexts, resulting in poor academic performance and high dropout rates (Dillon et al., 2017; Banerjee et al., 2010). Thus, inter-

[3] These results are reminiscent of descriptive evidence that concludes that students' commitment to goals are correlated with parents' goals for the academic performance of their children Zimmerman (1992).

[4] See for instance, Dweck (2006); Yeager and Dweck (2012)). Alan et al. (2019); Ganimian (2020); Islam et al. (2020); Dobronyi and Petronijevic (2019); Delavande (2019) test interventions designed to affect the mindset and goal setting behaviour to impact academic performance. For a detailed review, see (Paunesku et al., 2015; Yeager et al., 2016, 2018; Sriram, 2014; Yeager et al., 2014; Bettinger et al., 2018).

[5] Average monthly income in our sample is approximately PKR 36,000 (USD 225). The average household income in Punjab, Pakistan, in 2018 was approximately PKR 43,000 (USD 269) (Pakistan Bureau of Statistics, 2019)

ventions that aim to improve student mindset and self-belief have the potential to be particularly effective in these contexts. To the best of our knowledge, this is the first study that only focuses on women from low-income backgrounds and investigates a soft-touch, cost-effective intervention that can be easily scaled up and be used in developing country contexts.

Third, literature has largely considered non-cognitive skills immutable in older age and concentrated on younger, school-going children (Alan et al., 2019; Paunesku et al., 2015; Yeager et al., 2016, 2018). Two other studies have attempted to influence mindset among older secondary school students. Outes-León et al. (2020) test a mindset intervention in secondary schools in Peru and find positive impacts on student aspirations and test scores. However, Ganimian (2020) attempt an at-scale intervention (among more than 200 schools) and generally find no effect on any dimension of individual psychological outcomes or test scores (Bettinger et al., 2018; Polley, 2018). Finally, we test if these characteristics can be influenced among young adults enrolled in colleges and find that it is insufficient to improve average test scores while having a long-term change in mindset.

Finally, we have rich data that allows us to test how the effect of the intervention varies with student and household characteristics. We find, for instance, that student year of study, while not investigated extensively in literature (Alan et al., 2019; Islam et al., 2020), is an important dimension along which effects on grit, test scores, and goal outcomes vary.[6] We are also able to confirm findings from descriptive studies that support the role of educated mothers in providing a supportive environment that fosters education aspirations and outcomes of young women.(Roy and Bhattacharya, 2018).

Pakistan is a relevant setting for exploring the role of non-cognitive skills in student performance and well-being. Enrollment rates and adults' years of schooling have increased dramatically in almost all developing countries, with enrolment rates over 90% in Africa and South Asia (including Pakistan) (UNESCO, 2009). However, despite a significant increase in enrollment, learning outcomes reveal poor student performance. Filmer (2018) find that children from low-income countries such as India, Ethiopia, Peru, and Pakistan, perform worse in the international test than 95% of the students from high-income countries. A similar situation exists in higher education in Pakistan. Despite a significant increase in enrollment rates at the undergraduate level in Punjab, pass out rates are substantially low, i.e., 50 percent in Punjab and 49 percent for bachelors student in Lahore, one of the largest metropolitan centers of the country, with only a small fraction of students able to secure first division.[7] Our sample of women in the first, second and third year of intermediate and undergraduate degree enrolled in 10 public women-only colleges in urban areas of

[6] A few studies attempt to explore heterogeneity by year of enrolment. Morisano et al. (2010) find that the students in the early years have the lower test score as compared to senior years. Oreopoulos (2018) finds no heterogeneity in effects of an online planning module and coaching on test scores and credit accumulation across years of student enrollment.

[7] For details, see http://www.pu.edu.pk/.

the district of Lahore in Punjab, Pakistan. Women account for more than half of the total undergraduate students enrolled in the Lahore division, yet only less than half of the female candidates secure passing marks (PDS, 2018).

The existing learning crisis perpetuates inequalities later in life, determining access to work and income levels later in life (World Bank, 2018). Non-cognitive skills, such as grit and goal setting, have been shown to be instrumental in influencing various adult outcomes like retention rates among the United States Military Academy cadets (Maddi et al., 2012); teacher retention and effectiveness in class (Robertson-Kraft and Duckworth, 2014); labor force participation, earnings, and wages of unskilled workers (Lindqvist and Vestman, 2011); and fewer career changes (Duckworth, 2009). Investment in the non-cognitive skills of women can have potentially long-lasting, inter-generational effects on individual well-being.

The rest of the paper is organized as follows. In Section 2, we present the experiment design and treatment, along with study context and implementation details. Description of our sample and our empirical model. Section 3 describes the data and Section 4 summarizes the empirical strategy. We test the effect of our mindset intervention in section 5. Section 6 discusses variation in the effect by baseline characteristics and section 7 concludes with a discussion on potential program scale-up.

2 Experiment design and implementation

2.1 Study context

Our sample resides in Lahore, the second-largest city of Pakistan and capital of its' most populous province, Punjab. Lahore has a population of 11 million people, and with one out of every ten women, aged 15 - 29 years, estimated to live in Lahore, the city is an important education hub for women in Punjab (Census, 2017). The district of Lahore has a total of 63 intermediate and degree colleges, out of which 39 are for women only (PDS, 2018). Tuition fees in public colleges – approximately USD 22 per annum) – is much lower than the annual tuition in private colleges, which can range anywhere from USD 4252 to USD 280.[8] Female students usually outnumber men at the undergraduate and graduate levels in these colleges. An estimated 61,000 students were enrolled in intermediate and degree colleges in Lahore, out of which 62 percent were women.[9]

[8] Using 160 PKR = 1 USD exchange rate as of April 2020.

[9] At the undergraduate level, female students accounts for 6.9% of the total eligible female population, which is higher than that the proportion for males of around 3% (PDS, 2017).

Our study focuses on intermediate and undergraduate students from women-only public degree colleges in Lahore. While enrolment in colleges in Lahore has been steadily increasing, the pass rate remains consistently low.[10] Only about half of the female students enrolled in undergraduate public degree colleges in Punjab completed their degree in 2019, with similar but even lower completion rates in the Lahore division.[11] The grades of third-year students in public degree colleges show a dismal picture of learning gaps as out of the total female candidates who appeared for exams, less than fifth, scored more than 60% in end-of-year board exams and a little under a third of the sample had scores between 45 - 60%.[12] The test score for first and second-year students also presents a similar picture where only 13 percent students are able to score 70% or higher, while nearly half of the candidates who appeared for the annual board examinations in Punjab scored 50 - 60% in 2017 (PDS, 2018).

The Government of Punjab has introduced some initiatives to increase enrolment and boost academic performance, e.g., the Punjab Education Endowment Fund (PEEF) was introduced in 2009 to provide merit scholarships for intermediate, graduation and master students.[13] Evidence suggests that while government initiatives may have led to moderate growth in enrollment levels, their role in improving learning outcomes and performance are unclear (Fiszbein and Schady, 2009). Within this context, we attempt to test the role that non-cognitive skills can play in student mental state and academic performance.

2.2 Intervention

We borrow from the literature on supportive psychology and propose a soft touch, low-cost intervention provided through visual aids, discussions, and flyers to develop non-cognitive skills (particularly grit and perseverance) to overcome constraints to academic performance for college students. Our intervention consists of guided discussions that highlight the human brain's plasticity against the notion of innate ability and stresses the role of effort in achieving goals. We stress a constructive interpretation of failures and attempt to motivate students to persevere in pursuing short and long-term goals. We reiterate our message by providing examples, such as how young children learn the language and basic skills and how

[10] For example, there was an increase of 7.94% in total female enrolment at intermediate and undergraduate level in colleges between 2015-16 and 2016-17 PDS (2017), and PDS (2018).

[11] In 2019, out of the total female students who appeared in BA from public degree colleges, 50 percent in Punjab and 49 percent in Lahore district were able to pass (for more details, see http://www.pu.edu.pk/). Out of the total female students who appeared in FA, 73 percent in Punjab and 69 percent in Lahore board were able to pass in 2017 PDS (2018).

[12] Retrieved from http://www.pu.edu.pk/.

[13] For more details regarding the PEEF scholarship, please see: https://www.peef.org.pk/.

everyone has the capacity to improve performance through sustained effort and practice.[14] In the placebo discussion, we discussed the basic structure of the brain and were adopted from (Blackwell et al., 2007).

After the treatment and placebo activities, all students in our sample participate in a goal-setting activity at baseline to understand how participants formulate their goals and push them in a direction to reach their goals (Dobronyi and Petronijevic, 2019; Morisano et al., 2010). Students were asked to outline their short and long-term academic and professional goals. They are then informed about 'SMART' - Specific, Measurable, Ambitious, Realistic, Time-bound, goals; to think about whether their goals were SMART and how they could be modified to satisfy all five dimensions of SMART goals.

Treatment and placebo discussions were conducted after a short baseline student interview. Treatment was randomized by the research team using the survey software at the individual (student) level. Intervention discussions were followed by the goal-setting exercise. Table 1 summarizes treatment assignment.

Treatment Status	SMART goal discussion	N
Treatment: Growth mindset discussion	YES	187
Placebo: Placebo discussion	YES	179

Table 1: **Assignment to treatment**

2.3 Implementation timeline and protocol

We ran the randomized field experiment with 366 students enrolled in undergraduate programs in 10 women-only public degree colleges of Lahore, Pakistan. Our sample consists of students enrolled in the first, second, and third year of college, i.e., not final-year students near graduation. Sample students were Arts and Social Sciences majors, specializing in English, Urdu, Arabic, Persian, Economics, Education, Psychology, Sociology, Business Administration, Management Studies, accounting within their respective degrees.

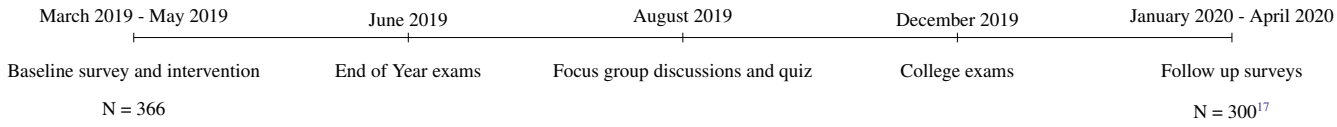
We recruited the study sample as follows: first, we requested the administration of our sample colleges for a list of students enrolled in the first, second, and third year. Next, we randomly draw out lists of students from each college, proportional to the size of the student body, and contact them for the baseline interview. All data was collected electronically. Finally, the survey software assigns each respondent to a treatment or

[14] The intervention was designed with the help of the materials available online at `mindsetkit.org` and `mindsetworks.com`, the two websites that make it easy for educators to implement growth mindset lessons and using the content and material from (Blackwell et al., 2007). The detailed content of treatment and placebo discussion is available on request.

placebo group with equal probability.[15] Treatment implementation is followed by the goal-setting exercise and short, post-intervention survey.

We collected detailed baseline data and implemented the treatment and placebo interventions via in-person interviews carried out by a team of trained enumerators between March and May 2019. In August 2019, we conducted focus group discussions with 10% randomly selected respondents and phone surveys with another 40% randomly selected treated and placebo students to record intervention recall. We conduct a follow-up survey later in January till April 2020. Finally, we collect administrative data on student performance in end-of-year, standardized board examinations held two months after our intervention, and internal college exams nine months after our intervention.[16] The timeline of our field activity is given in Figure 1:

Figure 1: Timeline of Field Activity



Follow-up surveys were interrupted when all higher education institutions were closed by the provincial government on March 13, 2020, in response to the global COVID-19. As a result, the follow-up data was collected using a mix of in-person and phone interviews. Out of 300 interviews at follow-up, 226 were conducted in-person before the lock-down, and 74 interviews were conducted on the phone between 14 March - 10 April 2020.

3 Data

3.1 Description of the sample and treatment balance

Our sample consists of 366 students from 10 selected female-only public degree colleges in urban areas of Lahore, Pakistan. All students in our sample are in their first, second, or third year of college, i.e., not

[15] Treatment assignment is done at the individual level using the random number generator in the survey software, SurveyCTO (www.surveycto.com).

[16] We planned to collect information on standardized board exams one year after the implementation. However, scheduled board exams were not administered due to country-wide closure in March 2019 due to the COVID-19 pandemic.

[17] Of the surveys successfully completed at followup; 226 were conducted in person during January 2020 - March 2020 and 74 over the phone from March 14, 2020 to April 10, 2020. Appendix section A6 provides details on the regulatory framework surrounding the lock-down.

final year students who will only be in college for another two to three months after baseline. Sample characteristics are summarized in Table 2.

The average respondent is 19 years of age, from households where parental education and labor force participation is unequal - fathers are both more educated and more likely to be working than the mothers of the students in this sample. Approximately two out of three respondents come from families where the primary earner, the father, is salaried; one out of three come from families where the father has his own business and is self-employed. The average monthly income for the household was not reported in 108 instances; for those who respond, the average income per month is approximately PKR 36,000 (USD 225) and the median of PKR 25,000 (USD 156).

Sample students are able to answer 4 out of 10 Raven's matrix questions correctly, while 72% of the students report having graduated in the 'first division' (more than 60% of total marks) from their previous degree. 91% plan to continue education after completing their current degree, and one out of every five reports they would like to join the labor force after graduating (from their terminal degree). 92% of the students wish to pursue a paid job. Students spend 3 hours every day studying at home (after college) in an average week. A relatively higher proportion of the day is spent on housework, approximately 6 hours daily in an average week. A high percentage of the students (71%) report discussing college assignments (homework, as it is locally known) and their education goals (93%) with their parents.

We test treatment balance by estimating equation 7 for a set of baseline covariates of interest denoted by y_{ki} and indexed by $k \in (1, ..., K)$.[18]

$$y_i = \beta_0 + \beta_1 . T_i + \epsilon_i \tag{1}$$

Where T_i as a dummy variable equal to 1 if a student (i) is assigned to treatment, and 0 for assignment to control. In column (5) of Table 2, we report the $p - value$ from a regression of the covariate on treatment status. Test of joint significance show that treatment is unrelated to respondent level characteristics measured at baseline ($F - stat(p - value)$ = 0.11) and treatment and placebo samples are balanced on all but three characteristics. The difference in average ages between the treated (19.2 years) and placebo group (19.5 years) is statistically significant at the 10% level; it is economically small. Furthermore, average values for the proportion of first-year students in control (31%) and treatment (40%) are statistically different at the 10% level. Finally, 96% of control group individuals wished to pursue a paid job compared to 89% of those in the treatment group, which is significant at the 5% level.

We find good recall of treatment discussions in focus group discussions conducted through phone surveys

[18] As specified in the PreAnalysis Plan (Haroon et al., 2020).

Table 2: Full sample description and balance at baseline

	N	Mean	Median	S.Dev.	$p-value$
	(1)	(2)	(3)	(4)	(5)
Age (years)	366	19.37	19.00	1.62	0.07*
Dummy: Enrolled in first year	366	0.36	0.00	0.48	0.10*
Dummy: Single	366	0.96	1.00	0.20	0.73
Father's education (years)	351	8.43	10.00	4.80	0.31
Dummy: Father is a business owner	366	0.25	0.00	0.44	0.91
Dummy: Father is a salaried worker	366	0.57	1.00	0.50	0.79
Mother's education (years)	362	6.64	8.00	5.15	0.58
Dummy: Mother is a business owner	366	0.02	0.00	0.13	0.44
Dummy: Mother is a salaried worker	366	0.05	0.00	0.23	0.41
Average monthly household income (000's)	258	35.56	25.00	33.09	0.40
Index: Household assets	366	0.00	0.30	1.39	0.57
Risk preferences (higher is more risk averse)	366	4.15	4.00	1.59	0.25
Ravens test score (out of 10)	366	4.08	4.00	2.11	0.93
Scale: Competitiveness (out of 75)	366	57.62	58.00	8.33	0.60
Dummy: Continue education after graduating	366	0.91	1.00	0.28	0.81
Dummy: Passed last degree in first division	363	0.72	1.00	0.45	0.29
Satisfaction with academic performance	366	2.37	2.00	0.91	0.95
Daily hours studying at home in average week	366	2.99	3.00	2.07	0.51
Daily hours doing household chores in average week	366	6.28	7.00	5.99	0.31
Dummy: Discusses homework with parents	366	0.71	1.00	0.46	0.45
Dummy: Discusses goals with parents	366	0.93	1.00	0.27	0.11
Dummy: Pursue paid job	366	0.92	1.00	0.27	0.03**
Dummy: Have a role model	366	0.33	0.00	0.47	0.34
P-value of F-statistic					0.11

Note: Column (5) shows $p-values$ from the balance test specified in equation 1. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. $***p<0.01, **p<0.05, *p<0.1$.

nearly three months after the initial intervention implementation.[19] In quick questions designed to test knowledge of treatment and placebo discussions, we found recall in treatment to be better than in the placebo group – Students in the placebo group correctly answered 2.9 out of 5 quiz questions correctly, compared to 4.7 out of 5 questions asked of the treated group (p = 0.000).

At follow-up, we were able to contact 300 students successfully. Table A1 in the appendix shows the sample is qualitatively similar, except on age and desire to pursue a paid job, as before. Specifically, treated students in this balanced sample (of 300) are younger when compared to the placebo sample; in addition, they are less likely to say at baseline that they plan to apply for a job in the future. The overall test of joint significance shows the sample is well balanced by treatment status ($p - value$ = 0.14). We also test for differences in characteristics of respondents interviewed in the person and those interviewed via phone at the end line and find no significant overall difference between the two samples ($p - value$ = 0.86).[20]

3.2 Attrition

We were unable to interview 66 students from the baseline sample at follow-up – 44% (29) of the attriters were from the treated group and 56% (37) from the placebo group. We test for differential attrition between the treatment and placebo group by coding a dummy variable for whether individual i attrited for the follow-up survey and then test if attrition is significantly related to baseline covariates of interest and to treatment status.[21]

As shown in Table A3, attrition does not vary by treatment status. Attriters vary by level of competitiveness and degree of satisfaction with their academic performance. In section 5.1, we show our results are robust to attrition.

[19] We conduct focus group discussions with 10% of the baseline sample (n=34) and phone surveys with an additional 40% (n = 148) of the sample. Both involved short questions regarding treatment and placebo discussion content (i.e., we have these measures for approximately 50% of the baseline sample). In the focus group discussions, we also ask detailed questions about components of the intervention they liked best and if they would like us to share any other information useful for them in the future, e.g., scholarships provided by the government.

[20] In Table A2, we show treatment balance in the balanced panel that participated in face-to-face surveys in both rounds. In section 5.1, we show our results are robust to restricting our sample to only the 226 who participated in face-to-face surveys.

[21] A complete list of baseline covariates for which we test balance (between treatment and control) is specified our pre analysis plan (Haroon et al., 2020).

4 Empirical strategy

4.1 Outcomes

We estimated the impact of our mindset treatment on two main sets of outcomes.[22]. Mindset interventions have shown in the literature to be highly predictive of psychometric scales measuring non-cognitive skills of the respondent and may affect academic performance (See, for instance, Alan et al. (2019)). Scales measuring students' non-cognitive skills - grit, growth (as opposed to fixed) mindset and self-efficacy; and performance in annual examination.

We estimated the effect of the mindset treatment on the following **primary outcomes** at two points in time: (i) once immediately after the intervention to test if the intervention had the intended effect on mindset; and (ii) we also test if the intervention has a longer-term effect on mindset one year later:

Grit: We measure grit post-intervention and at follow-up using a 12-item scale developed by Duckworth et al. (2007)[23]

Growth mindset: In both rounds of the survey, we also collect information on fixed versus growth mindset using a 15-item scale (Blackwell et al., 2007).

Self-efficacy: We measure self-reported respondent 'self-efficacy' using the 10-item scale by (Schwarzer and Jerusalem, 1995).

We correct for multiple hypothesis testing in two ways: i) we create an Anderson index using variables (i) - (iii) for *Psychological Mindset*; and (ii) by calculating the sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008).

Test score percentage: Approximately a few months after the intervention was first implemented, all sample students also take end-of-year, standardized exams. We collect administrative data on test scores from sample colleges. Since exams are administered by – and maximum that can be scored differ by – the year of enrollment, we calculate marks scored by each respondent in the exam and then calculate *test score percentage*, that is, the marks scored by each respondent as a percentage of the total marks for that

[22] The full set of outcomes is specified in Haroon et al. (2020) and is available in the Online Appendix

[23] One of the criticisms of using the self-reported measures of grit and similar other measures could be that it is difficult to capture them on a self-reported scale and in particular it could be more difficult in a high-stakes setting where the respondent can pretend to get the benefit. One could think of using more advanced or gameable measures of grit which could be similar to the effort task developed by (Alan et al., 2019).

year. Board exams next year were cancelled due to nation-wide COVID-19 lockdown. We collect data on non-standardized exams administered by colleges before the lockdown, approximately nine months after our intervention.

For the **secondary set of outcomes**, we investigate the immediate impact of the intervention on behavior. We do this by utilizing the goal-setting exercise implemented after the treatment and placebo discussion in which all students participated. We measure if students in the treated and control group have behaved differently in the type of goals being set.

Importance of academic goal: We ask respondents to rank the importance of their academic goals to them on a Likert scale ranging from 1 (not important at all) to 10 (extremely important agree).

Achieve academic goal: We ask respondents to rank how confident they feel about being able to achieve their academic goal, answered on a Likert scale ranging from 1 (not at all likely) to 10 (extremely likely).

Commitment to academic goal: This outcome measures respondent's self-reported commitment to achieving an academic goal, answered on a Likert scale ranging from 0 (not committed at all) to 10 (fully committed).

Willing to modify goals: We define a secondary outcome variable which is equal to 1 if the respondent sets a SMART goal after the goal-setting discussion.[24]

As before, we will calculate indices for the family of goal setting measuring ('Goal Index') and also report sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008) for each family of outcomes.[25]

[23] There are two governing bodies that are responsible for conducting the test for first, second-and third-year students across public colleges in the Lahore division. The Board of Intermediate and Secondary Education (BISE) Lahore is the governing body which under the Punjab University Act (Amendment) Ordinance 1954 is responsible for conducting standardized exams for first- and second-year students all across the Lahore division. On the other hand, the University of Punjab is responsible for conducting standardized exams for third-year students for affiliated public degree colleges all across Punjab. The examinations are standardized for all students and take place at the same time in the Lahore division (for first- and second-year students) and Punjab (for third-year students).

[23] Out of the total sample, we could not obtain information on 34 students. We checked if the students with missing score data were different from other students and if our treatment was correlated with attrition after nine months. The balance tests show that attrition was not related to treatment. However, it was positively related to if a father is a business owner and students who are most likely to continue education after graduation. In contrast, first-year students are less likely to drop out.

[24] In all cases, enumerators defined when the goal set was not Specific, Measurable, Ambitious, Realistic or Time-bound and then ask respondents if they wanted to modify their goal to one that could be considered to be all five.

[25] In face-to-face follow-up surveys, we also collect measures of happiness (Lyubomirsky and Lepper, 1999), self-discipline (Hagger et al., 2018), self-esteem (Rosenberg, 1965) and feelings of depression (Kroenke et al., 2001). These measures are only at follow-up and with individuals interviewed in person. We do not collect this information at baseline; hence, we have not included these results in our main analysis. However, results are available in Online Appendix, as specified in Haroon et al. (2020).

4.2 Estimation specification

We test for the effect of our intervention on outcome variables specified in 4.1. We test for both immediate effects measured post-intervention effects at baseline and longer-term impact on test scores after two and nine months and survey measures one year later.

For each outcome, we run the following basic specification:

$$y_{it} = \beta_0 + \beta_1 Treated_i + \beta_2 Post_{it} + \beta_3 Treated_i * Post_{it} + \epsilon_{it} \tag{2}$$

where y_{it} is an outcome variable, $Treated_i$ is a dummy variable capturing exposure to treatment (measured at the time of baseline activities, via a survey immediately after the intervention was implemented), $Post_t$ is a dummy variable for the follow-up time period. $Treated_i * Post_t$ is the average treatment effect in the longer run, i.e., approximately one year later for grit, growth mindset, and self-efficacy scales; and the mindset index created out of these measures.[26]

We test if our intervention has the intended immediate effect on the growth mindset and goals set. Our primary research question is to test if exposure to the treatment *i.e.* is able to change test performance in the short and medium term, and grit and mindset in the longer term. For outcomes for which we have data only after the intervention, we run a simplified specification: We measure if our treatment changes student performance in end-of-year examinations, 2-3 months after the intervention was implemented, college examinations in December, nine-month after our discussions, and secondary outcomes of interest – whether the intervention impacted respondent's self-reported likelihood of achieving academic goals, their commitment to meet academic goals, how important these goals are to them, and their willingness to set SMART goals. For these outcomes for which we have data only after the intervention. The test scores are measured after the intervention- two, and then nine months after the baseline data collection, and the secondary outcomes measured only at post-intervention - we run a simplified specification:

$$y_{it} = \beta_0 + \beta_1 Treated_i + \epsilon_{it} \tag{3}$$

where y_{it} is an outcome variable, $Treated_i$ is a dummy variable capturing exposure to treatment, as before. We report robust standard errors. In results discussed in section 5, we show estimates from equation 9 and

[26] Given our design involves individual-level randomisation and multiple rounds of data collection; we take guidance from Abadie et al. (2017) and cluster errors at the level of randomisation. In the analysis specified in our pre-analysis plan (Haroon et al., 2020), we show that the results are qualitatively similar (in fact more significant) when we cluster errors at the college level.

3. We also show estimates after controlling for college and enumerator fixed effects. Finally, in Appendix, we also show results from repeating all main estimations using 'post-double-selection' with LASSO (see Belloni et al. (2014a) and Belloni et al. (2014b)), where we use all of the baseline covariates specified in Haroon et al. (2020).

Finally, we will show robustness of results to attrition in two main ways: (i) We repeat all main estimations using 'post-double-selection' with LASSO (Belloni et al., 2014a,b) and (ii) we re-estimate our main regressions using only the sample of respondents who were interviewed face-to-face at followup.

5 Results

We measure the impact of mindset interventions on primary and secondary outcomes. We collect post-intervention data from 366 students on levels of grit, fixed vs. growth mindset, and self-efficacy at the time of baseline. We also collect this data from 300 students at the follow-up, nearly one year after the intervention was first implemented. We report both immediate (post-intervention) and longer-term impacts in Table A13, from simple estimations of equation 9. Given individual randomisation and the observed balance in the sample (as discussed in section OA7.1), Table 4 provides results from estimations without controls and those with college and enumerator fixed effects. In the Appendix, we also show the robustness of our results with controls selected using 'post-double-selection' with LASSO (Belloni et al., 2014a,b). Finally, we correct for multiple hypothesis testing by calculating the sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008), reported in square brackets in Table 4. We also create a 'Mindset index' for outcomes reported in columns (1) - (6).

We test if the treatment discussion led to treated respondents scoring higher in the standardized end-of-year exams two months later and college-level exams nine months later than their counterparts in the control group. Results are shown in Table 3. For ease of interpretation, we calculate and report effect on percentage scored out of the total (full) marks allowed for that year of study.[27] We find that students score 50% on average (Table 3, columns 1 and 2). Treated students score approximately 2.3 - 2.5 percentage points higher two months after the intervention; however, this improvement is small and is statistically insignificant (presented in column 1 and 2). We also find that treated students score approximately 2.9 percentage points lower nine months after the intervention; however, this effect is small and is statistically insignificant and not robust to inclusion of fixed effects (presented in Table 3 column 3 and 4). Note also

[27] The Total marks depend on the year the student is enrolled in. For instance, students are given marks out of 550 for those enrolled in the first year, out of 1100 for those enrolled in the second year, and out of 400 for those enrolled in the 3rd and final year.

that these exams were not standardized but administered by each college separately.

There is mixed evidence found in the literature regarding exposure to growth mindset treatment on test scores. Contrary to our findings, some of the existing literature documents the positive impact of growth mindset treatment on test scores in the case of children and adolescents in Norway (Bettinger et al., 2018); for secondary school students in Bangladesh (Polley, 2018); for 4^{th} grade students in Turkey (Alan et al., 2019). However, there is some recent evidence that document insignificant effect of growth mindset intervention on test score for the case of 12^{th} grade students in Argentina (Ganimian, 2020); for secondary school students in Tanzania (Islam et al., 2020) and college students in Canada (Dobronyi and Petronijevic, 2019).

Table 3: Impact of treatment on primary outcomes: exam performance two and nine months after treatment implementation

Time after baseline	2 months		9 months	
	Test percent	Test percent	Test percent	Test percent
	(1)	(2)	(3)	(4)
Treated	0.023	0.025	-0.293	0.068
	(0.030)	(0.033)	(2.516)	(2.307)
Mean	0.496	0.496	0.403	0.403
N	330	330	332	332
R^2	0.002	0.135	0.000	0.188
Controls	No	No	No	No
FE	No	Yes	No	Yes

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects. 'Test percent' are the marks scored in end-of-year and after nine months of our intervention examination as a percentage of the maximum (full) marks allowed for that enrollment year. These were measured 2 and 9 months after the baseline activities and are present after our intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We find that the intervention leads to a small and moderately significant decrease in the 'grit scale' post-intervention at baseline (Table 4, columns 1 and 2). That is, respondents in our treatment group first record lower levels of grit than the control group immediately after the discussion on mindset at baseline. We also see a small increase in the average levels of the 'growth mindset' among treated respondents at the time of baseline; however, this increase is not large or robust to the inclusion of college and enumerator fixed effects (and as shown in Appendix Table A14, to the inclusion of controls selected by PDS Lasso).

We find that the intervention has a significant and positive impact on grit levels in the longer run, i.e., grit levels are 4.0 - 4.2 % higher (from average scores for the control sample over the same period) among treated students one year later. This effect is robust to the inclusion of college and enumerator fixed effects

Table 4: Long term impact of treatment on primary outcomes

	Grit	Grit	Growth mindset	Growth mindset	Self efficacy	Self efficacy	Mindset index	Mindset index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-1.027	-1.112	1.192	0.987	0.058	0.163	0.029	0.011
	(0.654)	(0.654)*	(0.692)*	(0.678)	(0.546)	(0.536)	(0.107)	(0.109)
	[0.211]	[0.281]	[0.211]	[0.281]	[0.440]	[0.340]		
Post	-0.577	-0.670	-2.750	-2.798	0.074	0.104	-0.323	-0.333
	(0.528)	(0.521)	(0.714)***	(0.710)***	(0.472)	(0.479)	(0.090)***	(0.092)***
	[0.378]	[0.249]	[0.001]***	[0.001]***	[0.698]	[0.426]		
Treated*Post	1.648	1.727	0.244	0.303	-0.138	-0.221	0.165	0.170
	(0.747)**	(0.744)**	(0.976)	(0.977)	(0.684)	(0.687)	(0.133)	(0.134)
	[0.089]*	[0.089]*	[1.000]	[1.000]	[1.000]	[1.000]		
Mean	41.125	41.125	48.660	48.660	32.938	32.938	-0.143	-0.143
N	666	666	666	666	666	666	666	666
R^2	0.005	0.034	0.043	0.092	0.000	0.048	0.018	0.050
Controls	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. T*Post is the average treatment effect of our growth mindset intervention. 'Mean' is the average value of the outcome for the control group. N refers to the sample size. Values in squared brackets represent sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

and the inclusion of controls. Compared to other estimates in the literature, our long-term impact on grit is large. We note that (Alan et al., 2019) concludes that being randomly assigned to growth mindset treatment raises self-reported grit between the range of 0.28 - 0.35% for 4^{th} grade students from elementary school in Istanbul. While the intervention has successfully increased grit among respondents, we do not see similar impacts on growth mindset or self-efficacy levels.

On average, though the treatment was successful in increasing levels of grit on average, the higher levels of grit did not translate into improved average test score performance. We also find that the treatment had no effect on respondent levels of self-efficacy. In Appendix Table A9, we show that the treatment did not change self-reported measures of the importance of academic goals for treated respondents, their self-reported likelihood, or commitment to reaching academic goals. However, once we control for college and enumerator fixed effects, treated respondents are 5.5% more likely to modify their goals into SMART

goals than control respondents.[28] Our results are in line with existing literature which documents that students exposed to growth mindset interventions are more likely to set challenging goals (Alan et al., 2019).

To summarize, we find average treatment effects on students' grit level over the long run, but we do not find that this translates to other mindset or student test performance measures. These results are robust to the inclusion of controls and restricting the sample to a balanced panel that participates in face-to-face interviews in both rounds of the data collection. We discuss heterogeneity in these results for specific sub-samples of the study sample next, in section 6.

5.1 Robustness of results

5.1.1 Treatment and attrition balance

Our sample is balanced well on several characteristics measured at baseline (see section OA7.1). In addition, we successfully contact 82% of the respondents interviewed at baseline for the follow-up surveys one year later, and attrition is unrelated to treatment status. However, we re-estimate main regressions with controls selected using 'post-double-selection' with LASSO (Belloni et al., 2014a,b). Tables A10, A11, and A12 provide the results. Results are similar qualitatively and in statistical significance - the treatment effects on grit are positive and significant, while average effects on scores, other primary and secondary outcomes remain insignificant.[29]

5.1.2 Sensitivity to questionnaire administration

It is possible that the measurement of the mindset outcomes - grit, growth mindset, self-efficacy, is sensitive to the mode in which the survey was administered - phone or face-to-face. Next, we re-estimate the main equation 9 for primary outcomes on only the sub-sample of students who participated in the face-to-face

[28] At followup, we also collect self-reported information on self-discipline (Hagger et al., 2018), self esteem (Rosenberg, 1965), happiness (Lyubomirsky and Lepper, 1999) and depression (Kroenke et al., 2001) using validated scales. We also measure their willingness to learn a new local or foreign language. We do not collect these measures at baseline. However, we find no difference among treated and control groups on these dimensions at follow-up. Results are available in the Online Appendix with analysis as specified in our pre-analysis plan (Haroon et al., 2020).

[29] In Online Appendix Tables OA1 and OA2 we also show upper and lower bounds of the estimates using (Lee, 2009) bounds. The bounding estimates 2.2% excess responsiveness in test score results and 6.1% excess observations in the primary outcomes. The lower bound for the longer-term effect on grit is positive but no longer significant.

survey at both baseline and follow-up. Results are shown in Appendix Table A13.[30] Results remain qualitatively similar - the longer run treatment effect on respondent grit is positive, albeit smaller in size than the results summarized in Table A13. No other outcome is significantly affected by treatment either immediately, post-intervention, nor in the longer run.

5.1.3 Treatment spillover

We implement mindset and placebo discussions with randomly selected students in each of the sample colleges to increase statistical power given the small sample size. Since both treated and placebo group students belong to the same college, it is possible for the treated students to have communicated the content of the intervention to their friends, some of whom may be in the placebo group. Potential spillover of information could mean that the treatment effects we observe are an underestimate of the true effect – with many placebo students having access to the same information as the treated effect, their responses to psychometric scales measuring primary outcomes may be similar. In this case, the treatment effects estimated in Tables 3 and 4 are an underestimate of the true effect of mindset treatment. In order to test for spillover, we replicate the strategy followed by (Banerjee et al., 2010). At baseline, we ask each participant, subject i, to identify at least two of their friends, subjects j, who they talk to most (whether in this college or elsewhere). Then, we use our data to match these referrals/ friends, subject j, based on our sample data's names and college identification numbers. To evaluate the spillover effects, we estimate the indirect impact of our growth mindset treatment on subject i on subject j's outcomes. We estimate the following specification:

$$y_{jt} = \beta_1 Treated_i + \beta_2 Post_{jt} + \beta_3 Treated_i * Post_{jt} + \epsilon_{jt} \tag{4}$$

where y_{jt} is an outcome variable of a referred friend j who is a network member of participant i, $Treated_i$ is a dummy variable capturing exposure to the treatment of participant i (measured at the time of baseline activities), $Post_t$ is a dummy variable for the follow-up period. $Treated_i * Post_t$ is the indirect average treatment effect of our growth mindset treatment of subject i on the outcome for subject j.

We evaluate the spillover impact on our primary outcomes: test scores and one year later for grit, growth mindset, and self-efficacy scales; and the mindset index created out of these measures. In case of a positive spillover of our growth mindset treatment, we would expect j outcomes to respond positively due to indirect exposure of treatment. In case of no spillovers, there should not be any effect of exposure

[30] Estimates for primary effects for this restricted sample using controls selected via post-double-selection with LASSO are reported in Appendix A14.

of participants i treatment on j outcome, i.e., we should not see any treatment effect in placebo students just because they are friends with treated students. In addition, we estimate specification as in 9 using a matched sample of participants. In case of positive spillover from the treated group to their friends in the placebo group, we should find no significant differences in key outcomes among the treated and placebo group, i.e., the treatment effects in Tables 4 and 3 disappear or change.

Both estimation strategies substantially reduce our sample and the statistical power to detect significant effects. In the case of null results, we may not say whether there is no true effect or that we are underpowered to detect them. Therefore for each key outcome, we also compute the minimum detectable effect (MDE) size and use it to compare our coefficients to results found in the literature (with larger samples). This is the ex-post effect size that would have been detected at a 5 percent significance level and 80 percent power for our sample size.

Out of the total sample of 366 students in our sample, we are able to match 240 participants who are friends of other sample participants, which allows us to measure spillovers and indirect treatment effects potentially. We could not administer a separate survey for friends other than those already included in our sample due to limited availability of funds. Out of these 240 matched friends, 121 were also directly exposed to treatment, i.e., were part of the treatment group, and 119 identified participants were part of the placebo group who did not receive treatment directly.

Spillovers effect among friends: We estimate equation 4 for a sample of matched participants to measure indirect treatment effects. The results are presented in Appendix A2, Tables A4 and A5. We do not find any spillover effect and indirect treatment on test scores which is positive and insignificant. In Appendix Table A5, we find insignificant indirect effect on grit (columns 1-2) and growth mindset of placebo friends (columns 3-4) compared to their treated friends. We find similarly insignificant for self-efficacy (columns 5-6) and mindset index (columns 7-8).

We report MDEs for all outcomes. For grit, the MDE is about a 7% increase from the mean grit score of the placebo group. These appear in line with the effect sizes in typical behavioral interventions. For instance, in a recent summary of behavioral interventions, Della Vigna and Linos (2020) report average MDEs ranging from 8 - 35%. Many of the studies reported in this review are with larger samples, indicating that the lack of significance on spillover estimates is not due to the small sample size. These results are similar to (Banerjee et al., 2010), who detect positive knowledge spillovers in terms of passing on factual information to their friends but no significant spillover on attitudes or behavior of friends.

Treatment effects in a sub sample of matched peers: Next, we also re-estimate our main regressions on a sub-sample of friends – in the case of spillovers, the difference in treatment effects between treated

and placebo friends should be small and insignificant. Results are presented in Appendix Table A6 and A7. We find similar results for all outcomes as presented in 4 for the full sample, albeit smaller and less significant. This indicates that even among a sub-sample of matched participants where spillover is likely to exist and effects may be underestimated, there is still a statistically significant direct treatment effect of our intervention on grit (positive and significant). This is in line with the results presented in Table A5, where the effect among placebo students with treated friends is small and statistically not different from zero.

6 Heterogeneity in effects

We then test for the heterogeneous treatment effect of our intervention on our primary and secondary outcomes for both short run - immediate and post-intervention effects and longer term effects using data from the followup surveys, and characteristics that have often been investigated in literature.[31]. A similar analysis to understand heterogeneity of results has been documented in the literature. For details see; (Polley, 2018); (Islam et al., 2020); (Roy and Bhattacharya, 2018); (Alan et al., 2019); (Oreopoulos, 2018). For each of the interactions, we cannot claim causality but they indicate interesting avenues for future investigations.

We estimate the following regressions to measure heterogeneous treatment effects for primary outcomes - grit, growth mindset and self-efficacy:

$$\begin{aligned} y_{it} =& \beta_1.Treated_i + \delta Post_{it} + \gamma H_{i0} + \beta_2 H_{i0} * Treated_i + \\ & \beta_3 Treated_i * Post_{it} + \beta_4 H_{i0} * Treated_i * Post_{it} + \epsilon_{it} \end{aligned} \tag{5}$$

H_{i0} is the variable on which heterogeneous treatment effects are tested. We test the significance of coefficients β_3 and β_4 with interaction terms $Treated_i * Post_{it}$ and $H_{i0} * Treated_i * Post_{it}$. The rest are as defined in equation 9.

$$y_{it} = \beta_1.Treated_i + \gamma H_{i0} + \beta_2 H_{i0} * Treated_i + \epsilon_{it} \tag{6}$$

H_{i0} is the variable on which heterogeneous treatment effects are tested. We test for the significance of β_2

[31] The variables used to test for heterogeneous treatment effects are specified in Haroon and Said (2019)

with the interaction term $T_i * H_{i0}$. The rest are as defined in equation 3.

We find interesting heterogeneity in the average effects discussed earlier. We discuss heterogeneous effects in outcomes of interest one by one.[32]

Exam performance: We do not observe substantial heterogeneity in effects for test scores for end-of-year examination. In fact, we see negative and large effects of the treatment on first year students - the scores are 11% lower among students who are in the first year. This effect is significant at the 5% level. The literature provides mixed evidence of the heterogeneous effect of treatment on test scores based upon student year as (Morisano et al., 2010) find more than half a standard deviation increase in grades for upper-year students at McGill University and (Schippers and Dawson, 2015) find goal-setting to significantly reduce inequalities in achievement if implemented early in student academic career whereas Oreopoulos (2018) document no heterogeneity in treatment effects across students from first years as compared to others. We also find negative effects on the test scores of treated students from households with higher levels of household assets, which can be used as a proxy for household wealth. Our findings are contrary to the literature which suggests that mindset interventions helped students belonging to the weaker socio-economic backgrounds (Islam et al., 2020). We find no evidence of heterogeneity in effects on test scores at nine months after the intervention (Table 6).

Grit Among the primary outcomes, we find that the effect on grit varies by household and individual characteristics. We see indications of the treatment increasing grit among disadvantaged students. Results are given in Table 7. Specifically, the intervention is effective in increasing grit at follow-up among treated students with uneducated mothers (column (2)) and among students who spend less time in productive activities - household chores or studying (column (3) and (9)), i.e., grit increases for students who were spending greater time in leisure or rest. In addition, we find the treatment was able to substantially increase grit among treated students who, at baseline, were less likely to rank career goals to be important (column (4), those who scored low on our measures of cognitive ability - Raven's progressive matrices (column (5)) and students who are less competitive students (column (6)). Finally, grit substantially improves for treated students in students in the first year of college (column (8)).

Growth mindset: We measure heterogeneity in results in Table 8. In contrast to what we see in the

[32] We explore heterogeneity by the following characteristics: 'HH assets' is an index from the number of household assets, 'Mother education' is the years of education of respondents' mother, 'Time in hh chores' and 'Time in the study' are the numbers of hours a day devoting to household chores and study, respectively, 'Career goal importance' is the self-reported importance given to having a career, 'Ravens score' is the respondents' score in Raven's progressive matrices, 'Competitiveness' is the respondents score in (Ryckman et al., 1996) scale, 'First division' is if the respondent secured highest (first) grade in last exams before joining college, 'First year' is a binary variable equal to 1 if the respondent is enrolled in the first year of college, and 'Discuss work' is a binary variable equal to 1 if the respondent discusses studies and assignments with parents.

case of the grit scale, we find that treatment led to a large increase in the growth mindset scale of treated students with educated mothers - the scale decreases for treated respondents with uneducated mothers by approximately 7.9%. (Delavande, 2019) show evidence of heterogeneous treatment effect for female first-year university students in the U.K. as compared to their male counterparts with parents working in high skilled jobs requiring higher education. (Roy and Bhattacharya, 2018) suggests that educated mothers foster educational aspiration in young women. We see no other heterogeneity in the average null treatment effects on the growth mindset scale.

Other outcomes: In Appendix Tables A15, we show the estimations to explore heterogeneity in respondent self-efficacy. We find that self-efficacy increases among treated respondents with mothers who have high education levels. It is possible that educated mothers act as role models or take a greater interest in encouraging their daughters, which is reflected in higher treatment effects. Overall, the mother's education status and competitive levels have a significant impact on primary outcomes, as measured by the Mindset Index created out of grit, growth mindset, and self-efficacy scales using (Anderson, 2008) (Appendix Table A16). Consistent with our results, (Roy and Bhattacharya, 2018) find that the existence of a supportive or enabling environment, proxied by knowing other successful women, feeling connected to the broader social network, and educated parents, is found to play an important and independent role in facilitating education aspirations and outcomes of young women in case of India.

We also find that commitment to academic goals decreases among treated students who don't discuss their college work with parents but increases among students who do. The evidence is in line with the literature that suggests that parents take guidance from their personal experiences, and they are more likely to set realistic goals for their children based on their prior academic performance. Thus, it's more likely that when students discuss their goals with their parents, they are more committed to their goals which are realistic, based upon their prior academic performance, and more likely to commit to them Zimmerman (1992). Finally, though we detect the positive average effect of treatment on willingness to modify goals to SMART, we find that this effect may be driven by respondents who have higher cognitive abilities, as measured by their performance on the raven's score - willingness to modify scores is 21.1% higher among treated students who have better than median performance in Raven's test. We do not find significant evidence of heterogeneity of cognitive ability in modifying goals (as seen in Alan et al. (2019)). Overall, we find treatment effects vary by individual and household characteristics. In particular, students with educated mothers report no change in their level of grit; conversely, those with mothers with lower levels of education record higher levels of grit even one year later. However, a mother's education is important in promoting a growth mindset and higher levels of self-efficacy.

First year students who received treatment had higher grit but it did not translate in change in test scores.

Table 5: Heterogeneity in treatment impact in percentage scored in exams two months after the intervention

Independent variable: Exam percent scored at two months					
H:	HH assets (1)	Mother education (2)	Time in hh chores (3)	Career goal importance (4)	Raven's score (5)
Treated	0.080	0.045	0.028	0.036	0.050
	(0.056)	(0.040)	(0.045)	(0.041)	(0.047)
H	0.016	0.047	-0.029	0.010	0.025
	(0.027)	(0.027)*	(0.026)	(0.027)	(0.027)
H*Treated	-0.114	-0.049	-0.020	-0.028	-0.063
	(0.061)*	(0.061)	(0.053)	(0.060)	(0.055)
Mean	0.496	0.496	0.496	0.496	0.496
N	330	330	330	330	330
R^2	0.018	0.005	0.007	0.002	0.005
Independent variable: Exam percent scored at two months					
H:	Compet-itiveness (6)	First division (7)	First year (8)	Time in study (9)	Discuss work (10)
Treated	0.038	0.091	0.061	0.048	0.094
	(0.052)	(0.069)	(0.047)	(0.045)	(0.083)
H	0.010	0.089	0.059	0.004	-0.015
	(0.027)	(0.029)***	(0.028)**	(0.005)	(0.031)
H*Treated	-0.033	-0.088	-0.111	-0.008	-0.105
	(0.058)	(0.076)	(0.005)**	(0.009)	(0.086)
Mean	0.496	0.496	0.496	0.496	0.496
N	330	329	330	330	330
R^2	0.003	0.012	0.073	0.003	0.022

Note: Standard errors in parentheses. All errors are clustered at the individual level. Percent scored refer to the marks secured by respondent in end-of-year examinations (2 months after the intervention) as a percentage of maximum (full) marks allowed for respective year of study. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Heterogeneity in treatment impact in percentage scored in exams nine months after intervention

Independent variable: Exam percent scored at nine months					
H:	HH assets (1)	Mother education (2)	Time in hh chores (3)	Career goal importance (4)	Raven's score (5)
Treated	1.909	1.918	2.34	-2.031	-0.191
	(3.322)	(3.17)	(3.298)	(-3.61)	(3.321)
H	-1.368	1.083	3.916	0.327	2.282
	(3.571)	(3.583)	(3.494)	(3.503)	(3.52)
H*Treated	-3.883	-4.657	-6.662	3.841	-0.09
	(5.054)	(5.094)	(5.084)	(4.997)	(5.083)
Mean	40.346	40.346	40.346	40.346	40.346
N	332	332	332	332	332
R^2	0.007	0.003	0.005	0.004	0.002
Independent variable: Exam percent scored at nine months					
H:	Compet-itiveness (6)	First division (7)	First year (8)	Time in study (9)	Discuss work (10)
Treated	-1.14	5.448	0.753	-6.621	1.979
	(3.512)	(4.452)	(3.072)	(5.301)	(4.914)
H	0.337	6.803	0.387	-1.083	1.57
	(3.505)	(3.705)*	(3.844)	(0.991)	(3.849)
H*Treated	1.845	-7.661	-2.532	1.862	-3.182
	(5.027)	(5.39)	(5.339)	(1.351)	(5.707)
Mean	40.346	40.346	40.346	40.346	40.346
N	332	332	332	332	332
R^2	0.001	0.008	0.001	0.005	0.001

Note: Standard errors in parentheses. All errors are clustered at the individual level. Percent scored refer to the marks secured by respondent in end-of-year examinations (9 months after the intervention) as a percentage of maximum (full) marks allowed for respective year of study. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Heterogeneity in treatment impact on Grit

Panel A: Independent variable: Grit					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.496	-0.674	-1.538	-0.386	-0.945
	(0.909)	(0.834)	(0.789)*	(0.791)	(0.857)
Post	-0.600	-0.611	-0.580	-0.632	-0.631
	(0.530)	(0.529)	(0.527)	(0.521)	(0.525)
H	-0.662	2.020	0.145	3.236	1.082
	(0.823)	(0.811)**	(0.836)	(0.777)***	(0.810)
H*Treated	-0.937	-0.755	1.445	-1.248	-0.125
	(1.207)	(1.193)	(1.252)	(1.170)	(1.199)
Treated*Post	1.227	1.783	1.603	2.668	2.520
	(0.993)	(0.930)*	(0.839)*	(0.899)***	(0.884)***
H*treated*Post	0.830	-0.195	0.037	-1.983	-2.032
	(1.063)	(1.047)	(1.108)	(1.045)*	(1.038)*
Mean	41.125	41.125	41.125	41.125	41.125
N	666	666	666	666	666
R^2	0.012	0.023	0.013	0.045	0.012
Panel B: Independent variable: Grit					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.643	0.371	-0.457	-1.198	-1.124
	(0.729)	(1.115)	(0.752)	(1.139)	(1.146)
Post	-0.622	-0.658	-0.604	-0.527	-0.571
	(0.517)	(0.528)	(0.527)	(0.524)	(0.527)
H	3.191	1.795	2.412	0.545	0.281
	(0.789)***	(0.850)**	(0.926)***	(0.175)***	(0.926)
H*Treated	-0.624	-1.843	-1.945	0.027	0.155
	(1.185)	(1.311)	(1.301)	(0.289)	(1.326)
Treated*Post	2.650	1.485	0.808	2.914	1.478
	(0.902)***	(0.992)	(0.838)*	(1.190)*	(1.135)
H*Treated*Post	-2.119	0.284	2.129	-0.429	0.237
	(1.046)**	(1.073)	(1.097)*	(0.273)	(1.180)
Mean	41.125	41.125	41.125	41.125	41.125
N	666	662	666	666	666
R^2	0.049	0.012	0.031	0.031	0.006

Note: Standard errors in parentheses. All errors are clustered at the individual level. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale (Duckworth et al., 2007). The H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Heterogeneity in treatment impact on Growth Mindset

Independent variable: Growth mindset					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	0.697	2.317	0.318	1.050	0.927
	(0.915)	(0.916)**	(0.853)	(0.930)	(0.890)
Post	-2.764	-2.764	-2.725	-2.742	-2.645
	(0.718)***	(0.715)***	(0.713)***	(0.714)***	(0.712)***
H	-0.410	0.844	-1.176	-0.503	-2.103
	(0.900)	(0.865)	(0.885)	(0.869)	(0.853)**
H*Treated	0.993	-2.392	2.277	0.282	0.501
	(1.281)	(1.250)*	(1.297)*	(1.258)	(1.245)
Treated*Post	-0.376	-1.539	0.591	0.480	-0.023
	(1.207)	(1.140)	(1.062)	(1.157)	(1.155)
H*treated*Post	1.251	3.862	-1.042	-0.521	0.375
	(1.334)	(1.310)***	(1.420)	(1.335)	(1.319)
Mean	48.660	48.660	48.660	48.660	48.660
N	666	666	666	666	666
R^2	0.048	0.055	0.048	0.045	0.059

Independent variable: Growth mindset					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.151	0.189	0.882	-0.334	0.710
	(0.921)	(1.247)	(0.832)	(1.109)	(1.135)
Post	-2.716	-2.721	-2.756	-2.796	-2.759
	(0.709)***	(0.716)***	(0.713)***	(0.707)***	(0.712)***
H	-2.460	0.321	0.504	-0.491	-0.379
	(0.850)***	(0.997)	(0.924)	(0.188)***	(0.904)
H*Treated	2.785	1.614	0.678	0.466	0.680
	(1.246)**	(1.429)	(1.300)	(0.248)*	(1.343)
Treated*Post	0.211	1.808	0.718	-0.147	0.386
	(1.121)	(1.410)	(1.063)	(1.374)	(1.392)
H*treated*Post	-0.012	-2.344	-1.183	0.140	-0.194
	(1.343)	(1.453)	(1.407)	(0.292)	(1.440)
Mean	48.660	48.660	48.660	48.660	48.660
N	666	662	666	666	666
R^2	0.058	0.049	0.046	0.055	0.044

Note: Standard errors in parentheses. All errors are clustered at the individual level. Growth mindset is a scale constructed from the sum of 15 items on 7 to 6-point Likert scale (**?**). H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We also find that first year treated students are less likely to modify their goals.[33] One explanation for these results could be that may be there is additional planning activity required for first year students to improve their academic performance. Another related explanation could be that first year have unrealistic expectations of obstacles they may face in order to achieve their academic goals, that our treatment might not have been able to overcome. Indeed, we see indications of first-year students underestimating the difficulties or obstacles (to their academic goals) that they are likely to face. At baseline, first year students were less likely to report that they expect to face obstacles as compared to students from from advanced years – 22% of the first year students expect to face difficulty in retaining large amounts of information as compared to 44% of the students in higher years. Only about one out of every four first year students in our sample expect to face difficulty in managing time efficiently compared to half of all of senior students. We see similar difference with senior students reporting greater obstacles in the form of lack of motivation to study, especially when they compete with better performing students. It is possible that first year students, while experiencing an improvement in grit, were ill-equipped to translate this into an improved academic performance.

We find the year of the enrollment is an important factor for how effective the treatment can be. Finally, cognitive levels play an important role - grit increases for treated students with lower levels of cognition (as measured by Raven's test scores and whether they scored a First Division in the final exams in their previous degree). The cognitive ability seems particularly important for improving the willingness to modify goals to SMART goals; however, it is not a factor determining treatment effects on self-efficacy or test scores. While we cannot make any claims about the causality of these characteristics, they provide interesting avenues for future investigation and understand the sub-samples in the population where similar interventions may be particularly effective.

7 Discussion and Conclusion

In many developing countries context, improving learning outcomes have long been an important concern for development policy. The literature emphasizes the development of non-cognitive skills, such as grit and perseverance, can be an important predictor of academic performance, retention rates, and job market outcomes. This study uses a randomized control trial to develop a soft-touch, easily scalable, low-cost intervention that aims to foster grit, mindset, and academic performance. It does so by highlighting the

[33] In our followup phone survey, when we quizzed respondents on if they remember their goals, a higher fraction, approximately 62%, of higher year student remembered their goals as compared to 55% of students from first year who were less likely to remember their goals.

role of effort and persistence in enhancing skills and achieving goals. Students in our sample also participated in a goal-setting exercise and were introduced to the concept of "SMART" - Specific, Measurable, Ambitious, Realistic, Time-bound goals for their academic and career goals.

Our study adds to the literature in several ways. First, we contribute to a growing literature that documents the effect of mindset interventions on non-cognitive skills. Similar studies have been conducted with elementary school students in Istanbul (Alan et al., 2019) and secondary school students in Peru (Outes-León et al., 2020). Second, we join a small group of recent studies testing mindset interventions in a developing world context (Ganimian (2020); Outes-León et al. (2020)). We focus on female students from low to middle-income families in Pakistan and test a cost-effective intervention that can be easily scaled up to be administered as part of the higher-education curriculum in developing countries. Third, literature has largely considered non-cognitive skills as fixed in adults and concentrated on developing these skills in younger, school-going children (Alan et al., 2019; Yeager et al., 2016, 2018). Two studies, Outes-León et al. (2020) and Ganimian (2020), implement similar interventions with older secondary school students in developing countries. In this study, we are able to provide evidence on the malleability of these skills in young college students. Lastly, we exploit rich data to test how the effect of the intervention varies with student and household characteristics, such as parental education.

Our mindset intervention leads to a significant increase in grit levels of the treated students one year after the intervention. In addition, treated students were more receptive to changing their goals to SMART goals. However, the change in mindset and goals being set did not translate into a change in test scores two and nine months later. An analysis of results by student characteristics reveals some interesting insights: our findings suggest that the mindset intervention may be most effective for disadvantaged students – students that were less likely to rank academic career goals as important, those who scored lower on cognitive Raven's matrices, and students who were less competitive students at baseline experienced a greater increase in grit.

Second, we also find that grit substantially improved for freshmen treated students, but it is not sufficient to increase the test scores. First-year students are also less likely to modify their goals to SMART goals. This suggests that for mindset interventions to reduce learning gaps, it will need to be complemented by identifying obstacles and planning to achieve academic goals. Third, treated students with educated mothers score higher on growth mindset and self-efficacy scales, confirming evidence from earlier research that educated mothers can provide a supportive environment that foster education aspirations and outcomes of children, particularly among female children (Roy and Bhattacharya, 2018; Andrabi et al., 2012). In addition, commitment to self-reported academic goals is higher among students who discuss their homework and assignments with their parents.

Our results suggest that it may be possible to change the mindset and reduce learning gaps in urban areas of Pakistan, an encouraging result since a large fraction of students in different levels of education face a learning crisis. Our intervention is inexpensive and can easily be brought to scale, e.g., via specialized discussions in class or as part of college extra-curricular activities, with potentially sustained effects on student mindset. An important question that can be explored in future research, given the insignificant impact on test scores of first-year treated students, is the need to complement guided mindset discussions with the identification of obstacles. Planning strategies and repeated discussions to reinforce goals may translate grit into higher effort and improved test scores.

Our findings are important for policy focusing on educating women for improving incomes and a wider set of human development outcomes in low-income countries. The intervention is simple to implement and intuitive, lending itself easily to complement the existing curriculum in schools and colleges. We find even small discussions can have sustained effects on student mindset one year later and are particularly relevant for students who are disadvantaged compared to their counterparts. This is not just effective in the short run but can have long-term effects as well.

References

ABADIE, A., S. ATHEY, G. W. IMBENS, AND J. WOOLDRIDGE (2017): "When Should You Adjust Standard Errors for Clustering?" Working Paper 24003, National Bureau of Economic Research.

ALAN, S., T. BONEVA, AND S. ERTAC (2019): "Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit," The Quarterly Journal of Economics, 134, 1121–1162.

ALAN, S. AND S. ERTAC (2015): "Patience, self-control and the demand for commitment: Evidence from a large-scale field experiment," Journal of Economic Behaviour and Organization, 115, 111–122.

ANDERSON, M. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," Journal of the American Statistical Association, 103, 1481–1495.

ANDRABI, T., J. DAS, AND A. I. KHWAJA (2012): "What did you do all day? Maternal education and child outcomes," Journal of Human Resources, 47.

BANERJEE, A. V., R. BANERJI, E. DUFLO, R. GLENNERSTER, AND S. KHEMANI (2010): "Pitfalls of participatory programs: Evidece from a randomized evaluation in education in India," American Economics Journal: Economic Policy, 2, 1–30.

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2013): "Inference on Treatment Effects after Selection among High-Dimensional Controls," The Review of Economic Studies, 81, 608–650.

——— (2014a): "High-Dimensional Methods and Inference on Structural and Treatment Effects," The Journal of Economic Perspectives, 28, 29–50.

——— (2014b): "Inference on Treatment Effects After Selection Among High-dimensional Controls," The Review of Economic Studies, 81, 608–650.

BETTINGER, E., S. LUDVIGSEN, M. REGE, I. F. SOLLI, AND D. YEAGER (2018): "Increasing perseverance in math: Evidence from a field experiment in Norway," Journal of Economic Behavior Organization, 146, 1–15.

BLACKWELL, L. S., K. H. TRZESNIEWSKI, AND C. S. DWECK (2007): "Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention," Child Development, 78, 246–263.

CENSUS (2017): "Population Census of Pakistan," .

DELAVANDE, A., D. B. E. H. A.-. S. S. (2019): "Skills Accumulation with Malleable Ability: Evidence from a Growth Mindset Intervention." Unpublished manuscript.

DELLA VIGNA, S. AND E. LINOS (2020): "RCTs to Scale: Comprehensive Evidence from Two Nudge Units," Working Paper, UC Berkley.

DILLON, M. R., H. KANNAN, J. T. DEAN, E. S. SPELKE, AND E. DUFLO (2017): "Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics," Science, 375, 47–55.

DOBRONYI, C. R., O. P. AND U. PETRONIJEVIC (2019): "Goal setting, academic reminders, and college success: A large-scale field experiment," Journal of Research on Educational Effectiveness, 12, 38–66.

DUCKWORTH, ANGELA L., . Q. P. D. (2009): "Development and validation of the Short Grit Scale (GRIT-S)," Journal of personality assessment, 91, 166–174.

DUCKWORTH, A. L., C. PETERSON, M. D. MATTHEWS, AND D. R. KELLY (2007): "Grit: Perseverance and passion for long-term goals," Journal of Personality and Social Psychology, 92, 1087–1101.

DURLAK, J. A., R. P. WEISSBERG, A. B. DYMNICKI, R. D. TAYLOR, AND K. B. SCHELLINGER (2011): "The impact of enhancing students' social and emotional learning: A metaâĂŘanalysis of school-based universal interventions," Child Dvelopment, 82, 405–432.

DWECK, C. (2006): "Mindset: The new psycology of sucess," .

FILMER, D., L. M. S. R. . V. T. (2018): "Learning to Realize EducationâĂŹs Promise. World Development Report. The World Bank," .

FISZBEIN, A. AND N. SCHADY (2009): "Conditional Cash Transfers: Reducing Present and Future Poverty," Unpublished manuscript.

GANIMIAN, A. J. (2020): "Growth mindset interventions at scale: Experimental evidence from Argentina," Educational Evaluation and Policy Analysis, 42, 417–438.

GLEWWE, P. AND M. KREMER (2006): "Schools, Teachers, And Education Outcomes in Developing Countries," Handbook of the Economics of Education, 2.

GLEWWE, P. AND E. MIGUEL (2008): " The Impact of Child Health and Nutrition on Education in Less Developed Countries," Handbook of Development Economics, 4, 3561–3606.

HAGGER, M., C.-Q. ZHANG, E.-M. KANGRO, F. RIES, J. C. WANG, B. HERITAGE, AND D. K. C. CHAN (2018): "Trait Self-Control and Self-Discipline: Structure, Validity, and Invariance Across National Groups," Current Psychology, 1–16.

HAROON, M. AND F. SAID (2019): "The Role of Non-Cognitive Skills in Improving Academic Performance," AEA RCT.

HAROON, M., F. SAID, AND M. ZAFAR (2020): "The Role of Non-Cognitive Skills in Improving Academic Performance," AEA RCT.

HECKMAN, J. J. AND Y. RUBINSTEIN (2001): "The importance of noncognitive skills: Lessons from the GED testing program," American Economic Review, 91, 145–149.

HECKMAN, J. J., J. STIXRUD, AND S. URZUA (2006): "The effects of cognitive and noncognitive abilities on labor market outcomes and social behaviour," Journal of Labor Economics, 24, 411–482.

HUMPHRIES, J. E. AND K. FABIAN (2017): "On the intrepretation of non-cognitive skills - What is being measures and why it matters," Journal of Economic Behaviour and Organization, 136, 174–185.

ISLAM, A., S. KWON, E. MASOOD, N. PRAKASH, S. SABARWAL, AND D. SARASWAT (2020): "When Goal-Setting Forges Ahead but Stops Short," IZA Discussion Papers.

KAUTZ, T., J. J. HECKMAN, R. DIRIS, B. TER WEEL, AND L. BORGHANS (2014): "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success," National Bureau of Economic Research, w20749.

KROENKE, K., L. S. ROBERT, AND B. W. JANET (2001): "The PHQ-9: validity of a brief depression severity measure," Journal of general internal medicine, 16, 606–613.

LEE, D. S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," The Review of Economic Studies, 76, 1071–1102.

LINDQVIST, E. AND R. VESTMAN (2011): "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish Enlistment," American Economic Journal: Applied Economics, 3, 101–128.

LYUBOMIRSKY, S. AND H. S. LEPPER (1999): "A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation," Social Indicators Research, 46, 137–155.

MADDI, S. R., M. D. MATTHEWS, D. R. KELLY, B. VILLARREAL, AND M. WHITE (2012): "The role of hardiness and grit in predicting performance and retention of USMA cadets," Military Psychology, 24, 19–28.

MCEWAN, P. (2015): "Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments," Review of Educational Research, 85, 353–394.

MORISANO, D., J. B. HIRSH, J. B. PETERSON, R. O. PIHL, AND B. M. SHORE (2010): " Setting, Elaborating, and Reflecting on Personal Goals Improves Academic Performance," Journal of Applied Psychology, 95, 255–264.

OREOPOULOS, P., P. R. W. P. U. . P. N. G. (2018): "When Studying and Nudging DonâĂŹt Go as Planned: Unsuccessful Attempts to Help Traditional and Online College Students," National Bureau of Economic Research, w25036.

OUTES-LEÓN, I., A. SÁNCHEZ, AND R. VAKIS (2020): "The Power of Believing You Can Get Smarter: The impact of a growth-mindset intervention on academic achievement in Peru. World Bank Policy Paper 9141," .

PAKISTAN BUREAU OF STATISTICS (2019): "Pakistan Social and Living Standards Measurement Survey PSLM/HIES 2018-19 Report," .

PAUNESKU, D., G. M. WALTON, C. ROMERO, E. N. SMITH, D. S. YEAGER, AND C. S. DWECK (2015): "Mind-set interventions are a scalable treatment for academic underachievement," Psychological science, 26, 784–793.

PDS (2017): "Punjab Development Statistics Report," .

——— (2018): "Punjab Development Statistics 2018," .

POLLEY, T. (2018): "Believing is Achieving: Targeting Beliefs about Intelligence to Increase Learning in Bangladeshi Secondary Schools," Unpublished manuscript.

ROBERTS, B. W., N. R. KUNCEL, R. SHINER, A. CASPI, AND L. R. GOLDBERG (2007): "The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes," Perspectives and Psychological Science, 2, 313–345.

ROBERTSON-KRAFT, C. AND A. L. DUCKWORTH (2014): "True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers," Teachers College Record(1970), 116.

ROSENBERG, M. (1965): "Society and the adolescent self-image," .

ROY, SANCHARI, M. M. AND S. BHATTACHARYA (2018): "Hidden human capital: Self-efficacy, aspirations and achievements of adolescent and young women in India," World Development, 111, 161–180.

RYCKMAN, R. M., M. HAMMER, L. M. KACZOR, AND J. A. GOLD (1996): "Construction of a Personal Development Competitive Attitude Scale," Journal of Personality Assessment, 66, 374–385.

SCHIPPERS, M.C., W. M. AND J. DAWSON (2015): "Team Reflexivity and Innovation: The Moderating Role of Team Context," Journal of Management, 41(3), 769âĂŞ788.

SCHWARZER, R. AND M. JERUSALEM (1995): "Generalized Self-Efficacy scale," in Measures in health psychology: A user's portfolio. Causal and control beliefs, ed. by J. Weinman, S. Wright, and M. Johnston, Windsor, England: NFER-Nelson, 35–37.

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

UNESCO (2009): "EFA Global Monitoring Report: Overcoming Inequality: Why Governance Matters." .

WORLD BANK (2018): "The many faces of the learning crisis," .

YEAGER, D. S. AND C. S. DWECK (2012): "Mindsets that promote resilience: When students believe that personal characteristics can be developed," Educational Psychologist, 47, 302–314.

YEAGER, D. S., P. HANSELMAN, D. PAUNESKU, C. HULLEMAN, C. DWECK, C. MULLER, AND A. DUCKWORTH (2018): "Where and for whom can a brief, scalable mindset intervention improve adolescents' educational trajectories," PsyArXiv.

YEAGER, D. S., R. JOHNSON, B. J. SPITZER, K. H. TRZESNIEWSKI, J. POWERS, AND C. S. DWECK (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., C. ROMERO, D. PAUNESKU, C. S. HULLEMAN, B. SCHNEIDER, C. HINOJOSA, H. Y. LEE, AND J. E. A. O'BRIEN (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, 374–391.

ZIMMERMAN, B. J., B. A. . M.-P. M. (1992): "Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting," American educational research journal, 29(3), 663–676.

Appendix

A1 Sample balance and attrition

Sample balance

At followup, we were able to successfully contact 300 students. The sample is qualitatively similar, except on age and desire to pursue a paid job. Specifically, treated students in this balanced sample (of 300) are younger when compared to the placebo sample; in addition, they are less likely to purse a paid job. The overall test of joint significance shows the sample is well balanced by treatment status ($p - value$ = 0.14). We also test for differences in characteristics of respondents who were interviewed in the person and those interviewed via phone at end line and find no significant overall difference between the two samples ($p - value$ = 0.86)

Table A1: Full sample description and balance at baseline

	N	Mean	Median	S.Dev.	$p-value$ treatment	$p-value$ type
	(1)	(2)	(3)	(4)	(5)	(6)
Age (years)	300	19.32	19.00	1.54	0.09*	0.31
Dummy: Enrolled in first year	300	0.37	0.00	0.48	0.15	0.25
Dummy: Single	300	0.95	1.00	0.21	0.45	0.36
Father's education (years)	289	8.57	10.00	4.81	0.81	0.80
Dummy: Father is a business owner	300	0.24	0.00	0.43	0.70	0.54
Dummy: Father is a salaried worker	300	0.58	1.00	0.49	0.26	0.96
Mother's education (years)	297	6.69	8.00	5.19	0.86	0.06*
Dummy: Mother is a business owner	300	0.02	0.00	0.14	0.49	0.02**
Dummy: Mother is a salaried worker	300	0.06	0.00	0.23	0.60	0.06*
Average monthly household income (000's)	213	36.20	25.00	35.23	0.32	0.85
Index: Household assets	300	-0.02	0.00	1.37	0.77	0.42
Risk preferences (higher is more risk averse)	300	4.15	4.00	1.59	0.22	0.06*
Ravens test score (out of 10)	300	4.16	4.00	2.11	0.42	0.36
Scale: Competitiveness (out of 75)	300	58.00	58.00	8.31	0.63	0.40
Dummy: Continue education after graduating	300	0.91	1.00	0.28	0.59	0.50
Dummy: Passed last degree in first division	299	0.74	1.00	0.44	0.11	0.55
Satisfaction with academic performance	300	2.40	2.00	0.91	0.93	0.68
Daily hours studying at home in average week	300	2.94	3.00	2.09	0.37	0.72
Daily hours doing household chores in average week	300	6.43	7.00	6.12	0.37	0.53
Dummy: Discusses homework with parents	300	0.70	1.00	0.46	0.79	0.30
Dummy: Discusses goals with parents	300	0.91	1.00	0.28	0.18	0.84
Dummy: Pursue paid job	300	0.93	1.00	0.26	0.07*	0.34
Dummy: Have a role model	300	0.34	0.04	0.47	0.25	0.54
P-value of F-statistic (treatment)						0.14
P-value of F-statistic (type)						0.86

Note: Column (5) shows $p-values$ from the balance test specified in equation 7. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. Column (6) shows the $p-values$ from a test of baseline balance between the in-person and phone interview sample ('type' of survey), from a regression of the variable in the row on the type of interview. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. $***p<0.01, **p<0.05, *p<0.1$.

We show treatment balance in the balanced panel that participated in face-to-face surveys in both rounds. While restricting our sample to only the 226 who participated in face-to-face surveys, average characteristics varied by age, those enrolled in first year, and whether they pursue a paid job or not. The overall test of joint significance shows the sample is well balanced by treatment status ($p - value$ = 0.54)

Table A2: Sample description of face-to-face sample

	N	Mean	Median	S.Dev.	$p - value$
	(1)	(2)	(3)	(4)	(5)
Age (years)	226	19.37	19.00	1.61	0.03**
Dummy: Enrolled in first year	226	0.38	0.00	0.49	0.06*
Dummy: Single	226	0.95	1.00	0.22	0.42
Father's education (years)	217	8.61	10.00	4.71	0.38
Dummy: Father is a business owner	226	0.23	0.00	0.42	0.96
Dummy: Father is a salaried worker	226	0.58	1.00	0.58	0.61
Mother's education (years)	223	6.36	8.00	5.25	0.64
Dummy: Mother is a business owner	226	0.01	0.00	0.09	0.18
Dummy: Mother is a salaried worker	226	0.07	0.00	0.26	0.79
Average monthly household income (000's)	163	35.94	25.00	38.17	0.64
Index: Household assets	226	-0.06	0.00	1.35	0.67
Risk preferences (higher is more risk averse)	226	4.25	4.00	1.55	0.18
Ravens test score (out of 10)	226	4.22	4.00	2.19	0.78
Scale: Competitiveness (out of 75)	226	58.23	58.00	8.39	0.69
Dummy: Continue education after graduating	226	0.91	1.00	0.29	0.95
Dummy: Passed last degree in first division	226	0.75	1.00	0.44	0.25
Satisfaction with academic performance	226	2.42	2.00	0.93	0.78
Daily hours studying at home in average week	226	2.92	3.00	2.12	0.46
Daily hours doing household chores in average week	226	6.56	7.00	6.19	0.49
Dummy: Discusses homework with parents	226	0.71	1.00	0.45	0.89
Dummy: Discusses goals with parents	226	0.91	1.00	0.28	0.52
Dummy: Pursue paid job	226	0.94	1.00	0.24	0.05**
Dummy: Have a role model	226	0.35	0.00	0.48	0.59
P-value of F-statistic					0.54

Note: Column (5) shows $p - values$ from the balance test specified in equation 1 for 266 individuals who received the repeat treatment. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. $***p < 0.01, **p < 0.05, *p < 0.1$.

Attrition

We were unable to interview 66 students interviewed at baseline for in the followup survey. Attrition does not vary by treatment status, as shown in Table A3. Attriters vary by level of competitiveness and degree of satisfaction with their academic performance.

Table A3: Balance test for Attrition

	N	Mean	Median	S.Dev.	$p-value$
	(1)	(2)	(3)	(4)	(5)
Age (years)	366	19.37	19.00	1.62	0.13
Dummy: Enrolled in first year	366	0.36	0.00	0.48	0.33
Dummy: Single	366	0.96	1.00	0.20	0.24
Father's education (years)	351	8.43	10.00	4.80	0.24
Dummy: Father is a business owner	366	0.25	0.00	0.44	0.31
Dummy: Father is a salaried worker	366	0.57	1.00	0.50	0.14
Mother's education (years)	362	6.64	8.00	5.15	0.73
Dummy: Mother is a business owner	366	0.02	0.00	0.13	0.25
Dummy: Mother is a salaried worker	366	0.05	0.00	0.23	0.72
Average monthly household income (000's)	258	35.56	25.00	33.09	0.50
Index: Household assets	366	0.00	0.30	1.39	0.48
Risk preferences (higher is more risk averse)	366	4.15	4.00	1.59	0.88
Ravens test score (out of 10)	366	4.08	4.00	2.11	0.15
Scale: Competitiveness (out of 75)	366	57.62	58.00	8.33	0.07*
Dummy: Continue education after graduating	366	0.91	1.00	0.28	0.91
Dummy: Passed last degree in first division	363	0.72	1.00	0.45	0.11
Satisfaction with academic performance	366	2.37	2.00	0.91	0.09*
Daily hours studying at home in average week	366	2.99	3.00	2.07	0.32
Daily hours doing household chores in average week	366	6.28	7.00	5.99	0.31
Dummy: Discusses homework with parents	366	0.71	1.00	0.46	0.33
Dummy: Discusses goals with parents	366	0.93	1.00	0.27	0.26
Dummy: Pursue paid job	366	0.92	1.00	0.27	0.32
Dummy: Have a role model	366	0.33	0.00	0.47	0.57
Dummy: Treatment	366	0.51	1.00	0.51	0.20
P-value of F-statistic					0.11

Note: Column (5) shows p-values from the balance test specified in equation 1. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. ***p < 0.01, **p < 0.05,*p < 0.1.

A2 Spillovers results

In this section, we present the results from spillovers several robustness checks conducted to estimate treatment effects on primary and secondary outcomes.

Table A4: Spillover effects: Indirect impact of treatment on exam performance of peers at 2 months

	Test score percentage (1)	Test score percentage (2)
Treated	0.019	0.014
	(0.036)	(0.038)
	[0.101]	[0.106]
Mean	0.500	0.500
N	217	217
R^2	0.001	0.165

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects.Test score percentage' are the marks scored in end-of-year examination as a percentage of the maximum (full) marks allowed for that enrollment year. These were measured 2 months after the baseline activities and are present for only one round of the data. Treated is the indirect treatment effect of our growth mindset treatment of participant i on j outcomes. MDE is the ex post minimum detectable effect size at a significance level of 0.05 and power of 80 percent and is provided in square brackets for Treated. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Spillover effects: Indirect impact of treatment on primary outcomes on peers

	(1) Grit Grit	(2) Grit Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset index	(8) Mindset index
Treated	0.512	0.649	-0.570	-0.906	0.188	-0.065	0.008	-0.037
	(0.820)	(0.823)	(0.903)	(0.887)	(0.692)	(0.706)	(0.131)	(0.136)
Post	0.331	0.294	-1.816	-1.789	-1.095	-1.203	-0.255	-0.266
	(0.838)	(0.836)	(1.176)	(1.172)	(0.763)	(0.771)	(0.131)*	(0.133)**
Treated*Post	-0.597	-0.562	-0.322	-0.341	0.962	1.049	0.003	0.013
	(0.995)	(0.996)	(1.346)	(1.336)	(0.921)	(0.926)	(0.166)	(0.165)
	[2.786]	[2.788]	[3.768]	[3.741]	[2.578]	[2.592]	[0.464]	[0.462]
Mean	40.391	40.391	49.971	49.971	32.167	32.167	-0.153	-0.153
N	443	443	443	443	443	443	443	443
R^2	0.001	0.042	0.024	0.084	0.007	0.068	0.016	0.052
Controls	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. T*Post is the indirect average treatment effect of our growth mindset treatment of participant i on j outcomes. MDE is the ex post minimum detectable effect size at a significance level of 0.05 and power of 80 percent and is provided in square brackets for Treated*Post. 'Mean' is the average value of the outcome for the control group. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Spillover effects: Indirect impact of treatment on exam performance at 2 months on subsample of matched peers

	Test score percentage (1)	Test score percentage (2)
Treated	0.010	0.020
	(0.044)	(0.046)
	[0.123]	[0.129]
Mean	0.508	0.508
N	217	217
R^2	0.000	0.165

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects.Test score percentage' are the marks scored in end-of-year examination as a percentage of the maximum (full) marks allowed for that enrollment year. These were measured 2 months after the baseline activities and are present for only one round of the data. Treated is the average treatment effect of our growth mindset intervention. MDE is the ex post minimum detectable effect size at a significance level of 0.05 and power of 80 percent and is provided in square brackets for Treated. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A7: Spillover effects: Indirect impact of treatment on primary outcomes on subsample of matched peers

	(1) Grit Grit	(2) Grit Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset index	(8) Mindset index
Treated	-0.843	-0.851	0.377	0.293	-0.038	-0.055	-0.044	-0.055
	(0.756)	(0.769)	(0.845)	(0.845)	(0.650)	(0.653)	(0.125)	(0.131)
Post	-0.888	-0.915	-2.278	-2.272	-0.370	-0.372	-0.346	-0.348
	(0.642)	(0.635)	(0.804)***	(0.810)***	(0.566)	(0.570)	(0.103)***	(0.106)***
Treated*Post	1.607	1.631	0.477	0.485	-0.123	-0.211	0.186	0.181
	(0.902)*	(0.900)*	(1.161)	(1.165)	(0.855)	(0.856)	(0.161)	(0.162)
	[2.526]	[2.520]	[3.251]	[3.262]	[2.394]	[2.397]	[0.451]	[0.454]
Mean	40.609	40.609	49.173	49.173	32.645	32.645	-0.167	-0.167
N	443	443	443	443	443	443	443	443
R^2	0.005	0.045	0.024	0.081	0.002	0.065	0.018	0.054

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. T*Post is the average treatment effect of our growth mindset intervention. MDE is the ex post minimum detectable effect size at a significance level of 0.05 and power of 80 percent and is provided in square brackets for Treated*Post.'Mean' is the average value of the outcome for the control group. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A3 Secondary outcomes

After the first round of the implementation of our treatment protocols, we again re-iterated the same message of our treatment and placebo discussions. In order to focus the discussion in the repeat intervention, three months after the intervention was first implemented, we conducted focus group discussions with 10%, randomly selected students in our sample and quizzed them on what they remember from the baseline discussion. We found that recall was better in the treated than in the placebo group - students in the placebo group correctly answered 2.9 out of 5 quiz questions correctly, compared to 4.7 out of 5 questions asked of the treated group (p = 0.000). We also asked respondents about the examples used in the discussion that were most useful and clear in illustrate the main messages of the discussion. As a result, while the main discussions in the repeat interventions were the same as the intervention implemented at baseline, they were relatively shorter in length.

The second round of treatment/placebo discussions could only be conducted with 61% of our treated and control group participants who were available for interview while colleges were still open. The remainder of the sample, contacted via phone, only participated in a shortened survey, with no discussion included in the repeat intervention. In order to incentivise participation, in addition to the treatment or placebo activity, all students receive detailed information regarding the availability of scholarships for higher degrees in different colleges and universities across Lahore. The information included eligibility criterion for scholarships and process to apply for these scholarships. However, this discussion took place at the end of the survey, after all other questions and the treatment/placebo discussion had taken place. Treatment assignment for this is summarized in Table A8.

Treatment Status	Scholarship discussion	N
Treatment: Growth mindset discussion	YES	120
Placebo: Placebo discussion	YES	106

Table A8: **Assignment to treatment in round 2**

We estimate the immediate effect of the intervention on several post-intervention outcomes measured at baseline - these relate to outcomes measured during a goal setting activity where student's are asked to specify their academic and career goals. In this discussion, enumerators ask respondents about the importance of the academic goal to them, the perceived likelihood of them being able to achieve the goal they have set for themselves, their commitment in trying to achieve this goal and finally, their willingness to modify their goals to SMART (Specific, Measurable, Ambitious, Realistic and Time-based) goals.

Results, with an without college and enumerator fixed effects, are shown in Table A9. We find insignificant treatment effects, except on the willingness to modify goals to SMART goals, where the effect is small but positive and statistically significant.

Table A9: Immediate impact of treatment on secondary outcomes

Goal:	Importance	Importance	Achieve	Achieve	Commitment	Commitment	Modify	Modify	Index	Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treated	-0.065	-0.011	-0.070	-0.058	-0.084	-0.070	-0.003	0.055	-0.057	0.049
	(0.136)	(0.116)	(0.170)	(0.160)	(0.150)	(0.136)	(0.049)	(0.034)*	(0.107)	(0.090)
	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[1.000]	[0.656]		
Mean	8.648	8.648	7.883	7.883	8.330	8.330	0.687	0.687	-0.000	-0.000
N	366	366	366	366	366	366	366	366	366	366
R^2	0.001	0.276	0.000	0.153	0.001	0.254	0.000	0.544	0.001	0.362
Controls	No	No	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. Commitment to academic goal is respondent's level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. Goal Index is an Anderson Index created using the above 4 variables measured post intervention at the time of baseline. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. Values in squared brackets represent sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A4 Robustness checks

In this section, we discuss several robustness checks conducted to estimate treatment effects on primary and secondary outcomes.

First, we re-estimate equations 9 and 3 with controls selected from the list on which we show balance in Table 1 by the 'post-double-selection' with LASSO (Belloni et al., 2014a,b), as specified in the pre-analysis (Haroon et al., 2020). We find the results are qualitatively and statistically similar to those discussed in section 5. Treatment effects are significant for grit measured one year later after the treatment was first implemented. However, we see no other significant changes in other average outcomes - such as their performance in end-of-year examinations, important, commitment and willing to modify goals, for treated groups.

Table A10: Impact of treatment on end-of-year test performance, using PDS Lasso

	Test score percentage
	(1)
Treated	0.023
	(0.030)
Mean	0.496
N	330

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects.Test score percentage' are the marks scored in end-of-year examination as a percentage of the maximum (full) marks allowed for that enrollment year. These were measured 2 months after the baseline activities and are present for only one round of the data. N refers to the sample size. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Robustness check: Long term impact of treatment on primary outcomes, using PDS Lasso

	Grit	Growth mindset	Self efficacy	Mindset index
	(1)	(2)	(3)	(4)
Treated	-0.954	0.897	0.191	0.019
	(0.574)*	(0.658)	(0.500)	(0.103)
Post	-0.689	-2.753	0.047	-0.336
	(0.516)	(0.707)***	(0.473)	(0.092)***
T*Post	1.753	0.291	-0.219	0.172
	(0.739)**	(0.973)	(0.685)	(0.134)
Mean	41.125	48.660	32.938	-0.143
N	666	666	666	666

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. T*Post is the average treatment effect of our growth mindset intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A12: Immediate impact of treatment on secondary outcomes, using PDS Lasso

Goal:	Importance	Achieve	Commitment	Modify	Index
	(1)	(2)	(3)	(4)	(5)
Treated	-0.065	-0.070	-0.084	-0.003	-0.057
	(0.136)	(0.170)	(0.150)	(0.049)	(0.107)
Mean	8.648	7.883	8.330	0.687	-0.000
N	366	366	366	366	366

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. Commitment to academic goal is respondent's level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. Goal Index is an Anderson Index created using the above 4 variables measured post intervention at the time of baseline. Treated is the immediate effect of our treatment post intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Second, we also re-estimate equation 9 on only the balanced sub-sample of 226 students who participated in face-to-face survey (and repeat intervention) at the time of the baseline and followup. The longer term results remain qualitatively similar to the main regressions discussed earlier.

Table A13: Long term impact of treatment on primary outcomes, restricted to sample participating in face-to-face surveys

	Grit	Grit	Growth mindset	Growth mindset	Self efficacy	Self efficacy	Mindset index	Mindset index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-1.027	-1.079	1.192	1.001	0.058	0.156	0.035	0.017
	(0.654)	(0.655)*	(0.692)*	(0.680)	(0.546)	(0.537)	(0.106)	(0.109)
	[0.211]	[0.269]	[0.211]	[0.269]	[0.440]	[0.347]		
Post	-0.503	-0.512	-2.075	-2.114	-0.150	-0.093	-0.221	-0.225
	(0.548)	(0.534)	(0.724)***	(0.725)***	(0.513)	(0.505)	(0.095)**	(0.096)**
	[0.561]	[0.509]	[0.013]**	[0.013]**	[1.000]	[1.000]		
T*Post	1.791	1.875	-0.678	-0.798	0.029	-0.007	0.097	0.091
	(0.787)**	(0.782)**	(1.037)	(1.043)	(0.745)	(0.743)	(0.142)	(0.143)
	[0.075]*	[0.054]*	[1.000]	[0.799]	[1.000]	[1.000]		
Mean	41.193	41.193	49.105	49.105	32.849	32.849	-0.009	-0.009
N	592	592	592	592	592	592	592	592
R^2	0.007	0.041	0.035	0.093	0.000	0.048	0.020	0.069
Controls	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. T*Post is the average treatment effect of our growth mindset intervention. 'Mean' is the average value of the outcome for the control group. N refers to the sample size. Values in squared brackets represent sharpened $q - values$ that control for false discovery rate (FDR) (Anderson, 2008). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Robustness check: Long term impact of treatment on primary outcomes, restricted to sample participating in face-to-face surveys using PDS Lasso

	Grit	Growth mindset	Self efficacy	Mindset index
	(1)	(2)	(3)	(4)
Treated	-0.954	0.897	0.191	0.019
	(0.574)*	(0.658)	(0.500)	(0.103)
Post	-0.689	-2.753	0.047	-0.336
	(0.516)	(0.707)***	(0.473)	(0.092)***
T*Post	1.753	0.291	-0.219	0.172
	(0.739)**	(0.973)	(0.685)	(0.134)
Mean	41.125	48.660	32.938	-0.143
N	666	666	666	666

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy measured at the time of follow-up. T*Post is the average treatment effect of our growth mindset intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. Values in squared brackets represent sharpened $q-values$ that control for false discovery rate (FDR) (Anderson, 2008).* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A5 Heterogeneity

We explore heterogeneity in self-efficacy scores (Table A15), Mindset Index (Table A16) and secondary outcomes (Tables A18 - A20) by the following characteristics: 'HH assets' is an index from number of household assets, 'Mother education' is the years of education of respondents' mother, 'Time in hh chores' and 'Time in study' are the numbers of hours a day devoting to household chores and study, respectively, 'Career goal importance' is the self-reported importance given to having a career, 'Ravens score' is the respondents' score in Raven's progressive matrices, 'Competitiveness' is the respondents score in (Ryckman et al., 1996) scale, 'First division' is if the respondent secured highest (first) grade in last exams before joining college, 'First year' is a binary variable equal to 1 if the respondent is enrolled in the first year of college, and 'Discuss work' is a binary variable equal to 1 if the respondent discusses studies and assignments with parents.

As discussed in section 6, we find treatment effects on self-efficacy and commitment to academic goals differs by education level of mothers. We also find treatment effects on commitment to goals to vary by hours in a day spent studying at baseline and whether the respondent discusses their work from college with parents. Finally, the increase in willingness to modify goals to SMART among the treatment sample seems to be driven by respondents with high scores in the Raven's test.

Table A15: Heterogeneity in treatment impact in Self Efficacy

Independent variable: Self Efficacy					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.015	0.387	0.042	-0.101	0.548
	(0.741)	(0.716)	(0.682)	(0.790)	(0.720)
Post	0.086	0.066	0.068	0.057	0.046
	(0.465)	(0.473)	(0.472)	(0.471)	(0.474)
H	0.368	0.493	0.259	1.010	0.564
	(0.666)	(0.673)	(0.651)	(0.665)	(0.663)
H*Treated	0.097	-0.701	0.078	0.348	-1.166
	(1.033)	(1.034)	(1.078)	(1.024)	(1.037)
Treated*Post	-0.039	-0.938	0.008	0.509	-0.537
	(0.908)	(0.862)	(0.748)	(0.897)	(0.808)
H*treated*Post	-0.206	1.746	-0.389	-1.291	1.061
	(1.000)	(0.978)*	(1.078)	(0.981)	(0.996)
Mean	32.938	32.938	32.938	32.938	32.938
N	666	666	666	666	666
R^2	0.001	0.006	0.001	0.009	0.003
Independent variable: Self Efficacy					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	0.037	0.190	0.242	-1.052	-1.367
	(0.703)	(0.920)	(0.611)	(0.934)	(1.079)
Post	0.041	0.075	0.074	0.066	0.053
	(0.467)	(0.475)	(0.473)	(0.472)	(0.473)
H	2.341	-0.336	-0.010	-0.088	-0.944
	(0.647)***	(0.713)	(0.761)	(0.152)	(0.725)
H*Treated	0.208	-0.184	-0.465	0.328	2.015
	(1.004)	(1.111)	(1.137)	(0.227)	(1.199)*
Treated*Post	0.250	0.526	-0.384	0.560	-0.196
	(0.860)	(0.983)	(0.770)	(1.1011)	(1.191)
H*treated*Post	-0.833	-1.001	0.619	-0.227	0.114
	(0.980)	(1.052)	(1.038)	(0.213)	(1.213)
Mean	32.938	32.938	32.938	32.938	32.938
N	666	662	666	666	666
R^2	0.048	0.005	0.001	0.004	0.009

Note: Standard errors in parentheses. All errors are clustered at the individual level. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale (Schwarzer and Jerusalem, 1995). H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Heterogeneity in treatment impact in Mindset Index

Independent variable: Mindset Index					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	0.022	0.206	-0.107	0.060	0.057
	(0.147)	(0.136)	(0.126)	(0.145)	(0.137)
Post	-0.325	-0.328	-0.321	-0.329	-0.320
	(0.090)***	(0.090)***	(0.090)***	(0.090)***	(0.090)***
H	-0.068	0.320	-0.080	0.347	-0.057
	(0.125)	(0.121)***	(0.124)	(0.119)***	(0.123)
H*Treated	0.021	-0.377	0.371	-0.055	-0.072
	(0.198)	(0.195)*	(0.206)*	(0.193)	(0.196)
Treated*Post	0.073	-0.077	0.209	0.345	0.182
	(0.178)	(0.160)	(0.144)	(0.172)**	(0.163)
H*treated*Post	0.183	0.534	-0.138	-0.359	-0.051
	(0.197)	(0.192)***	(0.213)	(0.192)*	(0.191)
Mean	-0.143	-0.143	-0.143	-0.143	-0.143
N	666	666	666	666	666
R^2	0.020	0.043	0.025	0.038	0.021

Independent variable: Mindset Index					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.072	0.071	0.069	-0.245	-0.163
	(0.133)	(0.186)	(0.123)	(0.192)	(0.171)
Post	-0.327	-0.327	-0.326	-0.323	-0.325
	(0.090)***	(0.090)***	(0.090)***	(0.090)***	(0.090)***
H	0.272	0.168	0.274	-0.007	-0.101
	(0.121)**	(0.124)	(0.142)*	(0.027)	(0.121)
H*Treated	0.241	-0.027	-0.157	0.080	0.273
	(0.192)	(0.214)	(0.212)	(0.046)*	(0.205)
Treated*Post	0.292	0.369	0.111	0.310	0.158
	(0.171)*	(0.194)*	(0.149)	(0.208)	(0.200)
H*treated*Post	-0.278	-0.303	0.138	-0.047	0.013
	(0.192)	(0.208)	(0.205)	(0.045)	(0.213)
Mean	-0.143	-0.143	-0.143	-0.143	-0.143
N	666	662	666	666	666
R^2	0.048	0.023	0.030	0.025	0.022

Note: Standard errors in parentheses. All errors are clustered at the individual level.Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Heterogeneity in treatment impact on importance of achieving academic goal

Independent variable: Importance of academic goal					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.042	0.073	-0.079	-0.185	-0.188
	(0.193)	(0.182)	(0.178)	(0.181)	(0.181)
H	-0.220	-0.122	0.109	1.393	-0.081
	(0.189)	(0.189)	(0.189)	(0.157)***	(0.193)
H*Treated	-0.019	-0.292	0.053	0.276	0.296
	(0.273)	(0.272)	(0.276)	(0.219)	(0.272)
Mean	8.648	8.648	8.648	8.648	8.648
N	366	366	366	366	366
R^2	0.008	0.015	0.003	0.349	0.004

Independent variable: Importance of academic goal					
H:	Compet-itiveness (6)	First division (7)	First year (8)	Time in study (9)	Discuss work (10)
Treated	0.024	0.238	0.004	-0.135	-0.249
	(0.188)	(0.252)	(0.167)	(0.274)	(0.256)
H	0.193	0.411	-0.293	0.050	-0.400
	(0.188)	(0.213)*	(0.198)	(0.043)	(0.208)*
H*Treated	-0.183	-0.386	-0.112	0.018	0.246
	(0.273)	(0.300)	(0.283)	(0.069)	(0.302)
Mean	8.648	8.648	8.648	8.648	8.648
N	366	363	366	366	366
R^2	0.003	0.010	0.018	0.009	0.011

Note: Standard errors in parentheses. All errors are clustered at the individual level. Important academic goal is the respondent's self-reported importance of achieving academic goal using a Likert scale from 1 to 10measured post-intervention at the time of baseline activities. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. H*Treated is the post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Heterogeneity in treatment impact on likelihood of achieving academic goal

Independent variable: Achieve academic goal					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.124	0.084	-0.142	0.145	-0.229
	(0.257)	(0.260)	(0.211)	(0.209)	(0.226)
H	-0.231	0.445	-0.126	1.418	-0.133
	(0.251)	(0.246)*	(0.260)	(0.231)***	(0.259)
H*Treated	0.131	-0.328	0.185	-0.411	0.382
	(0.339)	(0.333)	(0.357)	(0.316)	(0.340)
Mean	7.883	7.883	7.883	7.883	7.883
N	366	366	366	366	366
R^2	0.003	0.010	0.001	0.143	0.004
Independent variable: Achieve academic goal					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.108	0.248	-0.027	0.023	-0.325
	(0.234)	(0.339)	(0.212)	(0.338)	(0.322)
H	0.161	0.548	-0.063	0.078	-0.443
	(0.253)	(0.283)*	(0.267)	(0.052)	(0.292)
H*Treated	0.096	-0.402	-0.231	-0.030	0.346
	(0.341)	(0.391)	(0.354)	(0.082)	(0.379)
Mean	7.883	7.883	7.883	7.883	7.883
N	366	363	366	366	366
R^2	0.005	0.012	0.005	0.008	0.008

Note: Standard errors in parentheses. All errors are clustered at the individual level. Achieve academic goal is respondent's self-reported likelihood of being able to achieve academic goal set on a scale from 1 to 10 measured post-intervention at the time of baseline activities. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. H*Treated is the post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A19: Heterogeneity in treatment impact on commitment to academic goal

Independent variable: Commitment to academic goal					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.015	0.178	-0.031	0.041	-0.120
	(0.208)	(0.222)	(0.186)	(0.192)	(0.204)
H	-0.098	0.478	-0.186	1.311	-0.016
	(0.220)	(0.214)**	(0.228)	(0.197)***	(0.221)
H*Treated	-0.119	-0.556	-0.174	-0.228	0.089
	(0.300)	(0.296)*	(0.312)	(0.274)	(0.300)
Mean	8.330	8.330	8.330	8.330	8.330
N	366	366	366	366	366
R^2	0.004	0.015	0.010	0.175	0.001

Independent variable: Commitment to academic goal					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.097	0.147	-0.024	0.084	-0.511
	(0.204)	(0.292)	(0.185)	(0.3289)	(0.273)*
H	0.074	0.458	-0.012	0.103	-0.558
	(0.221)	(0.249)*	(0.238)	(0.046)***	(0.233)**
H*Treated	0.035	-0.273	-0.149	-0.053	0.589
	(0.302)	(0.340)	(0.319)	(0.074)	(0.326)*
Mean	8.330	8.330	8.330	8.330	8.330
N	366	363	366	366	366
R^2	0.002	0.012	0.002	0.016	0.016

Note: Standard errors in parentheses. All errors are clustered at the individual level. Commitment to academic goal is respondent's self-reported level of commitment to achieve academic goal using a Likert scale from 0 to 10 measured post-intervention at the time of baseline activities. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. H*Treated is the post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A20: Heterogeneity in treatment effect on willingness to modify goals

Independent variable: Willing to modify goals					
H:	HH assets	Mother education	Time in hh chores	Career goal importance	Raven's score
	(1)	(2)	(3)	(4)	(5)
Treated	-0.066	-0.028	0.017	-0.049	-0.090
	(0.068)	(0.067)	(0.060)	(0.063)	(0.065)
H	-0.091	-0.016	-0.050	-0.182	-0.036
	(0.070)	(0.069)	(0.071)	(0.068)***	(0.070)
H*Treated	0.132	0.054	-0.063	0.090	0.213
	(0.097)	(0.097)	(0.101)	(0.096)	(0.096)**
Mean	0.687	0.687	0.687	0.687	0.687
N	366	366	366	366	366
R^2	0.006	0.001	0.008	0.024	0.019
Independent variable: Willing to modify goals					
H:	Compet-itiveness	First division	First year	Time in study	Discuss work
	(6)	(7)	(8)	(9)	(10)
Treated	-0.016	0.101	0.058	-0.097	-0.012
	(0.068)	(0.090)	(0.061)	(0.091)	(0.094)
H	0.050	0.033	0.117	0.008	0.075
	(0.069)	(0.081)	(0.071)*	(0.014)	(0.080)
H*Treated	0.033	-0.157	-0.177	0.027	0.017
	(0.097)	(0.107)	(0.100)*	(0.022)	(0.109)
Mean	0.687	0.687	0.687	0.687	0.687
N	366	363	366	366	366
R^2	0.005	0.008	0.009	0.011	0.007

Note: Standard errors in parentheses. All errors are clustered at the individual level. Modified academic goal is a binary variable equal to 1 if the respondent was willing to modify their goals into SMART goals with the help of the enumerators. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. H*Treated is the post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A6 The Covid-19 lockdown for educational institutions in Punjab, Pakistan

On 7 March, 2020, sample colleges closed for spring break. Before they could re-open, on 14 March, the Higher Education Commission implemented a complete closure of educational institutes across the country till 5 April in response to the global Covid-19 pandemic. All institutions will remain closed till 31 May, 2020. On 23 March, the Government of Punjab announced a complete lockdown due to the covid-19 pandemic. Finally, on 26 March, the federal government announced all educational institutes will remain closed till 31 May. At the time of registration of this document, the lock-down will be in place till 30 April, 2020.[34] Before the lockdown, we had completed 61% of the surveys face-to-face. The in-person interaction involved repeat interventions with the treated and placebo participants. Please see (Haroon and Said, 2019) for experiment details including description of the sample and the intervention.

[34] All schools and colleges were closed down under the Government of Punjab's Notification No. SO(A-II)1-1/2001(P). The date for reopening schools and colleges was extended to May 31 by order of the Government of Pakistan, Notification No.F.1-1/2020-FEPT. The general lockdown in Punjab was implemented by a special order (NO(IS-III)1-1/2004) under section 144 of the national Code of Criminal Procedure 1898.

Appendix: For Online Publication

OA6.1 Upper and Lower Bound estimates using Lee (2009)

Using our binary indicator of attrition, this bounding method estimates 2.2% excess responsiveness for test scores few months after the baseline and 6.1% excess observations are trimmed. The worst case estimate is 0.363 for Grit (p = 0.702), compared to an unadjusted estimate of 1.354. Table OA1 reports these results.

Table OA1: Lee bounds estimate of primary outcomes in the immediate and longer term

Longer run effect	Grit	Growth mindset	Self Efficacy	Mindset Index
	(1)	(2)	(3)	(4)
Treated	1.354	0.451	-0.281	0.145
Lower bound	0.363	-0.588	-1.328	-0.025
	(0.949)	(1.022)	(0.854)	(0.161)
Upper bound	2.166	1.726	0.607	0.292
	(0.970)**	(1.193)	(0.786)	(0.146)**
Trimming prop.				6.1%
N	366	366	336	366

Note: This provides the bounds on change in the outcome variable over the longer run, at the time of the follow-up, when 300 respondents could be surveyed out of the 366 respondents interviewed at baseline. Standard errors in parentheses. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and Self-efficacy measured at the time of follow-up. Coefficients reported in the 'Treated' row are from a simple OLS regression on change in the outcome variable since baseline on treatment status. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA2: Lee bounds estimates of test score outcomes

	Test score percentage
Treated	0.023
Lower bound	0.018
	-0.031
Upper bound	0.038
	-0.027
Trimming prop.	2.2%
N	360

Note: Standard errors in parentheses. Test score percentage' are the marks scored in end-of-year examination as a percentage of the maximum (full) marks allowed for that enrollment year. These were measured 9 months after the baseline activities and are present for only one round of the data. Coefficients reported in the 'Treated' row are from a simple OLS regression on change in the outcome variable since baseline on treatment status. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA7 PAP Analysis

OA7.1 Description of the sample and treatment balance

We test balance by running the regression 7 for a set of baseline covariates of interest specified in the PreAnalysis Plan, denoted by y_{ki} and indexed by $k \in (1, ..., K)$, with errors clustered at the college level.

$$y_i = \beta_0 + \beta_1 . T_i + \epsilon_i \tag{7}$$

Denote T_i as a dummy variable equal to 1 if a student (i) is assigned to treatment, and 0 for assignment to control. A p-value which is greater than 0.10 (on β_1) will ensure randomization balance. An imbalance is detected when the p-value is smaller than 0.10. Results of this test are given in column (5) of Table OA3. Treatment and placebo samples are balanced on all characteristics except age,discuss education goals with parents and pursue paid job. The difference in average ages between treated (19.2 years) and placebo group (19.5 years) is statistically significant at the 1% level. 94% of control group individuals discussed their goals with the parents in comparison to 90% of those in treatment group, and this difference is significant at the 10% level. 96% of control group individuals wished to pursue a paid job in comparison to 89% of those in treatment group, and this difference is also significant at the 1% level.

At followup, we were able to successfully contact 300 students. Table OA4 shows the sample is qualitatively similar, except on except on age, enrolled in first year, average marital status, desire to work, performance in their previous degree, discuss education goals with parents and pursue a paid job. Specifically, treated students in this balanced sample (of 300) are likely to be single and younger when compared to the placebo sample; in addition, they express a lower desire to work after graduation and are less likely to have passed their previous degree in the first division. The overall test of joint significance shows the sample is well balanced by treatment status ($p - value$ = 0.14). We also test for differences in characteristics of respondents who were interviewed in the person and those interviewed via phone at end line and find no significant overall difference between the two samples ($p - value$ = 0.86). As specified in (Haroon et al., 2020), we will test for robustness of results by re-running all main regressions using controls selected by PDS Lasso (Belloni et al., 2013).

We repeated treatment and placebo interventions with 226 (61% of the individuals from the sample at

baseline), out of which 120 (53%) belonged to the treated group. The treatment status been randomly and individually assigned at baseline. We test if the sample with whom we were able to conduct in person interviews are significantly different from the sample we could only interview over the phone due to closure of colleges. As shown in column 5 of Table OA5, average characteristics between the phone and in-person sample differ by age, enrolled in first year and whether they desire to pursue paid job or not. The overall test of joint significance shows the sample is well balanced by treatment status ($p - value$ = 0.54). In analysis, we will show the robustness of our results using using controls selected by PDS Lasso (Belloni et al., 2013).

Table OA3: Full sample description and balance at baseline

	N	Mean	Median	S.Dev.	$p-value$
	(1)	(2)	(3)	(4)	(5)
Age (years)	366	19.37	19.00	1.62	0.03**
Dummy: Enrolled in first year	366	0.36	0.00	0.48	0.14
Dummy: Single	366	0.96	1.00	0.20	0.47
Father's education (years)	351	8.43	10.0	4.80	0.23
Dummy: Father is a business owner	366	0.25	0.00	0.44	0.90
Dummy: Father is a salaried worker	366	0.57	1.00	0.50	0.82
Mother's education (years)	362	6.64	8.00	5.15	0.59
Dummy: Mother is a business owner	366	0.02	0.00	0.13	0.57
Dummy: Mother is a salaried worker	366	0.05	0.00	0.23	0.26
Average monthly household income (000's)	258	35.56	25.00	33.09	0.18
Index: Household assets	366	0.00	0.30	1.39	0.60
Risk preferences (higher is more risk averse)	366	4.15	4.00	1.59	0.54
Ravens test score (out of 10)	366	4.08	4.00	2.11	0.92
Scale: Competitiveness (out of 75)	366	57.62	58.00	8.33	0.59
Dummy: Continue education after graduating	366	0.91	1.00	0.28	0.82
Dummy: Passed last degree in first division	363	0.72	1.00	0.45	0.12
Satisfaction with academic performance	366	2.37	2.00	0.91	0.96
Daily hours studying at home in average week	366	2.99	3.00	2.07	0.46
Daily hours doing household chores in average week	366	6.28	7.00	5.99	0.11
Dummy: Discusses homework with parents	366	0.71	1.00	0.46	0.21
Dummy: Discusses goals with parents	366	0.93	1.00	0.27	0.07*
Dummy: Purse paid job	366	0.92	1.00	0.27	0.01***
Dummy: Have a role model	366	0.33	0.00	0.47	0.15
F-statistics					0.11

Note: Column (5) shows p-values from the balance test specified in equation 7. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. ***p < 0.01, **p < 0.05,*p < 0.1.

Table OA4: Full sample description and balance at baseline

	N	Mean	Median	S.Dev.	$p-value$ treatment	$p-value$ type
	(1)	(2)	(3)	(4)	(5)	(6)
Age (years)	300	19.32	19.00	1.54	0.04**	0.54
Dummy: Enrolled in first year	300	0.37	0.00	0.48	0.09*	0.48
Dummy: Single	300	0.95	1.00	0.21	0.08*	0.03**
Father's education (years)	289	8.57	10.00	4.81	0.77	0.54
Dummy: Father is a business owner	300	0.24	0.00	0.43	0.51	0.60
Dummy: Father is a salaried worker	300	0.58	1.00	0.49	0.34	0.97
Mother's education (years)	297	6.69	8.00	5.19	0.86	0.02**
Dummy: Mother is a business owner	300	0.02	0.00	0.14	0.60	0.17
Dummy: Mother is a salaried worker	300	0.06	0.00	0.23	0.58	0.02**
Average monthly household income (000's)	213	36.20	25.00	35.23	0.17	0.81
Index: Household assets	300	-0.02	0.00	1.37	0.80	0.44
Risk preferences (higher is more risk averse)	300	4.15	4.00	1.59	0.41	0.03**
Ravens test score (out of 10)	300	4.16	4.00	2.11	0.44	0.29
Scale: Competitiveness (out of 75)	300	58.00	58.00	8.31	0.59	0.59
Dummy: Continue education after graduating	300	0.91	1.00	0.28	0.59	0.39
Dummy: Passed last degree in first division	299	0.74	1.00	0.44	0.01***	0.51
Satisfaction with academic performance	300	2.40	2.00	0.91	0.94	0.69
Daily hours studying at home in average week	300	2.94	3.00	2.09	0.26	0.61
Daily hours doing household chores in average week	300	6.43	7.00	6.12	0.16	0.59
Dummy: Discusses homework with parents	300	0.70	1.00	0.46	0.63	0.24
Dummy: Discusses goals with parents	300	0.91	1.00	0.28	0.09*	0.82
Dummy: Purse paid job	300	0.93	1.00	0.26	0.03**	0.30
Dummy: Have a role model	300	0.34	0.04	0.47	0.22	0.50
P-value of F-statistic (treatment)						0.14
P-value of F-statistic (type)						0.86

Note: Column (5) shows $p-values$ from the balance test specified in equation 7. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. Column (6) shows the $p-values$ from a test of baseline balance between the in-person and phone interview sample ('type' of survey), from a regression of the variable in the row on the type of interview. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. $***p<0.01, **p<0.05, *p<0.1$.

Table OA5: Sample description and balance at baseline - repeat intervention sample

	N (1)	Mean (2)	Median (3)	S.Dev. (4)	$p-value$ (5)
Age (years)	226	19.37	19.00	1.61	0.01***
Dummy: Enrolled in first year	226	0.38	0.00	0.49	0.04**
Dummy: Single	226	0.95	1.00	0.22	0.10
Father's education (years)	217	8.61	10.00	4.71	0.21
Dummy: Father is a business owner	226	0.23	0.00	0.42	0.96
Dummy: Father is a salaried worker	226	0.58	1.00	0.58	0.68
Mother's education (years)	223	6.36	8.00	5.25	0.59
Dummy: Mother is a business owner	226	0.01	0.00	0.09	0.14
Dummy: Mother is a salaried worker	226	0.07	0.00	0.26	0.83
Average monthly household income (000's)	163	35.94	25.00	38.17	0.55
Index: Household assets	226	-0.06	0.00	1.35	0.57
Risk preferences (higher is more risk averse)	226	4.25	4.00	1.55	0.20
Ravens test score (out of 10)	226	4.22	4.00	2.19	0.78
Scale: Competitiveness (out of 75)	226	58.23	58.00	8.39	0.69
Dummy: Continue education after graduating	226	0.91	1.00	0.29	0.94
Dummy: Passed last degree in first division	226	0.75	1.00	0.44	0.23
Satisfaction with academic performance	226	2.42	2.00	0.93	0.82
Daily hours studying at home in average week	226	2.92	3.00	2.12	0.37
Daily hours doing household chores in average week	226	6.56	7.00	6.19	0.31
Dummy: Discusses homework with parents	226	0.71	1.00	0.45	0.81
Dummy: Discusses goals with parents	226	0.91	1.00	0.28	0.49
Dummy: Purse paid job	226	0.94	1.00	0.24	0.04**
Dummy: Have a role model	226	0.35	0.00	0.48	0.59
F-statistic					0.54

Note: Column (5) shows $p-values$ from the balance test specified in equation 1 for 266 individuals who received the repeat treatment. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment.F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. $***p < 0.01, **p < 0.05, *p < 0.1$.

OA8 Results

We then test for the effect of our intervention on several outcome variables specified in Haroon et al. (2020). We test for both short run - immediate and post-intervention effects at baseline, and longer term effects using data from the followup surveys.

For each outcome, we run the following basic specification for immediate or post intervention effects:

$$y_i = \beta_1 . T_i + \mu_c + \epsilon_{it} \tag{8}$$

where y_i is an outcome variable, T_i is a dummy variable capturing exposure to treatment (as described earlier), μ_c denote college fixed effects. All errors are clustered at the college levels. The main hypothesis we test that exposure to the treatment, *i.e.* growth mindset discussion has no effect on y_i; $H_0 : \beta_1 = 0$.

Our main estimating specification for long term effects is:

$$y_{it} = \beta_1 Ti + \beta_2 Post_t + \beta_3 T_i * Post_t + \epsilon_{it} \tag{9}$$

where y_{it} is an outcome variable, T_i is a dummy variable capturing exposure to treatment (as described earlier), $Post_t$ is a dummy variable for the post-intervention time period. $T_i * Post_t$ is the average treatment effect of our growth mindset treatment in the post intervention time period. All errors are clustered at the college levels. The main hypothesis we propose to test is that exposure to the treatment *i.e.* growth mindset discussion has no effect; $H_0 : \beta_3 = 0$.

Next, we discuss results, for each round of the survey separately.

OA8.1 Primary outcomes

Table OA6: Immediate impact of treatment on primary outcomes

	(1) Grit	(2) Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset Index	(8) Mindset Index
Treated	-1.027	-1.039	1.192	0.938	0.058	0.177	0.029	0.014
	(0.622)*	(0.623)*	(0.611)*	(0.527)*	(0.589)	(0.663)	(0.129)	(0.130)
Mean	41.380	41.380	49.877	49.877	32.905	32.905	-0.000	-0.000
N	366	366	366	366	366	366	366	366
R^2	0.007	0.007	0.008	0.005	0.000	0.000	0.000	0.000
Controls	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured post-intervention. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured post intervention. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured post intervention. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy measured post intervention. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA7: Long term impact of treatment on primary outcomes

	(1) Grit Grit	(2) Grit Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset Index	(8) Mindset Index	(9) Test scores	(10) Test scores
Treated	-1.027	-1.112	1.192	0.987	0.058	0.163	0.029	0.011	17.085	5.703
	(0.622)*	(0.628)*	(0.611)*	(0.491)**	(0.589)	(0.611)	(0.129)	(0.126)	(21.604)	(21.244)
Post	-0.577	-0.670	-2.750	-2.798	0.074	0.104	-0.323	-0.333	0.000	
	(0.267)**	(0.292)**	(0.733)***	(0.752)***	(0.495)	(0.510)	(0.091)***	(0.092)***	(.)	
T*Post	1.648	1.727	0.244	0.303	-0.138	-0.221	0.165	0.170	0.000	
	(0.643)**	(0.663)***	(0.864)	(0.908)	(0.570)	(0.557)	(0.154)	(0.156)	(.)	
Mean	41.125	41.125	48.660	48.660	32.938	32.938	-0.143	-0.143	259.633	259.633
N	666	666	666	666	666	666	666	666	333	333
R^2	0.005	0.006	0.043	0.043	0.000	0.000	0.018	0.018	0.003	0.000
Controls	No	No	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy measured at the time of follow-up. Test scores are the scores of participants in end-of-year exams which are not measured at the time of baseline. T*Post is the average treatment effect of our growth mindset intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA8.2 Secondary outcomes

Table OA8: Immediate impact of treatment on secondary outcomes

Goal:	(1) Importance	(2) Importance	(3) Achieve	(4) Achieve	(5) Commitment	(6) Commitment	(7) Modify	(8) Modify	(9) Index	(10) Index
Treated	-0.065	-0.011	-0.070	-0.058	-0.084	-0.070	-0.003	0.055	-0.057	0.049
	(0.112)	(0.102)	(0.091)	(0.062)	(0.170)	(0.189)	(0.024)	(0.016)***	(0.108)	(0.073)
Mean	8.648	8.648	7.883	7.883	8.330	8.330	0.687	0.687	-0.000	-0.000
N	366	366	366	366	366	366	366	366	366	366
R^2	0.001	0.000	0.000	0.000	0.001	0.001	0.000	0.007	0.001	0.001
Controls	No	No	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. Commitment to academic goal is respondentâĂŹs level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. Goal Index is an Anderson Index created using the above 4 variables measured post intervention at the time of baseline. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA9: Immediate effects on secondary outcomes

	(1) Happiness	(2) Happiness	(3) Self discipline	(4) Self discipline	(5) Self esteem	(6) Self	(7) Depression	(8) Depression	(9) Wellbeing index	(10) Wellbeing index
Treated	0.471	0.471	-0.008	-0.008	-0.630	-0.630	0.677	0.677	-0.197	-0.197
	(0.798)	(0.798)	(0.525)	(0.525)	(0.282)**	(0.282)**	(0.582)	(0.582)	(0.137)	(0.137)
Mean	19.821	19.821	34.858	34.858	12.755	12.755	7.840	7.840	-0.000	-0.000
N	226	226	226	226	226	226	226	226	226	226
R^2	0.003	0.003	0.000	0.000	0.012	0.012	0.004	0.004	0.008	0.008
Controls	No	No	No	No	No	No	No	No		
FE	No	Yes	No	Yes	No	Yes	No	Yes		

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Happiness is measured using a 4-item scale ranging from a 1 to 7-point Likert scale measured post intervention at the time of follow-up. Self-discipline is a 10-item scale, each item is based on responses ranging from a 1 to 5-point Likert scale measured post intervention at the time of follow-up. Self Esteem is a 10-item scale based on responses ranging from a 0 to 4-point Likert scale measured post intervention at the time of follow-up. Depression is a 9-item scale based on responses ranging from a 0 to 3 (point Likert scale measured post intervention at the time of follow-up. Psychological well-being is an Anderson index using all variables above. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA10: Immediate effects on secondary outcomes

	(1)	(2)	(3)	(4)
Willing to learn language:	Chinese	Chinese	Regional	Regional
Treated	0.018	0.018	0.052	0.052
	(0.050)	(0.050)	(0.059)	(0.059)
Mean	0.774	0.774	0.415	0.415
N	226	226	226	226
R^2	0.000	0.000	0.003	0.003
Controls	No	No	No	No
FE	No	Yes	No	Yes

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Willingness to opt for a Chinese language is a binary variable equal to 1 if the respondent is willing to opt for a Chinese language measured post intervention at the time of follow-up. Willingness to opt for a regional language is a binary variable equal to 1 if the respondent is willing to opt for a regional language measured post intervention at the time of follow-up. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA8.3 Heterogeneity in effects

We then test for the heterogeneous treatment effect of our intervention on our primary and secondary outcomes for both short run - immediate and post-intervention effects and longer term effects using data from the followup surveys. The variables used to test for heterogeneous treatment effects are specified in Haroon and Said (2019). [35]

Table OA11: Heterogeneity in impact on Grit

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.705	-0.661	-0.590	-0.509	-0.521	0.295	-0.433	-1.651	-0.259
	(0.849)	(0.747)	(0.497)	(0.509)	(0.572)	(0.867)	(0.509)	(0.635)***	(0.569)
H	-1.113	-0.026	2.200	3.465	2.053	1.694	2.488	-0.133	3.490
	(0.950)	(0.621)	(1.036)**	(0.709)***	(1.314)	(0.683)**	(0.608)***	(0.709)	(0.820)***
H*Treated	-0.486	-1.034	-0.935	-0.899	-1.096	-1.742	-2.021	1.722	-1.502
	(0.992)	(1.292)	(1.251)	(1.061)	(1.695)	(1.226)	(1.054)*	(0.809)**	(0.978)
N	366	366	366	366	366	363	366	366	366
R^2	0.019	0.010	0.027	0.065	0.022	0.013	0.024	0.014	0.057

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.763	-1.132	-1.029	-1.708	-0.837	-1.385	-1.717	1.167	-0.493
	(0.876)	(1.319)	(0.511)**	(0.762)**	(0.536)	(0.597)**	(0.736)**	(0.944)	(0.566)
H	0.636	0.270	0.873	0.800	2.558	1.481	-0.148	0.000	0.000
	(0.299)**	(0.754)	(0.714)	(1.680)	(1.140)**	(1.460)	(1.278)	(.)	(.)
H*Treated	-0.104	0.166	-0.101	1.580	-0.198	0.879	1.641	-3.062	-1.387
	(0.349)	(1.413)	(1.460)	(1.820)	(0.996)	(1.934)	(1.913)	(0.751)***	(1.038)
N	366	366	366	366	366	366	366	366	366
R^2	0.045	0.007	0.011	0.026	0.045	0.030	0.014	0.031	0.013

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[35] Note, we record 'Stroop measures' only at follow-up and use them to test heterogeneity on the assumption that this measure of cognitive ability will not have changed due to the treatment. However, we cannot prove this and do not discuss these results in detail.

Table OA12: Heterogeneity in impact on Grit

H:	HH assets	Father Educ	Mother Educ	Competitive	Raven	First Div	First year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.496	-0.45	-0.674	-0.643	-0.945	0.371	-0.457	-1.538	-0.386
	(0.777)	(0.640)	(0.496)	(0.507)	(0.604)	(0.894)	(0.500)	(0.641)**	(0.544)
Post	-0.6	-0.586	-0.611	-0.622	-0.631	-0.658	-0.604	-0.58	-0.632
	(0.266)**	(0.269)**	(0.267)**	(0.263)**	(0.247)**	(0.269)**	(0.280)**	(0.273)**	(0.299)**
T*Post	1.227	1.293	1.783	2.65	2.52	1.485	0.808	1.603	2.668
	(0.901)	(0.589)**	(0.786)**	(0.971)***	(0.721)***	(1.127)	(0.826)	(0.609)***	(0.829)***
H	-0.662	0.626	2.02	3.191	1.082	1.795	2.412	0.145	3.236
	(0.807)	(0.828)	(0.830)**	(0.511)***	(1.289)	(0.594)***	(0.554)***	(0.655)	(0.890)***
H*Treated	-0.937	-1.687	-0.755	-0.624	-0.125	-1.843	-1.945	1.445	-1.248
	(0.906)	(1.451)	(1.270)	(0.875)	(1.739)	(1.181)	(0.924)**	(0.701)**	(1.006)
H*Treated*Post	0.83	1.03	-0.195	-2.119	-2.032	0.284	2.129	0.037	-1.983
	(0.800)	(0.687)	(0.593)	(0.867)**	(0.669)***	(1.423)	(1.613)	(1.133)	(0.698)***
N	666	666	666	666	666	662	666	666	666
R^2	0.012	0.008	0.023	0.049	0.012	0.012	0.031	0.013	0.045

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-1.198	-1.124	-0.734	-1.673	-1.048	-1.481	-1.367	0.293	-0.494
	(0.811)	(1.294)	(0.526)	(0.696)**	(0.585)*	(0.512)***	(0.645)**	(0.806)	(0.561)
Post	-0.527	-0.571	-0.58	-0.601	-0.589	-0.581	-0.561	-0.577	-0.577
	(0.294)*	(0.263)**	(0.249)**	(0.266)**	(0.267)**	(0.270)**	(0.266)**	(0.267)**	(0.267)**
T*Post	2.914	1.478	2.035	1.609	2.216	1.943	2.392	1.552	1.492
	(1.248)**	(1.164)	(0.707)***	(0.608)***	(1.141)*	(0.694)***	(0.986)**	(0.637)**	(0.614)**
H	0.545	0.281	1.586	0.891	2.142	1.218	0.723	0.000	0.000
	(0.282)*	(0.697)	(0.653)**	(1.108)	(0.894)**	(0.994)	(1.049)	(.)	(.)
H*Treated	0.027	0.155	-0.814	1.489	0.218	1.142	0.77	-1.842	-1.385
	(0.368)	(1.400)	(1.469)	(1.285)	(0.790)	(1.493)	(1.588)	(0.597)***	(0.817)*
H*Treated*Post	-0.429	0.237	-0.813	0.118	-1.175	-0.867	-1.812	0.000	0.000
	(0.282)	(1.510)	(0.980)	(1.129)	(1.310)	(0.700)	(1.375)	(.)	(.)
N	666	666	666	666	666	666	666	666	666
R^2	0.031	0.006	0.014	0.027	0.032	0.022	0.011	0.015	0.011

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$)

Table OA13: Heterogeneity in impact on Growth Mindset

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.649	0.835	2.219	-0.263	0.424	-0.792	0.951	0.052	1.188
	(1.022)	(0.717)	(0.944)**	(0.959)	(0.948)	(1.074)	(0.709)	(0.732)	(0.675)*
H	-0.512	-0.711	0.635	-2.691	-3.259	-0.991	0.725	-1.828	-0.225
	(0.724)	(1.224)	(0.541)	(0.710)***	(0.958)***	(1.373)	(1.496)	(0.896)**	(0.663)
H*Treated	1.095	1.072	-2.183	3.016	1.656	2.927	0.457	2.929	0.004
	(1.588)	(1.225)	(1.246)*	(1.194)**	(1.401)	(1.496)*	(1.468)	(1.278)**	(0.903)
N	366	366	366	366	366	363	366	366	366
R^2	0.010	0.010	0.016	0.029	0.044	0.021	0.013	0.020	0.008

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-1.132	1.005	0.932	1.114	0.066	1.204	0.452	-0.066	1.158
	(1.471)	(0.968)	(0.680)	(1.218)	(0.995)	(0.892)	(1.005)	(0.949)	(0.657)*
H	-0.638	0.027	-1.610	-1.092	-2.119	-0.822	-0.747	0.000	0.000
	(0.269)**	(1.053)	(0.866)*	(1.821)	(0.667)***	(0.972)	(1.457)	(.)	(.)
H*Treated	0.705	0.273	0.749	0.289	2.224	0.006	1.788	1.756	0.090
	(0.330)*	(1.478)	(0.720)	(1.981)	(1.282)*	(1.364)	(2.186)	(1.146)	(1.075)
N	366	366	366	366	366	366	366	366	366
R^2	0.032	0.008	0.017	0.013	0.021	0.012	0.013	0.015	0.008

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA14: Heterogeneity in impact on Growth Mindset

H:	HH assets	Father Educ	Mother Educ	Competitive	Raven	First Div	First year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.697	1.052	2.317	-0.151	0.927	0.189	0.882	0.318	1.05
	(0.886)	(0.659)	(0.887)***	(1.141)	(0.894)	(0.844)	(0.757)	(0.761)	(0.552)*
Post	-2.764	-2.75	-2.764	-2.716	-2.645	-2.721	-2.756	-2.725	-2.742
	(0.733)***	(0.729)***	(0.733)***	(0.728)***	(0.724)***	(0.694)***	(0.732)***	(0.734)***	(0.737)***
T*Post	-0.376	-0.249	-1.539	0.211	-0.023	1.808	0.718	0.591	0.48
	(1.129)	(0.705)	(0.787)*	(1.259)	(1.180)	(1.436)	(0.954)	(1.114)	(0.569)
H	-0.41	-0.042	0.844	-2.46	-2.103	0.321	0.504	-1.176	-0.503
	(0.850)	(0.805)	(0.474)*	(0.883)***	(0.613)***	(1.047)	(0.904)	(0.898)	(0.455)
H*Treated	0.993	0.403	-2.392	2.785	0.501	1.614	0.678	2.277	0.282
	(1.608)	(0.825)	(0.992)**	(1.397)**	(1.191)	(1.419)	(1.085)	(1.443)	(0.999)
H*Treated*Post	1.251	1.523	3.862	-0.012	0.375	-2.344	-1.183	-1.042	-0.521
	(1.034)	(2.591)	(1.881)**	(1.027)	(0.910)	(2.816)	(1.086)	(1.230)	(1.009)
N	666	666	666	666	666	662	666	666	666
R^2	0.048	0.047	0.055	0.058	0.059	0.049	0.046	0.048	0.045

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.334	0.71	1.282	1.311	0.205	1.426	0.574	-0.242	0.924
	(1.130)	(0.765)	(0.677)*	(1.029)	(0.916)	(0.768)*	(0.946)	(0.894)	(0.713)
Post	-2.796	-2.759	-2.749	-2.735	-2.74	-2.75	-2.76	-2.75	-2.75
	(0.748)***	(0.732)***	(0.722)***	(0.729)***	(0.735)***	(0.733)***	(0.722)***	(0.733)***	(0.733)***
T*Post	-0.147	0.386	-0.899	0.846	0.72	0.785	1.269	0.348	0.323
	(1.078)	(1.280)	(0.823)	(1.142)	(1.525)	(1.248)	(1.090)	(0.850)	(0.812)
H	-0.491	-0.379	-0.762	-0.565	-1.846	-0.21	-0.444	0.000	0.000
	(0.161)***	(0.754)	(0.974)	(1.128)	(0.585)***	(0.723)	(1.098)	(.)	(.)
H*Treated	0.466	0.68	-0.099	-0.237	1.951	-0.606	1.486	2.002	0.698
	(0.223)**	(1.129)	(0.961)	(1.425)	(1.039)*	(1.068)	(1.793)	(0.942)**	(0.972)
H*Treated*Post	0.140	-0.194	2.372	-1.466	-1.037	-1.286	-2.461	0000	0.000
	(0.225)	(1.943)	(1.918)	(0.745)**	(1.522)	(1.311)	(0.966)**	(.)	(.)
N	666	666	666	666	666	666	666	666	666
R^2	0.005	0.044	0.048	0.051	0.053	0.049	0.048	0.052	0.044

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA15: Heterogeneity in impact on Self Efficacy

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.541	-0.240	0.250	0.076	0.484	-0.110	0.175	-0.196	0.002
	(1.085)	(0.494)	(0.529)	(0.657)	(0.607)	(0.815)	(0.741)	(0.703)	(0.604)
H	-0.766	-0.778	0.201	2.421	0.418	-0.737	-0.226	-0.325	1.217
	(1.067)	(0.814)	(0.808)	(0.678)***	(0.497)	(0.542)	(0.967)	(0.555)	(0.549)**
H*Treated	1.231	0.906	-0.410	0.128	-1.020	0.217	-0.249	0.663	0.140
	(1.306)	(1.242)	(0.726)	(0.655)	(0.964)	(0.520)	(0.781)	(0.685)	(0.588)
N	366	366	366	366	366	363	366	366	366
R^2	0.004	0.002	0.000	0.056	0.002	0.003	0.001	0.001	0.015

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-1.097	-1.245	-0.246	0.611	0.994	0.329	0.581	1.038	0.451
	(0.895)	(1.545)	(0.683)	(0.556)	(0.663)	(0.598)	(0.586)	(0.447)**	(0.305)
H	-0.128	-0.777	1.521	2.156	2.741	1.357	1.716	0.000	0.000
	(0.108)	(0.672)	(0.750)**	(0.978)**	(1.120)**	(0.761)*	(0.557)***	(.)	(.)
H*Treated	0.342	1.848	0.460	-1.540	-1.769	-0.774	-1.322	-1.369	-1.023
	(0.299)	(1.536)	(1.069)	(1.526)	(1.458)	(0.713)	(0.822)	(0.550)**	(1.346)
N	366	366	366	366	366	366	366	366	366
R^2	0.005	0.007	0.028	0.021	0.038	0.009	0.013	0.007	0.005

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA16: Heterogeneity in impact on Self Efficacy

H:	HH assets	Father Educ	Mother Educ	Competitive	Raven	First Div	First year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.015	-0.14	0.387	0.037	0.548	0.19	0.242	0.042	-0.101
	(1.073)	(0.415)	(0.484)	(0.535)	(0.617)	(0.794)	(0.781)	(0.713)	(0.681)
Post	0.086	0.08	0.066	0.041	0.046	0.075	0.074	0.068	0.057
	(0.490)	(0.500)	(0.492)	(0.498)	(0.497)	(0.505)	(0.500)	(0.493)	(0.493)
T*Post	-0.039	-0.29	-0.938	0.25	-0.537	0.526	-0.384	0.008	0.509
	(0.475)	(0.482)	(0.533)*	(0.668)	(0.601)	(1.150)	(0.944)	(0.800)	(0.963)
H	0.368	-0.472	0.493	2.341	0.564	-0.336	-0.01	0.259	1.01
	(1.123)	(0.582)	(0.599)	(0.497)***	(0.266)**	(0.413)	(0.784)	(0.431)	(0.554)*
H*Treated	0.097	0.6	-0.701	0.208	-1.166	-0.184	-0.465	0.078	0.348
	(1.349)	(0.943)	(0.629)	(0.690)	(1.001)	(0.463)	(0.789)	(0.689)	(0.684)
H*Treated*Post	-0.206	0.453	1.746	-0.833	1.061	-1.001	0.619	-0.389	-1.291
	(0.439)	(0.821)	(0.596)***	(0.905)	(0.788)	(0.973)	(1.016)	(0.947)	(0.941)
N	666	666	666	666	666	662	666	666	666
R^2	0.001	0.002	0.006	0.048	0.003	0.005	0.001	0.001	0.009

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-1.052	-1.367	-0.555	0.343	0.66	0.346	0.301	0.531	0.353
	(0.836)	(1.341)	(0.796)	(0.610)	(0.655)	(0.613)	(0.541)	(0.520)	(0.381)
Post	0.066	0.053	0.072	0.035	0.062	0.07	0.096	0.074	0.074
	(0.494)	(0.490)	(0.499)	(0.486)	(0.487)	(0.491)	(0.488)	(0.495)	(0.495)
T*Post	0.560	-0.196	0.552	-0.494	-0.067	-0.226	-0.031	-0.172	-0.224
	(0.334)*	(1.165)	(0.910)	(0.825)	(0.860)	(0.737)	(0.632)	(0.563)	(0.535)
H	-0.088	-0.944	0.774	1.441	2.084	1.402	1.019	0.000	0.000
	(0.094)	(0.558)*	(0.729)	(0.774)*	(0.740)***	(0.480)***	(0.527)*	(.)	(.)
H*Treated	0.328	2.015	1.207	-0.825	-1.112	-0.82	-0.626	-0.661	-0.767
	(0.270)	(1.308)	(1.189)	(1.354)	(1.119)	(0.548)	(0.839)	(0.327)**	(0.998)
H*Treated*Post	-0.227	0.114	-1.48	0.937	-0.121	0.192	-0.302	0.000	0.000
	(0.184)	(1.264)	(0.937)	(1.113)	(0.845)	(0.794)	(0.846)	(.)	(.)
N	666	666	666	666	666	666	666	666	666
R^2	0.004	0.009	0.013	0.015	0.024	0.01	0.005	0.002	0.003

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA17: Heterogeneity in impact on Psychological Index

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.052	-0.001	0.191	-0.067	0.040	-0.062	0.072	-0.167	0.095
	(0.239)	(0.111)	(0.105)*	(0.126)	(0.135)	(0.182)	(0.132)	(0.124)	(0.091)
H	-0.228	-0.147	0.288	0.282	-0.096	-0.011	0.283	-0.227	0.418
	(0.161)	(0.160)	(0.144)**	(0.087)***	(0.086)	(0.152)	(0.182)	(0.153)	(0.137)***
H*Treated	0.181	0.097	-0.345	0.231	-0.032	0.152	-0.166	0.517	-0.126
	(0.246)	(0.215)	(0.123)***	(0.179)	(0.149)	(0.207)	(0.183)	(0.202)**	(0.190)
N	366	366	366	366	366	363	366	366	366
R^2	0.006	0.003	0.010	0.041	0.003	0.003	0.010	0.015	0.031

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.288	-0.122	-0.026	0.010	0.023	0.023	-0.060	0.200	0.113
	(0.207)	(0.231)	(0.099)	(0.146)	(0.133)	(0.124)	(0.098)	(0.089)**	(0.069)*
H	-0.017	-0.046	0.064	0.169	0.285	0.184	0.074	0.000	0.000
	(0.056)	(0.105)	(0.124)	(0.149)	(0.190)	(0.142)	(0.196)	(.)	(.)
H*Treated	0.093	0.218	0.109	0.031	0.037	0.009	0.207	-0.239	-0.217
	(0.073)	(0.228)	(0.214)	(0.211)	(0.227)	(0.166)	(0.322)	(0.144)*	(0.240)
N	366	366	366	366	366	366	366	366	366
R^2	0.011	0.003	0.004	0.008	0.022	0.008	0.010	0.006	0.006

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Psychological mindset index is an Anderson index constructed from Grit, Growth Mindset and Self-efficacy variables measured at the baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA18: Heterogeneity in impact on Psychological Index

H:	HH assets	Father Educ	Mother Educ	Competitive	Raven	First Div	First year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.022	0.05	0.206	-0.072	0.057	0.071	0.069	-0.107	0.06
	(0.220)	(0.092)	(0.087)**	(0.145)	(0.133)	(0.161)	(0.139)	(0.114)	(0.110)
Post	-0.325	-0.323	-0.328	-0.327	-0.32	-0.327	-0.326	-0.321	-0.329
	(0.091)***	(0.091)***	(0.091)***	(0.090)***	(0.093)***	(0.088)***	(0.089)***	(0.090)***	(0.088)***
T*Post	0.073	0.068	-0.077	0.292	0.182	0.369	0.111	0.209	0.345
	(0.181)	(0.102)	(0.097)	(0.226)	(0.182)	(0.151)**	(0.122)	(0.152)	(0.151)**
H	-0.068	0.01	0.32	0.272	-0.057	0.168	0.274	-0.08	0.347
	(0.126)	(0.061)	(0.116)***	(0.088)***	(0.113)	(0.129)	(0.130)**	(0.135)	(0.121)***
H*Treated	0.021	-0.06	-0.377	0.241	-0.072	-0.027	-0.157	0.371	-0.055
	(0.217)	(0.223)	(0.150)**	(0.207)	(0.190)	(0.167)	(0.158)	(0.202)*	(0.217)
H*Treated*Post	0.183	0.292	0.534	-0.278	-0.051	-0.303	0.138	-0.138	-0.359
	(0.126)	(0.282)	(0.202)***	(0.172)	(0.153)	(0.319)	(0.166)	(0.159)	(0.123)***
N	666	666	666	666	666	662	666	666	666
R^2	0.02	0.021	0.043	0.048	0.021	0.023	0.03	0.025	0.038

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.245	-0.163	0.008	0.008	-0.015	0.038	-0.041	0.053	0.08
	(0.192)	(0.225)	(0.090)	(0.144)	(0.139)	(0.117)	(0.095)	(0.095)	(0.091)
Post	-0.323	-0.325	-0.323	-0.327	-0.324	-0.323	-0.32	-0.323	-0.323
	(0.089)***	(0.091)***	(0.092)***	(0.088)***	(0.091)***	(0.090)***	(0.089)***	(0.091)***	(0.091)***
T*Post	0.310	0.158	0.152	0.188	0.272	0.238	0.347	0.163	0.15
	(0.196)	(0.230)	(0.110)	(0.146)	(0.275)	(0.208)	(0.200)*	(0.149)	(0.138)
H	-0.007	-0.101	0.144	0.163	0.212	0.225	0.119	0.000	0.000
	(0.043)	(0.092)	(0.119)	(0.105)	(0.137)	(0.108)**	(0.157)	(.)	(.)
H*Treated	0.008	0.273	0.028	0.037	0.111	-0.032	0.162	-0.033	-0.132
	(0.068)	(0.216)	(0.209)	(0.212)	(0.179)	(0.110)	(0.278)	(0.104)	(0.175)
H*Treated*Post	-0.047	0.013	0.022	-0.047	-0.225	-0.192	-0.444	0.000	0.000
	(0.020)**	(0.248)	(0.202)	(0.133)	(0.293)	(0.194)	(0.162)***	(.)	(.)
N	666	666	666	666	666	666	666	666	666
R^2	0.025	0.022	0.024	0.025	0.031	0.026	0.026	0.018	0.02

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Psychological mindset index is an Anderson index constructed from grit, growth mindset and self-efficacy variables measured at the time of follow-up. H*Treated*Post is the long-term heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA19: Heterogeneity in impact on Importance of Goal

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.042	0.132	0.073	0.024	-0.188	0.238	0.004	-0.079	-0.185
	(0.149)	(0.169)	(0.155)	(0.147)	(0.179)	(0.118)**	(0.133)	(0.116)	(0.137)
H	-0.22	0.113	-0.122	0.193	-0.081	0.411	-0.293	0.109	1.393
	(0.166)	(0.225)	(0.156)	(0.070)***	(0.157)	(0.143)***	(0.224)	(0.14)	(0.109)***
H*Treated	-0.019	-0.568	-0.292	-0.183	0.296	-0.386	-0.112	0.053	0.276
	(0.175)	(0.432)	(0.324)	(0.17)	(0.267)	(0.141)***	(0.156)	(0.208)	(0.171)
N	366	366	366	366	366	363	366	366	366
R^2	(0.008)	(0.016)	(0.015)	(0.003)	(0.004)	(0.01)	(0.018)	(0.003)	(0.349)

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.135	-0.249	-0.134	-0.197	0.092	-0.132	-0.097	-0.195	-0.1
	(0.136)	(0.236)	(0.159)	(0.131)	(0.202)	(0.2)	(0.165)	(0.2)	(0.191)
H	0.050	-0.4	0.07	-0.034	0.224	-0.003	0.31	0	0
	(0.034)	(0.154)***	(0.194)	(0.202)	(0.127)*	(0.249)	(0.178)*	(.)	(.)
H*Treated	0.018	0.246	0.138	0.323	-0.316	0.176	0.062	0.181	0.091
	(0.040)	(0.242)	(0.19)	(0.150)**	(0.235)	(0.444)	(0.248)	(0.218)	(0.252)
N	366	366	366	366	366	366	366	366	366
R^2	0.009	0.011	0.004	0.007	0.005	0.003	0.017	0.003	0.001

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA20: Heterogeneity in impact on Achieve Academic Goal

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.124	0.132	0.084	-0.108	-0.229	0.248	0.027	-0.142	0.145
	(0.113)	(0.137)	(0.109)	(0.17)	(0.111)**	(0.358)	(0.082)	(0.151)	(0.217)
H	-0.231	0.556	0.445	0.161	-0.133	0.548	-0.063	-0.126	1.418
	(0.193)	(0.162)***	(0.119)***	(0.184)	(0.305)	(0.377)	(0.163)	(0.172)	(0.164)***
H*Treated	0.131	-0.618	-0.328	0.096	0.382	-0.402	-0.231	0.185	-0.411
	(0.162)	(0.338)*	(0.163)**	(0.248)	(0.209)*	(0.499)	(0.120)*	(0.264)	(0.439)
N	366	366	366	366	366	363	366	366	366
R^2	0.003	0.013	0.010	0.005	0.004	0.012	0.005	0.001	0.143

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.023	-0.325	-0.242	-0.337	-0.006	-0.004	0.011	0.042	-0.065
	(0.132)	(0.32)	(0.117)**	(0.134)**	(0.157)	(0.112)	(0.165)	(0.114)	(0.19)
H	0.078	-0.443	0.108	-0.218	0.373	0.377	0.289	0	0
	(0.038)**	(0.187)**	(0.117)	(0.19)	(0.307)	(0.124)***	(0.321)	(.)	(.)
H*Treated	-0.030	0.346	0.353	0.669	-0.105	-0.189	-0.206	-0.156	-0.012
	(0.028)	(0.363)	(0.176)**	(0.172)***	(0.235)	(0.33)	(0.31)	(0.174)	(0.32)
N	366	366	366	366	366	366	366	366	366
R^2	0.008	0.008	0.011	0.012	0.010	0.008	0.005	0.001	0.000

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA21: Heterogeneity in impact on Commitment to Academic Goal

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.015	0.008	0.178	-0.097	-0.12	0.147	-0.024	-0.031	0.041
	(0.287)	(0.197)	(0.256)	(0.234)	(0.298)	(0.156)	(0.219)	(0.167)	(0.16)
H	-0.098	0.201	0.478	0.074	-0.016	0.458	-0.012	-0.186	1.311
	(0.264)	(0.138)	(0.189)**	(0.11)	(0.178)	(0.187)**	(0.174)	(0.227)	(0.207)***
H*Treated	-0.119	-0.277	-0.556	0.035	0.089	-0.273	-0.149	-0.174	-0.228
	(0.314)	(0.24)	(0.377)	(0.236)	(0.351)	(0.162)*	(0.264)	(0.207)	(0.177)
N	366	366	366	366	366	363	366	366	366
R^2	0.004	0.003	0.015	0.002	0.001	0.012	0.002	0.010	0.175

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.084	-0.511	-0.003	-0.393	0.037	-0.092	0.008	-0.065	-0.112
	(0.131)	(0.38)	(0.229)	(0.156)**	(0.234)	(0.176)	(0.185)	(0.182)	(0.188)
H	0.103	-0.558	0.106	-0.145	0.581	0.062	0.424	0	0
	(0.049)**	(0.236)**	(0.266)	(0.269)	(0.158)***	(0.2)	(0.347)	(.)	(.)
H*Treated	-0.053	0.589	-0.184	0.765	-0.209	0.018	-0.237	-0.025	0.074
	(0.072)	(0.314)*	(0.204)	(0.243)***	(0.258)	(0.363)	(0.328)	(0.134)	(0.123)
N	366	366	366	366	366	366	366	366	366
R^2	0.016	0.016	0.002	0.025	0.029	0.001	0.013	0.001	0.001

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Commitment to academic goal is respondentâĂŹs level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA22: Heterogeneity in impact on Willing to Modify Goal

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.066	-0.025	-0.028	-0.016	-0.09	0.101	0.058	0.017	-0.049
	(0.040)*	(0.058)	(0.045)	(0.051)	(0.033)***	(0.077)	(0.020)***	(0.034)	(0.042)
H	-0.091	-0.047	-0.016	0.05	-0.036	0.033	0.117	-0.05	-0.182
	(0.065)	(0.072)	(0.044)	(0.033)	(0.097)	(0.103)	(0.083)	(0.049)	(0.084)**
H*Treated	0.132	0.067	0.054	0.033	0.213	-0.157	-0.177	-0.063	0.09
	(0.047)***	(0.125)	(0.095)	(0.093)	(0.070)***	(0.126)	(0.041)***	(0.075)	(0.073)
N	366	366	366	366	366	363	366	366	366
R^2	0.006	0.001	0.001	0.005	0.019	0.008	0.009	0.008	0.024

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.097	-0.012	-0.031	0.066	0.029	-0.029	0.031	-0.065	-0.044
	(0.074)	(0.058)	(0.058)	(0.058)	(0.051)	(0.043)	(0.047)	(0.058)	(0.037)
H	0.008	0.075	0.073	0.094	0.078	-0.016	0.082	0	0
	(0.015)	(0.067)	(0.043)*	(0.105)	(0.068)	(0.066)	(0.062)	(.)	(.)
H*Treated	0.027	0.017	0.051	-0.176	-0.061	0.071	-0.084	0.086	0.107
	(0.018)	(0.102)	(0.093)	(0.104)*	(0.071)	(0.084)	(0.093)	(0.072)	(0.061)*
N	366	366	366	366	366	366	366	366	366
R^2	0.011	0.007	0.012	0.009	0.004	0.002	0.004	0.004	0.006

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA23: Heterogeneity in impact on Goal Index

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.143	0.017	0.043	-0.075	-0.261	0.299	0.088	-0.028	-0.061
	(0.177)	(0.153)	(0.148)	(0.196)	(0.144)*	(0.141)**	(0.095)	(0.095)	(0.117)
H	-0.26	0.135	0.203	0.169	-0.105	0.387	0.113	-0.145	0.685
	(0.195)	(0.118)	(0.096)**	(0.091)*	(0.143)	(0.210)*	(0.114)	(0.116)	(0.153)***
H*Treated	0.193	-0.221	-0.213	0.052	0.494	-0.485	-0.391	-0.1	0.023
	(0.154)	(0.224)	(0.231)	(0.245)	(0.143)***	(0.280)*	(0.120)***	(0.112)	(0.153)
N	366	366	366	366	366	363	366	366	366
R^2	0.009	0.003	0.006	0.010	0.020	0.015	0.011	0.010	0.116

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.141	-0.293	-0.127	-0.129	0.071	-0.097	0.036	-0.144	-0.134
	(0.131)	(0.214)	(0.18)	(0.154)	(0.179)	(0.123)	(0.155)	(0.133)	(0.171)
H	0.070	-0.226	0.181	0.046	0.422	0.08	0.377	0	0
	(0.026)***	(0.060)***	(0.168)	(0.124)	(0.103)***	(0.135)	(0.157)**	(.)	(.)
H*Treated	0.022	0.33	0.126	0.171	-0.237	0.102	-0.238	0.121	0.2
	(0.030)	(0.199)*	(0.206)	(0.158)	(0.189)	(0.206)	(0.25)	(0.135)	(0.203)
N	366	366	366	366	366	366	366	366	366
R^2	0.027	0.007	0.016	0.007	0.026	0.005	0.019	0.002	0.005

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Goal Index is an Anderson Index created using the willing to modify goal, commitment to achieve goal, importance of goal and achieve academic goal measured post intervention at the time of baseline. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA24: Heterogeneity in impact on Happiness

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.531	0.920	0.753	0.603	0.295	0.766	0.545	-0.035	0.790
	(0.959)	(0.731)	(0.583)	(0.788)	(0.921)	(0.892)	(0.945)	(0.556)	(0.941)
H	-0.808	0.822	1.165	1.333	-0.191	0.937	1.606	-0.720	0.304
	(0.387)**	(0.459)*	(0.453)**	(0.657)**	(0.438)	(0.953)	(1.237)	(1.055)	(0.516)
H*Treated	0.015	-1.368	-0.630	-0.193	0.389	-0.325	-0.608	1.193	-0.621
	(0.946)	(0.374)***	(0.604)	(0.581)	(0.794)	(1.041)	(1.619)	(1.333)	(0.800)
N	452	452	452	452	452	452	452	452	452
R^2	0.013	0.010	0.016	0.027	0.004	0.010	0.028	0.009	0.005

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.193	-0.648	0.857	0.994	0.903	1.544	0.559	0.896	0.658
	(1.024)	(1.415)	(0.679)	(1.029)	(0.642)	(0.601)**	(0.667)	(1.009)	(0.738)
H	-0.146	-0.436	0.536	1.162	1.220	1.804	0.324	0.000	0.000
	(0.103)	(0.862)	(1.024)	(0.822)	(0.754)	(0.600)***	(0.452)	(.)	(.)
H*Treated	0.201	1.574	-0.891	-1.279	-0.863	-2.637	-0.244	-0.762	-4.478
	(0.115)*	(1.026)	(0.754)	(0.916)	(1.115)	(1.019)***	(0.951)	(0.828)	(1.487)***
N	452	452	452	452	452	452	452	452	452
R^2	0.007	0.013	0.006	0.013	0.015	0.031	0.004	0.008	0.030

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Happiness is measured using a 4-item scale ranging from a 1 to 7-point Likert scale measured post intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA25: Heterogeneity in impact on Self Discipline

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.577	-0.677	-0.742	1.628	0.901	3.125	-0.866	-0.732	0.447
	(0.802)	(1.175)	(1.195)	(0.760)**	(0.550)	(1.638)*	(0.493)*	(0.659)	(1.092)
H	-1.453	-0.685	0.563	3.751	-0.073	2.984	2.070	0.941	1.587
	(1.093)	(1.378)	(1.242)	(0.744)***	(1.221)	(1.184)**	(1.500)	(1.573)	(1.033)
H*Treated	1.358	2.046	1.593	-3.306	-2.106	-4.096	1.375	1.830	-0.820
	(1.687)	(2.886)	(2.198)	(1.043)***	(1.199)*	(1.495)***	(1.070)	(1.246)	(1.052)
N	452	452	452	452	452	452	452	452	452
R^2	0.006	0.007	0.016	0.042	0.015	0.021	0.050	0.027	0.009

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.163	-0.586	1.079	0.216	0.867	0.881	-0.426	0.443	0.133
	(2.095)	(1.396)	(1.129)	(0.825)	(1.174)	(1.496)	(0.904)	(0.901)	(0.584)
H	0.012	1.337	0.942	1.803	3.486	2.617	0.914	0.000	0.000
	(0.354)*	(1.534)	(0.775)	(1.252)	(1.299)***	(1.210)**	(0.851)	(.)	(.)
H*Treated	0.044	0.831	-2.428	-0.656	-1.699	-2.345	1.012	-0.809	-3.391
	(0.569)	(1.704)	(1.541)	(1.499)	(2.095)	(2.887)	(1.542)	(1.442)	(1.938)*
N	452	452	452	452	452	452	452	452	452
R^2	0.000	0.017	0.010	0.013	0.046	0.019	0.014	0.002	0.006

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Self-discipline is a 10-item scale, each item is based on responses ranging from a 1 to 5-point Likert scale measured post intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA26: Heterogeneity in impact on Self Esteem

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.231	-0.998	-0.881	-0.644	-0.163	-1.974	-1.226	-0.217	-0.543
	(0.457)	(0.212)***	(0.176)***	(0.376)*	(0.248)	(0.516)***	(0.238)***	(0.398)	(0.383)
H	1.314	-0.700	-0.313	-0.840	0.881	-1.257	-1.934	0.817	-0.389
	(0.416)***	(0.317)**	(0.217)	(0.547)	(0.324)***	(0.434)***	(0.492)***	(0.518)	(0.525)
H*Treated	-1.002	1.123	0.552	-0.029	-0.999	1.760	1.879	-0.960	-0.195
	(0.470)**	(0.712)	(0.439)	(0.430)	(0.456)**	(0.639)***	(0.356)***	(0.515)*	(0.524)
N	452	452	452	452	452	452	452	452	452
R^2	0.038	0.021	0.014	0.034	0.023	0.031	0.059	0.022	0.020

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	-0.262	-0.600	-1.190	-0.649	-0.693	-0.855	-0.797	-1.245	-0.650
	(0.682)	(0.450)	(0.263)***	(0.247)***	(0.261)***	(0.266)***	(0.281)***	(0.339)***	(0.290)**
H	-0.024	-0.342	-0.566	-0.918	-1.179	-1.788	-0.302	0.000	0.000
	(0.122)	(0.581)	(0.412)	(0.395)**	(0.535)**	(0.483)***	(0.268)	(.)	(.)
H*Treated	-0.105	-0.046	1.263	0.119	0.075	0.752	0.446	1.103	0.496
	(0.160)	(0.479)	(0.547)**	(0.316)	(0.624)	(0.466)	(0.432)	(0.248)***	(0.498)
N	452	452	452	452	452	452	452	452	452
R^2	0.016	0.015	0.024	0.034	0.051	0.072	0.014	0.031	0.013

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Self Esteem is a 10-item scale based on responses ranging from a 0 to 4-point Likert scale measured post intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA27: Heterogeneity in impact on Depression

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.362	1.228	0.803	1.081	1.555	1.734	0.773	1.416	0.707
	(0.584)	(0.542)**	(0.822)	(1.008)	(0.753)**	(0.762)**	(0.630)	(0.836)*	(1.064)
H	-0.803	2.372	0.331	0.655	0.493	-0.094	-0.370	-0.677	0.340
	(1.064)	(0.941)**	(0.988)	(1.239)	(0.724)	(0.621)	(0.485)	(0.660)	(1.213)
H*Treated	0.752	-1.656	-0.278	-0.836	-1.982	-1.484	-0.116	-1.852	-0.040
	(1.448)	(1.828)	(1.339)	(1.630)	(1.188)*	(0.767)*	(0.843)	(0.806)**	(1.927)
N	452	452	452	452	452	452	452	452	452
R^2	0.007	0.027	0.005	0.006	0.015	0.014	0.006	0.035	0.005

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.706	-0.052	-0.397	1.375	1.359	0.262	1.105	0.425	0.473
	(0.995)	(0.518)	(0.837)	(0.880)	(0.721)*	(0.871)	(0.656)*	(0.582)	(0.564)
H	0.223	-1.293	-2.039	-0.943	-0.742	-2.244	0.781	0.000	0.000
	(0.293)	(1.104)	(0.762)***	(1.288)	(0.850)	(0.696)***	(0.953)	(.)	(.)
H*Treated	-0.022	1.014	2.558	-1.503	-1.499	1.238	-1.141	0.452	4.887
	(0.264)	(0.521)*	(0.780)***	(1.490)	(1.285)	(1.002)	(0.993)	(0.532)	(1.640)***
N	452	452	452	452	452	452	452	452	452
R^2	0.011	0.010	0.022	0.036	0.030	0.029	0.007	0.005	0.022

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Depression is a 9-item scale based on responses ranging from a 0 to 3 point Likert scale measured post intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA28: Heterogeneity in impact on Mindset Index

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.079	-0.357	-0.315	-0.069	-0.049	-0.416	-0.572	-0.248	-0.083
	(0.180)	(0.136)***	(0.137)**	(0.161)	(0.141)	(0.212)**	(0.215)***	(0.161)	(0.141)
H	0.262	-0.355	0.088	0.189	0.230	-0.023	-0.222	0.307	0.026
	(0.123)**	(0.217)	(0.220)	(0.171)	(0.159)	(0.204)	(0.149)	(0.106)***	(0.176)
H*Treated	-0.276	0.486	0.257	-0.266	-0.323	0.303	0.479	0.144	-0.226
	(0.204)	(0.279)*	(0.334)	(0.176)	(0.239)	(0.244)	(0.226)**	(0.227)	(0.277)
N	452	452	452	452	452	452	452	452	452
R^2	0.015	0.020	0.022	0.012	0.014	0.015	0.016	0.039	0.012

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.146	-0.347	-0.157	-0.159	-0.129	-0.009	-0.310	-0.287	-0.149
	(0.285)	(0.216)	(0.159)	(0.150)	(0.167)	(0.240)	(0.117)***	(0.219)	(0.137)
H	0.057	0.033	0.119	0.084	0.136	0.049	-0.033	0.000	0.000
	(0.071)	(0.184)	(0.144)	(0.139)	(0.143)	(0.139)	(0.094)	(.)	(.)
H*Treated	-0.102	0.212	-0.102	-0.092	-0.138	-0.425	0.290	0.161	-1.143
	(0.066)	(0.190)	(0.116)	(0.174)	(0.255)	(0.413)	(0.138)**	(0.223)	(0.261)***
N	452	452	452	452	452	452	452	452	452
R^2	0.015	0.014	0.009	0.009	0.010	0.024	0.015	0.011	0.031

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Psychological well-being is an Anderson index constructed from happiness, self-discipline, self-esteem and depression variables measured post-intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. . * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA29: Heterogeneity in impact on Final Year Test Scores

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	44.743	28.491	49.249	13.920	29.195	75.390	26.555	12.519	29.439
	(29.495)	(14.412)**	(19.068)***	(24.799)	(27.859)	(29.284)**	(30.675)	(26.572)	(25.791)
H	32.155	57.063	79.497	0.545	5.130	86.065	49.152	-32.858	-4.021
	(23.686)	(35.337)	(28.931)***	(16.084)	(16.354)	(33.392)***	(44.244)	(21.875)	(20.938)
H*Treated	-59.156	-35.994	-70.196	6.807	-29.102	-75.242	-36.591	7.302	-26.015
	(39.626)	(50.629)	(36.008)*	(29.338)	(28.096)	(38.599)*	(32.320)	(23.311)	(33.630)
N	666	666	666	666	666	664	666	666	666
R^2	0.010	0.017	0.030	0.003	0.005	0.027	0.012	0.010	0.007

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	21.085	45.809	11.375	31.842	18.205	20.005	14.958	44.815	-0.506
	(44.821)	(40.823)	(19.291)	(29.148)	(20.526)	(23.754)	(20.824)	(54.412)	(30.086)
H	4.667	-3.931	14.756	-3.434	-12.098	-40.784	44.368	0.000	0.000
	(5.365)	(16.106)	(24.739)	(25.624)	(24.370)	(17.077)**	(17.560)**	(.)	(.)
H*Treated	-1.401	-42.344	10.339	-34.979	-3.814	-3.281	4.121	-40.269	55.429
	(10.133)	(42.051)	(45.722)	(26.207)	(17.427)	(24.788)	(30.749)	(51.765)	(49.736)
N	666	666	666	666	666	666	666	666	666
R^2	0.005	0.011	0.006	0.009	0.004	0.017	0.021	0.009	0.014

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Final year test score consists of the final year test scores after 2 months. The final year score was not available at the time baseline. H*Treated is the long-run heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA30: Heterogeneity in impact on Willingness to Learn Chinese Language

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.039	-0.037	0.024	0.152	0.015	0.142	0.072	0.008	0.130
	(0.041)	(0.042)	(0.020)	(0.067)**	(0.043)	(0.097)	(0.043)*	(0.070)	(0.047)***
H	-0.147	-0.089	-0.005	0.256	0.050	0.155	0.030	0.016	0.101
	(0.091)	(0.117)	(0.100)	(0.073)***	(0.066)	(0.073)**	(0.087)	(0.068)	(0.046)**
H*Treated	0.137	0.169	-0.013	-0.274	0.012	-0.158	-0.130	0.026	-0.218
	(0.101)	(0.117)	(0.123)	(0.112)**	(0.071)	(0.118)	(0.052)**	(0.071)	(0.074)***
N	452	452	452	452	452	452	452	452	452
R^2	0.015	0.010	0.001	0.046	0.005	0.012	0.009	0.002	0.018

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.055	-0.005	0.016	-0.002	0.072	0.065	0.061	0.019	0.026
	(0.114)	(0.106)	(0.068)	(0.058)	(0.076)	(0.060)	(0.061)	(0.051)	(0.054)
H	0.005	-0.083	0.068	-0.010	0.067	0.170	0.034	0.000	0.000
	(0.013)	(0.087)	(0.084)	(0.070)	(0.069)	(0.056)***	(0.053)	(.)	(.)
H*Treated	0.021	0.031	-0.005	0.046	-0.112	-0.128	-0.111	-0.001	-0.200
	(0.022)	(0.108)	(0.072)	(0.046)	(0.120)	(0.085)	(0.050)**	(0.066)	(0.165)
N	452	452	452	452	452	452	452	452	452
R^2	0.009	0.006	0.007	0.002	0.005	0.021	0.006	0.000	0.005

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Willingness to opt for a Chinese language is a binary variable equal to 1 if the respondent is willing to opt for a Chinese language measured post intervention at the time of follow-up. H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA31: Heterogeneity in impact on Willingness to Learn Regional Language

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.016	0.087	0.059	0.113	0.110	0.008	0.050	0.034	0.097
	(0.089)	(0.043)**	(0.072)	(0.073)	(0.080)	(0.136)	(0.062)	(0.093)	(0.068)
H	-0.103	0.063	-0.073	0.123	0.129	-0.136	-0.178	-0.004	-0.009
	(0.083)	(0.068)	(0.068)	(0.051)**	(0.094)	(0.092)	(0.067)***	(0.041)	(0.055)
H*Treated	0.088	-0.108	-0.016	-0.132	-0.118	0.049	0.052	0.043	-0.091
	(0.128)	(0.123)	(0.097)	(0.078)*	(0.115)	(0.151)	(0.111)	(0.104)	(0.069)
N	452	452	452	452	452	452	452	452	452
R^2	0.008	0.005	0.009	0.010	0.011	0.012	0.024	0.003	0.008

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	0.160	0.148	-0.032	0.021	-0.010	0.085	0.090	0.151	0.063
	(0.069)**	(0.090)*	(0.093)	(0.065)	(0.076)	(0.061)	(0.060)	(0.068)**	(0.061)
H	0.003	0.068	-0.033	0.045	-0.098	0.039	0.033	0.000	0.000
	(0.021)	(0.074)	(0.089)	(0.080)	(0.059)*	(0.081)	(0.049)	(.)	(.)
H*Treated	-0.032	-0.135	0.180	0.065	0.127	-0.079	-0.101	-0.178	-0.278
	(0.011)***	(0.109)	(0.136)	(0.069)	(0.092)	(0.078)	(0.089)	(0.074)**	(0.163)*
N	452	452	452	452	452	452	452	452	452
R^2	0.009	0.006	0.015	0.010	0.008	0.004	0.006	0.019	0.009

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Willingness to opt for a regional language is a binary variable equal to 1 if the respondent is willing to opt for a regional language measured post intervention at the time of follow-up . H*Treated is the immediate post-intervention heterogeneous treatment effect. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA32: Heterogeneity in impact on exam score percentage 2 months after implementation

H:	HH assets	Father Educ	Mother Educ	Competitiveness	Raven	First Division	First Year	HH Chores	Career Goal Importance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	44.743	28.491	49.249	13.92	29.195	75.39	90.385	12.519	29.439
	(29.495)	(14.412)**	(19.068)***	-24.799	-27.859	(29.284)**	-65.341	(26.572)	(25.791)
H	32.155	57.063	79.497	0.545	5.13	86.065	-128.417	-32.858	-4.021
	(23.686)	(35.337)	(28.931)***	-16.084	-16.354	(33.392)***	(60.374)**	(21.875)	(20.938)
H*Treated	-59.156	-35.994	-70.196	6.807	-29.102	-75.242	-83.721	7.302	-26.015
	(39.626)	(50.629)	(36.008)*	-29.338	-28.096	(38.599)*	-64.243	(23.311)	(33.63)
N	666	666	666	666	666	664	666	666	666
R^2	0.01	0.017	0.03	0.003	0.005	0.027	0.189	0.01	0.007

H:	Study Time	Discuss HW	Extraversion	Agreeable.	Conscient.	Stable.	Open.	Stroop Error	Stroop Time
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treated	4.722	45.809	11.375	31.842	18.205	20.005	14.958	44.815	-0.506
	-45.091	-40.823	-19.291	-29.148	-20.526	-23.754	-20.824	-54.412	-30.086
H	-0.444	-3.931	14.756	-3.434	-12.098	-40.784	44.368	0	0
	-5.327	-16.106	-24.739	-25.624	-24.37	(17.077)**	(17.560)**	(.)	(.)
H*Treated	3.608	-42.344	10.339	-34.979	-3.814	-3.281	4.121	-40.269	55.429
	-10.045	-42.051	-45.722	-26.207	-17.427	-24.788	-30.749	-51.765	-49.736
N	666	666	666	666	666	666	666	666	666
R^2	0.003	0.011	0.006	0.009	0.004	0.017	0.021	0.009	0.014

Note: Standard errors in parentheses. All errors are clustered at the college level. All regressions include college and enumerator fixed effects with errors clustered at the college level. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. Fixed effects include college and enumerator fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA8.4 Attrition

We test for differential attrition between the treatment and placebo group by coding a dummy variable for whether individual i attrited for the follow-up survey and then test if attrition is significantly related to baseline covariates of interest. A complete list of baseline covariates for which we test balance (between treatment and control) is specified our preanalysis plan (Haroon et al., 2020).

We test balance by running the following regression:

$$ATT_i = \beta_0 + \beta_1.x_i + \epsilon_i \tag{10}$$

$$ATT_i = \beta_0 + \beta_1.T_i + \epsilon_i \tag{11}$$

where ATT_i is a dummy variable for whether individual i attrited between baseline and follow-up survey, μ_c denote college fixed effects. All errors are clustered at the college level. The main hypothesis we propose to test is that attrition is unrelated to baseline covariates and treatment status *i.e.* growth mindset discussion has no effect; $H_0 : \beta_1 = 0$. A significant imbalance is detected when the p-value is smaller than 0.10.

Results are given in Table OA33. We find that attrition varies by respondent age, occupation of parents, ravens score, competitiveness, measures of past academic performance and whether the respondent discusses education goal with parents. Attrition does not vary by treatment status. The overall test of joint significance i.e., F-statistic is 0.11 thereby implying a well balanced sample. However, we show robustness of results in two ways (i) Table OA8.6, we show results for primary outcomes using the Lee (2009) bounds. (ii) As discussed in section OA8, we also repeat all main estimations using 'post-double-selection' with LASSO (see Belloni et al. (2014a) and Belloni et al. (2014b)), where we use all of the baseline covariates specified in Haroon et al. (2020).

Table OA33: Balance test for Attrition

	N	Mean	Median	S.Dev.	$p-value$
	(1)	(2)	(3)	(4)	(5)
Age (years)	366	19.37	19.00	1.62	0.08*
Dummy: Enrolled in first year	366	0.36	0.00	0.48	0.19
Dummy: Single	366	0.96	1.00	0.20	0.35
Father's education (years)	351	8.43	10.00	4.80	0.38
Dummy: Father is a business owner	366	0.25	0.00	0.44	0.20
Dummy: Father is a salaried worker	366	0.57	1.00	0.50	0.04**
Mother's education (years)	362	6.64	8.00	5.15	0.74
Dummy: Mother is a business owner	366	0.02	0.00	0.13	0.00***
Dummy: Mother is a salaried worker	366	0.05	0.00	0.23	0.48
Average monthly household income (000's)	258	35.56	25.00	33.09	0.27
Index: Household assets	366	0.00	0.30	1.39	0.59
Risk preferences (higher is more risk averse)	366	4.15	4.00	1.59	0.89
Ravens test score (out of 10)	366	4.08	4.00	2.11	0.03**
Scale: Competitiveness (out of 75)	366	57.62	58.00	8.33	0.07*
Dummy: Continue education after graduating	366	0.91	1.00	0.28	0.90
Dummy: Passed last degree in first division	363	0.72	1.00	0.45	0.06*
Scale:Satisfaction with academic performance baseline	366	2.37	2.00	0.91	0.15
Daily hours studying at home in average week	366	2.99	3.00	2.07	0.52
Daily hours doing household chores in average week	366	6.28	7.00	5.99	0.24
Dummy: Discusses homework with parents	366	0.71	1.00	0.46	0.21
Dummy: Discusses education goals with parents	366	0.93	1.00	0.27	0.06*
Dummy: Pursue paid job	366	0.92	1.00	0.27	0.30
Dummy: Have a role model	366	0.33	0.00	0.47	0.59
Dummy: Treatment	366	0.51	1.00	0.51	0.19
F-statistics					0.11

Note: Column (5) shows p-values from the balance test specified in equation 7. The cells show the coefficient on treatment assignment when the variable in the row is regressed on the treatment assignment. Column (6) shows the p-values from a test of baseline balance between the in-person and phone interview sample ('type' of survey), from a regression of the variable in the row on the type of interview. F-statistic explains the overall significance of the model- the null hypothesis is that all coefficients on the independent variables are equal to zero. If the null is rejected, there is a well balanced sample. ***p < 0.01, **p < 0.05,*p < 0.1.

OA8.5 Primary outcomes

Selection of controls using PDS Lasso

Table OA34: Robustness check: Immediate impact of treatment on primary outcomes, using PDS Lasso

	(1) Grit	(2) Growth mindset	(3) Self efficacy	(4) Mindset Index
Treated	-0.745	1.030	0.390	0.071
	(0.552)	(0.527)*	(0.582)	(0.122)
Mean	41.380	49.877	32.905	-0.000
N	366	366	366	366

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured post-intervention. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured post intervention. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured post intervention. Mindset Index is an Anderson index constructed from grit, growth mindset and self-efficacy measured post intervention. Treated is the immediate effect of our treatment post intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA35: Robustness check: Long term impact of treatment on primary outcomes, using PDS Lasso

	(1) Grit	(2) Growth mindset	(3) Self efficacy	(4) Mindset Index	(5) Test scores
Treated	-0.954	0.897	0.191	0.019	18.299
	(0.591)	(0.429)**	(0.597)	(0.113)	(16.629)
Post	-0.689	-2.753	0.047	-0.336	0.000
	(0.374)*	(0.802)***	(0.511)	(0.091)***	(.)
T*Post	1.753	0.291	-0.219	0.172	0.000
	(0.614)***	(1.003)	(0.560)	(0.147)	(.)
Mean	41.125	48.660	32.938	-0.143	259.633
N	666	666	666	666	333

Standard errors in parentheses

Controls are selected using PDS Lasso (Belloni et al., 2013). All errors are clustered at the college level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Restrict sample to students who have received repeat intervention

Table OA36: Immediate impact of treatment on primary outcomes - robustness check using restricted sample

	(1) Grit	(2) Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset Index	(8) Mindset Index
Treated	-0.362	-0.207	0.502	0.182	0.620	0.788	0.072	0.069
	(0.555)	(0.517)	(0.644)	(0.642)	(0.371)*	(0.425)*	(0.071)	(0.065)
Mean	41.104	41.104	50.557	50.557	32.689	32.689	0.024	0.024
N	226	226	226	226	226	226	226	226
R^2	0.001	0.000	0.001	0.000	0.004	0.006	0.001	0.001
Controls	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes

Note: Sample in this regression is restricted to the balanced sample of individuals who have received the repeat intervention. Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured post-intervention. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured post intervention. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured post intervention. Mindset Index is an Anderson index constructed from grit, growth mindset and Self-efficacy measured post intervention. Treated is the immediate effect of our treatment post intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA37: Long term impact of treatment on primary outcomes

	(1) Grit Grit	(2) Grit Grit	(3) Growth mindset	(4) Growth mindset	(5) Self efficacy	(6) Self efficacy	(7) Mindset Index	(8) Mindset Index	(9) Test scores	(10) Test scores
Treated	-0.362	-0.269	0.502	0.151	0.620	0.690	0.071	0.050	16.312	8.090
	(0.555)	(0.531)	(0.644)	(0.623)	(0.371)*	(0.360)*	(0.070)	(0.065)	(17.940)	(24.017)
Post	-0.226	-0.226	-2.755	-2.755	0.066	0.066	-0.296	-0.296		
	(0.278)	(0.278)	(0.774)***	(0.774)***	(0.583)	(0.583)	(0.092)***	(0.092)***		
T*Post	1.126	1.126	0.013	0.013	-0.533	-0.533	0.061	0.061		
	(0.568)**	(0.568)**	(0.674)	(0.674)	(0.724)	(0.724)	(0.103)	(0.103)		
Mean	40.991	40.991	49.179	49.179	32.722	32.722	0.000	0.000	241.534	241.534
N	452	452	452	452	452	452	452	452	225	225
R^2	0.003	0.004	0.042	0.043	0.002	0.003	0.020	0.020	0.003	0.001
Controls	No	No	No	No	No	No	No	No	No	No
FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Sample in this regression is restricted to the balanced sample of individuals who have received the repeat intervention. Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and Self-efficacy measured at the time of follow-up. Test scores are the scores of participants in end-of-year exams which are not measured at the time of baseline. T*Post is the average treatment effect of our growth mindset intervention. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Upper and Lower Bound estimates using Lee (2009)

Table OA38: Robustness of Impact on Primary Outcomes to Attrition

	(1) Grit	(2) Growth Mindset	(3) Self Efficacy	(4) Psych Index	(5) Test Score
Treated	-1.027	1.192	0.058	0.071	17.085
	(0.622)*	(0.611)*	(0.589)	(0.07)	(21.604)
Upper bound	-0.011	1.842	1.201	0.182	22.134
	(0.778)	(0.810)**	(0.714)*	(0.135)	(22.119)
Lower bound	-1.519	0.261	-0.249	-0.072	-23.124
	(0.811)*	(0.865)	(0.656)	(0.135)	(24.665)
N	366.000	366.000	366.000	366.000	363.000

Note: Standard errors in parentheses. All errors are clustered at the college level. Grit is measured using 12-item scale ranging from a 1 to 5-point Likert scale measured at the time of follow-up. Growth mindset is a scale constructed from the sum of 15 items on 1 to 6-point Likert scale measured at the time of follow-up. Self-efficacy is measured using 10-item scales on 1 to 4 Likert scale measured at the time of follow-up. Mindset Index is an Anderson index constructed from grit, growth mindset and Self-efficacy measured at the time of follow-up. Test scores are the scores of participants in end-of-year exams which are not measured at the time of baseline. Treated is the coefficient from our main tables. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA8.6 Secondary outcomes

Selection of controls using PDS Lasso

Table OA39: Immediate impact of treatment on secondary outcomes, using PDS Lasso

	(1) Importance of goal	(2) Will achieve goal	(3) Commitment to goal	(4) Modified goal	(5) Goal Index
Treated	-0.065	-0.070	-0.084	-0.003	-0.057
	(0.112)	(0.091)	(0.170)	(0.024)	(0.108)
Mean	8.648	7.883	8.330	0.687	-0.000
N	366	366	366	366	366

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. Commitment to academic goal is respondentâĂŹs level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. Goal Index is an Anderson Index created using the above 4 variables measured post intervention at the time of baseline. Treated is the immediate effect of our treatment post intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA40: Immediate impact of treatment on secondary outcomes, using PDS Lasso

	(1) Happiness	(2) Self discipline	(3) Self esteem	(4) Depression	(5) Mindset Index
Treated	0.471	-0.008	-0.630	0.677	-0.197
	(0.798)	(0.525)	(0.282)**	(0.582)	(0.137)
Mean	19.821	34.858	12.755	7.840	-0.000
N	226	226	226	226	226

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Happiness is measured using a 4-item scale ranging from a 1 to 7-point Likert scale measured post intervention at the time of follow-up. Self-discipline is a 10-item scale, each item is based on responses ranging from a 1 to 5-point Likert scale measured post intervention at the time of follow-up. Self Esteem is a 10-item scale based on responses ranging from a 0 to 4-point Likert scale measured post intervention at the time of follow-up. Depression is a 9-item scale based on responses ranging from a 0 to 3 (point Likert scale measured post intervention at the time of follow-up. Psychological well-being is an Anderson index using all variables above. Treated is the immediate effect of our treatment post intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA41: Immediate impact of treatment on secondary outcomes, using PDS Lasso

	(1) Willingness to Learn Chinese Language	(2) Willingness to Learn Regional language
Treated	0.018	0.052
	(0.050)	(0.059)
Mean	0.774	0.415
N	226	226

Note: Standard errors in parentheses. All errors are clustered at the college level. Fixed effects include college and enumerator fixed effects. Willingness to opt for a Chinese language is a binary variable equal to 1 if the respondent is willing to opt for a Chinese language measured post intervention at the time of follow-up. Willingness to opt for a regional language is a binary variable equal to 1 if the respondent is willing to opt for a regional language measured post intervention at the time of follow-up. Treated is the immediate effect of our treatment post intervention. Controls are selected using PDS Lasso (Belloni et al., 2013). N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Upper and Lower Bound estimates using Lee (2009)

Table OA42: Robustness of Impact on Secondary Outcomes to Attrition

	(1) Important Academic Goal	(2) Achieve Academic Goal	(3) Commitment to Academic Goal	(4) Willing to Modify Goal	(5) Goal Index
Treated	-0.065	-0.07	-0.084	-0.003	-0.057
	(0.112)	(0.091)	(0.170)	(0.024)	(0.108)
Upper bound	0.042	0.273	0.078	0.045	0.118
	(0.178)	(0.188)	(0.204)	(0.062)	(0.149)
Lower bound	-0.247	-0.102	-0.230	-0.020	-0.133
	(0.189)	(0.224)	(0.213)	(0.054)	(0.127)
N	366.000	366.000	366.000	366.000	366.000

Note: Standard errors in parentheses. All errors are clustered at the college level. Important academic goal is the importance academic goal for a respondent on a Likert scale from 1 to 10 measured post intervention at the time of baseline. Achieve academic goal is a response how likely a respondent is likely to achieve her goal on a scale from 1 to 10 measured post intervention at the time of baseline. Commitment to academic goal is respondentâĂŹs level of commitment to achieve goal ranked on a Likert scale from 0 to 10 measured post intervention at the time of baseline. Willing to modify goal is a binary variable if a respondent is willing to change her goal to SMART goals measured post intervention at the time of baseline. Goal Index is an Anderson Index created using the above 4 variables measured post intervention at the time of baseline. Treated is the coefficient from our main tables. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA43: Robustness of Impact on Secondary Outcomes to Attrition

	(1) Happiness	(2) Self Discipline	(3) Self Esteem	(4) Depression	(5) Mindset Index
Treated	0.471	-0.008	-0.630	0.677	-0.197
	(0.798)	(0.525)	(0.282)âĹŮâĹŮ	(0.582)	(0.137)
Upper bound	1.119	1.038	-0.146	1.358	-0.008
	(0.604)*	(1.006)	(0.469)	(0.860)	(0.176)
Lower bound	-0.150	-1.040	-1.037	-0.256	-0.388
	(0.643)	(1.156)	(0.471)**	(0.736)	(0.190)**
N	292.000	292.000	292.000	292.000	292.000

Note: Standard errors in parentheses. All errors are clustered at the college level. Happiness is measured using a 4-item scale ranging from a 1 to 7-point Likert scale measured post intervention at the time of follow-up. Self-discipline is a 10-item scale, each item is based on responses ranging from a 1 to 5-point Likert scale measured post intervention at the time of follow-up. Self Esteem is a 10-item scale based on responses ranging from a 0 to 4-point Likert scale measured post intervention at the time of follow-up. Depression is a 9-item scale based on responses ranging from a 0 to 3 (point Likert scale measured post intervention at the time of follow-up. Psychological well-being is an Anderson index using variables the above variables. Treated is the coefficient from our main tables. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table OA44: Robustness of Impact on Secondary Outcomes to Attrition

	(1) Willingness to learn Chinese Language	(2) Willingness to learn Regional Language
Treated	0.018	0.052
	(0.05)	(0.059)
Upper bound	0.087	0.092
	(0.084)	(0.079)
Lower bound	0.000	0.005
	(0.065)	(0.078)
N	292.000	292.000

Note: Standard errors in parentheses. All errors are clustered at the college level. Willingness to opt for a Chinese language is a binary variable equal to 1 if the respondent is willing to opt for a Chinese language measured post intervention at the time of follow-up. Willingness to opt for a regional language is a binary variable equal to 1 if the respondent is willing to opt for a regional language measured post intervention at the time of follow-up. Treated is the coefficient from our main tables. N refers to the sample size. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OA9 Supplementary Material: Goal Setting Activity

Begin survey > Goal setting > Organizing and Evaluating Your Goal			
Begin survey > Goal setting > Organizing and Evaluating Your Goal > Evaluate Goals			
evaluate_g	On a scale from 1 to 10 where 1 means "not important at all" and 10 means "extremely important", how important overall, each one of the goals is to you? Rank each of them according to how you really feel. There are no right or wrong answers.		
academicgo	How important is your academic goal to you?	1	1. Not important at all
		2	2
		3	3
		4	4
		5	5
		6	6
		7	7
		8	8
		9	9
		10	10. Extremely important goal
careergoal_	How important is your career goal to you?	1	1. Not important at all
		2	2
		3	3
		4	4
		5	5
		6	6
		7	7
		8	8
		9	9
		10	10. Extremely important goal
Begin survey > Goal setting > Organizing and Evaluating Your Goal > Achieve Goals			
achieve_go	On a scale from 1 to 10 where 1 means "not at all likely" and 10 means "extremely likely", how likely is it that you will ever achieve your goals?		
academic_g	How likely is it that you will ever achieve your academic goal?	1	1. Not at all likely
		2	2
		3	3
		4	4
		5	5
		6	6
		7	7
		8	8
		9	9
		10	10. Extremely likely

career_goa	How likely is it that you will ever achieve your career goal?	1	1. Not at all likely
		2	2
		3	3
		4	4
		5	5
		6	6
		7	7
		8	8
		9	9
		10	10. Extremely likely
achieve_acad	We will now focus on your Academic Goal. You said that your academic goal is [title_academic_goal_short]. Do you think you can achieve this goal in 6 months?	1	Yes
		0	No
academic_go	Can you set a goal related to your academic goal that you think you could achieve within 6 months?		
academic_go	Academic goal title new		
academic_go	Academic goal description new		
achieve_marl	If you were to achieve this goal this year (within 6 months), what percentage of marks do you think you will score in your final examinations?		
Begin survey > Goal setting > Organizing and Evaluating Your Goal > Committed goal			
commit_no	On a scale from 0 to 10 where 0 is "not committed at all" and 10 is "fully committed", how committed do you think you are to achieving your short term academic and career goal? How likely is it that you will ever achieve this goal within 6 months?		
academic_g	Academic Goal	0	0. Not committed at all
		1	1
		2	2
		3	3
		4	4
		5	5. Neither committed nor uncommitted
		6	6
		7	7
		8	8
		9	9
		10	10. Fully committed

career_goa	How likely is it that you will ever achieve your career goal?	1	1. Not at all likely
		2	2
		3	3
		4	4
		5	5
		6	6
		7	7
		8	8
		9	9
		10	10. Extremely likely
achieve_acad	We will now focus on your Academic Goal. You said that your academic goal is [title_academic_goal_short]. Do you think you can achieve this goal in 6 months?	1	Yes
		0	No
academic_go	Can you set a goal related to your academic goal that you think you could achieve within 6 months?		
academic_go	Academic goal title new		
academic_go	Academic goal description new		
achieve_marl	If you were to achieve this goal this year (within 6 months), what percentage of marks do you think you will score in your final examinations?		
Begin survey > Goal setting > Organizing and Evaluating Your Goal > Committed goal			
commit_no	On a scale from 0 to 10 where 0 is "not committed at all" and 10 is "fully committed", how committed do you think you are to achieving your short term academic and career goal? How likely is it that you will ever achieve this goal within 6 months?		
academic_g	Academic Goal	0	0. Not committed at all
		1	1
		2	2
		3	3
		4	4
		5	5. Neither committed nor uncommitted
		6	6
		7	7
		8	8
		9	9
		10	10. Fully committed

smart_goal	We will familiarize you with importance of goal setting and how effective it can be enhancing students' performance and life in general. A goal directs a student's attention toward the goal, increase effort and persistence and motivate a student to find strategies to attain a goal. Effective goal setting helps people achieve more and feel better while doing it. It focuses your activities, increases motivation towards achieving specific aims and makes you more likely to work through setbacks. Research shows that student' s exposure to goal setting exercises has increased their performance by 30%. We will introduce you with a new strategy of developing goals which can enable you achieve to your targets and goals easily and you should be able to keep track of them. There are 5 important criteria that your goal should satisfy:	
smart_goal	1. Specific: Goals should be precise and clear description of what you want to achieve or what is your objective. A specific goal should answer: a. What do I want to accomplish? b. Why: Specific reasons, purpose or benefits of accomplishing the goal. c. Who is involved? d. Where: Identify a location. e. Which: Identify requirements and constraints.	
smart_goal	2. Measurable: it should contain quantities, numbers, etc. so that you can come back to it in the future and be able to say whether you reached your goals or not. You should be easily able to identify how much you are able to achieve. A measurable goal should answer: a. How much? b. How many? c. How will I know when it is accomplished?	
smart_goal	3. Ambitious: it is important to challenge yourself and aim at something that you have not done, that would take effort, but that could be very rewarding if you achieve it. a. Does this seem worthwhile? b. Is this the right time? c. Does this match our other efforts/needs? d. Are you the right person? e. Is it applicable in current socio- economic- technical environment?	
smart_goal	4. Realistic: not so ambitious that you will never be able to achieve it, something that you think you can achieve if you put effort and God helps you. a. How: How can the goal be accomplished?	
smart_goal	5. Time-bound: Goals should be linked to a timeframe that creates a practical sense of urgency, or results in tension between the current reality and the vision of the goal. While setting the goal one should specify the exact date when you think your goal should be reached. A time-bound goal should answer: a. When? b. What can I do six months from now? c. What can I do six weeks from now? d. What can I do today?	
Begin survey > Goal setting > Setting SMART Goals > SMART goals example		

ex_note	Let's discuss next using few examples.	
ex1	Example 1: Sadia is studying in first year of undergraduate program. She has final exams six months from now on and is deciding what goals to set for her academics. She is currently scoring 50% of marks in her exams. Imagine she has the following long- term and short-term goals:	
ex1_longte	Long-term goal: I want to improve my marks in the final exams by 10% at the end of 6 months. In order to achieve this, I will spend two hours studying one of the subjects each day. Is this goal smart? Specific: Yes, because this describes what she wants to achieve - improve marks in all subjects. Measurable: Yes, improve marks by 10% and spend two hours. Ambitious: Yes, because she wants to increase marks in all subjects by 10%. She can even do more? increase by more than 10%. Realistic: Yes, because she already has 50% marks and improving them by 10% would be a realistic goal. Time-bound: Yes, because she specified that it would be in 6 months	
ex1_shortg	Short-term goal: In order to achieve her long-term goals of improving marks by 10%. Sadia sets short term goals as well. Her goals are: "In my up-coming class test which are due in 2 weeks, I will improve my performance by one grade (or 5 %) in all subjects. For the next two weeks, I will take careful notes and review them at least two days before test/quiz so that I can ask the teacher questions about what I don't understand. I will complete my homework and before handing in my assignment I will ask the teacher about anything I'm not sure about. When I get anything wrong, I will make sure to ask the teacher, or one of my classmates how they got the right answer. Every day I will spend 2 hours on one of the subjects and cover all subjects in a week's time." Is this goal smart? Specific: Yes, because this describes what she wants to achieve. Measurable: Yes, improve marks by 5% and spend two hours every day Ambitious: Yes, she wants to increase marks in all subjects by 5%. She can even do more? increase by more than 10%. Realistic: Yes, she can improve her marks by 5% and she has developed a strategy to achieve it Time-bound: Yes, she wants to achieve them in next two weeks.	
ex2	Example 2: Ayesha is an undergraduate student who plans to start working after her bachelors. There is a job interview which is coming up next month. For this she thinks she needs to develop her interview and communication skills.	

ex2_1	Goal: I want to learn job interviews skills and work on my communication skills so that when the time of the interview comes, I can give interviews confidently and can speak well in English. For that on weekly basis, I will search and read at least two new tips regarding job interviews both from Internet, books and will also be consulting career-counseling center of my college. I will practice my speaking skills by asking questions and volunteering to answer questions in class. I will spend half an hour to practice some of the important questions asked in the interview with my friend. Is this goal smart?		
ex2_2	Specific: Yes, she wants to learn job interviews skills and work on my communication skills to become confident in giving interviews Measurable: Yes, since she plans to search and read at least two new tips regarding job interviews both from Internet, books and will also be consulting career-counseling center of my college. She will search the type of questions asked in an interview and practice speaking skills by asking questions and volunteering to answer questions in class. She also plans to spend half an hour to practice some of the important questions asked in the interview with friend. Ambitious: Yes, since she wants to develop skills Realistic: Yes, she aims to achieve the goal by practicing speaking skills in class Time bound: Yes, since she plans to improve her job interview skills within 1 months period.		
Begin survey > Goal setting > Setting SMART Goals > Are Goals SMART			
goal_smart	Now let's think about your goal. Is your goal formulated in a SMART way? *Enumerator: Remind the respondent of the 6-month academic goal they set*		
goal_S *(req*	Specific	1	Yes
		0	No
goal_M *(rec*	Measurable	1	Yes
		0	No
goal_A *(req*	Ambitious	1	Yes
		0	No
goal_R *(req*	Realistic	1	Yes
		0	No
goal_T *(req*	Time-bound	1	Yes
		0	No
Begin survey > Goal setting > Setting SMART Goals > Making goals SMART			
specific *(rec*	Your goal was not specific. Re-write your goal making it specific here. *Enumerator: Read all aspects for which respondent said "no" and record response* *Question relevant when: selected(${goal_S} , '0')*		
measurable	Your goal was not measurable. Re-write your goal making it measurable here. *Enumerator: Read all aspects for which respondent said "no" and record response*		

	Question relevant when: selected(${goal_M} , '0')	
ambitious (	Your goal was not ambitious. Re-write your goal making it ambitious here. *Enumerator: Read all aspects for which respondent said "no" and record response* *Question relevant when: selected(${goal_A} , '0')*	
realistic (re	Your goal was not realistic. Re-write your goal making it realistic here. *Enumerator: Read all aspects for which respondent said "no" and record response* *Question relevant when: selected(${goal_R} , '0')*	
time_boun	Your goal was not time bound. Re-write your goal making it time bound here. *Enumerator: Read all aspects for which respondent said "no" and record response* *Question relevant when: selected(${goal_T} , '0')*	
change_goa	Was the respondent willing and able to change her goals after her discussion? *Enumerator: Do not ask this of the respondent. Record your own observation here*	1 Yes 0 No
plan_strategi	In order to achieve this academic goal you need to have a plan. For this plan to work it has to be as specific and detailed as possible. That is, you need to have a strategy for what to do if something unexpected happens that challenges your plan. Many things related to the environment, the social world and yourself might stand in your way. It is useful to anticipate these difficulties, so that you can plan to overcome them. Such strategy could be, for example, to agree with your family members that you will sit down together and discuss what to do if something unexpected happens that challenges your plan. Another strategy could be to find people who can help you if you are not able to understand few concepts or spend time with friends in group study before exams.	
obstacles	I will now tell you about obstacles that are commonly mentioned by individuals who are in the same situation as you are, that is, who are currently studying in bachelors. For each of these obstacles, tell me if it is the one that you also face. Please tell us if you face the same obstacles or not.	
Begin survey > Goal setting > Setting SMART Goals > Obstacles		
obstacle1 (I	I have difficulty in organizing large amounts of information.	1 Yes 0 No
obstacle2 (I	I have difficulty in managing my time.	1 Yes 0 No
obstacle3 (I	I feel I lack the motivation to study certain subjects	1 Yes 0 No

obstacle4 (r	I forget things that I read or hear in class.	1	Yes
		0	No
obstacle5 (r	I suffer from test anxiety	1	Yes
		0	No
Begin survey > Goal setting > Setting SMART Goals > obstacles_list2			
obstacle7 (r	I feel that other students will always be better students than me.	1	Yes
		0	No
obstacle9 (r	Please specify any other obstacle that you may face.		
Begin survey > Goal setting > Setting SMART Goals > Overcoming obstacles and monitoring your progress			
obs_note	We need to know, concretely, whether or not we are progressing towards valued goals. Of course, this is not an easy process. When we want to complete very specific tasks, feedback on our performance is relatively easy to monitor. However, if our goals are less short-term, or more abstract, this becomes a little more abstract. You will be asked to identify how will you overcome potential obstacles and personal benchmarks that will allow you to evaluate your own performance. *Enumerator: Ask for each obstacle that the person has recognized as relevant to her. "What works as a good strategy can be very different from person to person, so what can you do to avoid giving up on your goal if you encounter this obstacle?"*		
overcomeo	What can you do? Obstacle 1 *Obstacle1: I have difficulty in organizing large amounts of information.* *Question relevant when: selected(${obstacle1} , '1')*		
succesobst	How would you measure success? Obstacle 1 *Obstacle1: I have difficulty in organizing large amounts of information.* *Question relevant when: selected(${obstacle1} , '1')*		
timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 1 *Obstacle1: I have difficulty in organizing large amounts of information.* *Question relevant when: selected(${obstacle1} , '1')*		
overcomeo	What can you do? Obstacle 2 *Obstacle2: I have difficulty in managing my time.* *Question relevant when: selected(${obstacle2} , '1')*		
succesobst	How would you measure success? Obstacle 2 *Obstacle2: I have difficulty in managing my time.* *Question relevant when: selected(${obstacle2} , '1')*		
timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 2 *Obstacle2:I have difficulty in managing my time.* *Question relevant when: selected(${obstacle2} , '1')*		
overcomeo	What can you do? Obstacle 3 *Obstacle3: I feel I lack the motivation to study certain subjects* *Question relevant when: selected(${obstacle3} , '1')*		

[redacted]	succesobst	How would you measure success? Obstacle 3 *Obstacle3: I feel I lack the motivation to study certain subjects* *Question relevant when: selected(${obstacle3} , '1')*	
[redacted]	timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 3 *Obstacle3: I feel I lack the motivation to study certain subjects* *Question relevant when: selected(${obstacle3} , '1')*	
[redacted]	overcomeo	What can you do? Obstacle 4 *Obstacle4: I forget things that I read or hear in class.* *Question relevant when: selected(${obstacle4} , '1')*	
[redacted]	succesobst	How would you measure success? Obstacle 4 *Obstacle4:I forget things that I read or hear in class.* *Question relevant when: selected(${obstacle4} , '1')*	
[redacted]	timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 4 *Obstacle4: I forget things that I read or hear in class.* *Question relevant when: selected(${obstacle4} , '1')*	
[redacted]	overcomeo	What can you do? Obstacle 5 *Obstacle5: I suffer from test anxiety* *Question relevant when: selected(${obstacle5} , '1')*	
[redacted]	succesobst	How would you measure success? Obstacle 5 *Obstacle5: I suffer from test anxiety* *Question relevant when: selected(${obstacle5} , '1')*	
[redacted]	timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 5 *Obstacle5:I suffer from test anxiety* *Question relevant when: selected(${obstacle5} , '1')*	
[redacted]	overcomeo	What can you do? Obstacle 7 *Obstacle7: I feel that other students will always be better students than me.* *Question relevant when: selected(${obstacle7} , '1')*	
[redacted]	succesobst	How would you measure success? Obstacle 7 *Obstacle7: I feel that other students will always be better students than me.* *Question relevant when: selected(${obstacle7} , '1')*	
[redacted]	timesucces	When do you think you can achieve this indicator of success? Let's mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 7 *Obstacle7: I feel that other students will always be better students than me.* *Question relevant when: selected(${obstacle7} , '1')*	
[redacted]	overcomeo	What can you do? Obstacle 9 *Obstacle9:[obstacle9]*	
[redacted]	succesobst	How would you measure success? Obstacle 9 You said that your other obstacle is [obstacle9]. Do you think you can achieve this goal in this year? *Obstacle9:[obstacle9]*	

timesucces	When do you think you can achieve this indicator of success? Lets mark the expected date on this calendar [mark on calendar] to help keep track of you plan. Obstacle 9 *You said that your other obstacle is [obstacle9]. Do you think you can achieve this goal in this year?*	
consolidate	Consolidation of plan on calendar. "Now please think of how long it will take you to achieve these milestones, and use this calendar to mark dates when you will review your plan and monitor your achievement of each milestone. Some obstacles may not be overcome once and for all, but may require regular and constant monitoring. Lets mark on this calendar, your plan for dealing with stated obstacles and achieving your goal. *Enumerator: on a calendar, mark together with the respondent the dates she identifies as suitable for achieving and monitoring milestones. Mark on calendar the days identified for monitoring the achievement of milestones.*	
reflection *(*	How would it make you feel to achieve this goal?	
reinforcem	Great work! You now have determined plans and strategies that will help you achieve your goal.	
reinforcem	Remind me again, what are you going to do today or over the next weeks in order to achieve your goal of [academic_goal_des]?	
reinforcem	Enumerator: Pause for response and provide support if needed. Record if the respondent was able to report her strategy correctly or not. "Keep the calendar as a reminder of your plan. We hope you will us it to help you achieve your goal. For example, you can put it on the wall of your room as a reminder of your plan. You can also show to your other family members to engage them in working towards your goal. Keep it and not dispose it. If you achieve any of the milestone please put a star on it.	

Lahore School of Economics
Centre for Research in Economics and Business

Recent Working Papers

No. 05-20
Impact of Parental Occupation in Career Aspirations
Fatima Shah and Zonia Saif Tirmazee

No. 04-20
Home-bias among Female Entrepreneurs: Experimental Evidence on Preferences from Pakistan
Farah Said, Mahreen Mahmud, Giovanna d' Adda and Azam Chaudhry

No. 03-20
Gender Differences in Willingness to Compete
Muhammad Ahmed Nazif and Farah Said

No. 02-20
Can Competition Reduce Moral Hazard? A Laboratory Experiment
Raja Abdar Rahman and Farah Said

No. 01-20
Macroeconomic and Spatial Determinants of Remittances: A Cross Country Analysis
Gul Noor Ismail

No. 03-19
The Impact of Social Facilities on Drop outs in Pakistans
Mahniya Zafar and Farah Said

No. 02-19
Corruption, Tax Evasion and Economic Development in Economies With Hierarchial Tax Administrative System
Anum Ellahi

No. 01-19
Political Connections, Allocation of Development Projects, and Voting Pattern in the Faisalabad District
Tahreem Zia and Azam Chaudhry

No. 03-18
The Impact of Early Childbearing and Child Health in Punjab
Aeman Nadeem and Uzma Afzal

No. 02-18
Spousal and Nonspousal Trust and Reciprocity: Evidence from a Field Experiment
Marjan Nasir and Shemiah Parshad

No. 01-18
The Impact of Prenatal Exposure to Fasting on Child Health Outcomes
Azka Sarosh Mir and Theresa Thompson Chaudhry

No. 02-17
Homophily and Strategic Behavior in Social Interactions: Evidence from a Lab Experiment
Zeenat Ansari and Farah Said

No. 01-17
The Theory of the Natural Resource Curse: A Political Economy View
Sadia Hussain and Waqar Wadho

No. 03-16
Connections and Elections in Lahore: How Network Centrality Affects Electoral Politics in Pakistan
Mahnoor Asif and Azam Chaudhry

Policy Papers

No. 01-11
Pakistan and Lessons from East Asia: Growth, Equity, and Governance
Khalid Ikram

No. 01-10
A Strategy for Reversing Pakistan's Dismal Export Performance
Hamna Ahmed, Mahreen Mahmud, Naved Hamid and Talal-Ur-Rahim

These papers can be accessed at: www.creb.org.pk

The Lahore School of Economics (established in 1993) is one of Pakistan's leading centres of learning for teaching and research in economics, finance and business administration. Its objectives are (i) to train young Pakistanis as professional economists, finance managers, accountants, financial analysts, bankers, and business executives, and (ii) to undertake research in economics, management, finance, and banking to deepen the understanding of major facts, issues, and policies.

The Centre for Research in Economics and Business (CREB) is an independent research centre at the Lahore School of Economics. CREB's mission is to conduct and facilitate research, coordinate and manage the Lahore School's postgraduate program, and promote discussion on policy issues facing Pakistan. The research focus at CREB is on the management of the Pakistan economy, income distribution and poverty, and the role of the modern services sector in the area of economics; and financial markets in the area of business management.

The Lahore School's publication program comprises the Lahore Journal of Economics, Lahore Journal of Policy Studies, Lahore Journal of Business, a Text Book Series, Lahore School Case Study Journal, the CREB Working Paper Series, and CREB Policy Paper Series. The program encourages both in-house and external contributors.

Lahore School of Economics
Intersection Main Boulevard Phase VI, DHA and Burki Road
Burki Lahore 53200, Pakistan