Too Little, Too Late: Improving Post-primary Learning Outcomes in India*

Gaurav Chiplunkar
University of Virginia

Diva Dhar
University of Oxford

Radhika Nagesh

January 2021

Abstract

We evaluate an innovative blended learning after-school intervention for post-primary students, which aims to bridge this gap by improving students' learning outcomes and non-cognitive skills in government secondary schools in India. Using a randomized controlled trial, we find that over one academic year, the program translates into large improvements in basic Math and Reading scores for the beneficiary Grade 9 students, who were found to have baseline competencies at Grade 2 level, on average. However, these improvements fail to translate into better scores on a standardized national exam in grade 10. We conclude that while such a remedial program can improve foundational skills, it is not sufficient to overcome the extent of accumulated academic deficiencies at the post-primary level, which underlines both the need to rethink the strategy for post-primary education initiatives to overcome deficiencies, and the importance of improving learning outcomes at an earlier stage.

*We thank Avanti Fellows, especially Akshay Saxena and Deepak Kamble for partnering with us on this project. We are grateful for funding from MacArthur Foundation, especially to Dipa Nag Chowdhury. We thank Clare Leaver, Anandi Mani, Alejandro Ganimian and Karthik Muralidaran and participants in the RISE Conference in 2020 for their comments. Contact information: Chiplunkar: ChiplunkarG@darden.virginia.edu; Dhar (corresponding author): diva.dhar@bsg.ox.ac.uk (corresponding author); Nagesh: radhika.n89@gmail.com

1 Introduction

Investment in education to improve human capital development and boost "inclusive growth" has been the cornerstone of development policy across many developing countries. However, despite an increase in student enrolments worldwide, learning outcomes have stagnated (Angrist et al., 2019). 61% of adolescents globally are unable to achieve minimum proficiency levels in Math and Reading, when they should be completing lower secondary school, leading to a 'learning crisis' (Bank, 2018; UNESCO, 2017).

The Indian context is no exception. India has achieved near-universal school enrollment rates in primary education and 85% enrollment in secondary education (Pratham, 2016), with significant improvements in areas of school access, enrollment, infrastructure, pupil-teacher ratios and teacher salaries over the last decade (Muralidharan, 2013). However, learning at the post-primary level continues to remain poor. For example, more than half the students in grade 8 (55% to be exact) could not read a simple English sentence and 57% could not do a simple mathematical operation (Pratham, 2016). This suggests that many students fall behind academically, even as they progress through the school system, absorbing and learning very little from their grade level classroom instruction, which would remain above their learning levels (Banerjee and Duflo, 2011; Muralidharan et al., 2019). These learning deficiencies become particularly salient in the presence of national exams (like grade 10 in India), which play an important role in determining further education and labor market outcomes.

There is an extensive literature which provides multiple explanations for the continuing deficiencies in learning levels despite high enrollment (see Kremer et al. (2013) and Glewwe and Muralidharan (2016) for extensive reviews). Kremer et al. (2013) conclude that additional schooling inputs (such as extra-teachers, more textbooks or toilets) may not be as effective in changing learning outcomes relative to improvements in pedagogical and remedial instruction, as well as accountability reforms. Similarly, a more recent review by Evans and Popova (2016) concludes that pedagogical interventions that adapt the teaching to student learning levels are more effective at increasing test scores. To that effect, a growing body of research has evaluated the impact of pedagogical interventions to improve learning outcomes and test scores in India. These include Teaching at the Right Level (Banerjee et al., 2007, 2016), adaptive technology-aided instruction

(Muralidharan et al., 2019), model schools (Kumar, 2020), free bicycles for girls (Muralidharan and Prakash, 2017) and performance-based incentives for teachers (Muralidharan and Sundararaman, 2011). This complements emerging evidence on "what works" to improve learning outcomes in developing countries, including provision of information on school quality (Andrabi et al., 2015); student level scholarships and incentives (Blimpo, 2014; Kremer et al., 2009); curricular simplification and pacing (Mbiti and Rodrigue-Segura, 2020; Pritchett and Beatty, 2012); public rankings of schools (Cilliers et al., 2019); and bundled interventions such as combining school grants with teacher incentives (Mbiti et al., 2019).

While most of the above literature has focused on improving learning outcomes in primary school, there is little evidence on how interventions impact learning outcomes at the post-primary or secondary level, in India or elsewhere. This can be particularly important from a policy perspective in targeting education interventions across different grades, given the high levels of public investment and expenditure for secondary and higher education. This paper reports on the experimental evaluation of one such technology-aided, after-school instruction program at the secondary school level (Grade 9) in India. Avanti's after-school program was developed to incorporate technology into a remedial instruction program. Avanti trains facilitators to make effective use of technology aids (video recordings, presentations, worksheets etc.) along with peer learning and counseling to enhance students' cognitive and non-cognitive development. The incorporation of technology aids in an otherwise remedial program aims to improve the quality of education for marginalized students at minimal costs to the schools.

In this paper, we conduct an experiment to evaluate the impact of Avanti's after-school blended learning program run by trained facilitators at the secondary level (Grade 9) across 24 public schools in Chennai, India. We examine whether the program improves basic learning outcomes (as measured by ASER test scores), as well as non-cognitive outcomes (such as teamwork, decision making, communication etc), gender attitudes of students and parental investment in education. We find that the program leads to large gains in basic reading and math ASER scores (40.7% and 15.5% respectively). These effects are comparable to the Teaching at the Right Level evaluation (Banerjee et al., 2016) and the MindSpark program (Muralidharan et al., 2019). However, we find that these large gains in foundational learning outcomes do not translate into improvements in exams scores in a mandatory standardized national exam in Grade 10. Lastly, we also do not find any effect of the program on improving non-cognitive abilities and parental support.

Put together, we conclude that while carefully designed remedial education interventions can bring large gains in basic learning levels even at the post-primary level (Grade 9, in our case), it is too late to translate these gains into any improvements in more advanced grade-level exams which play a crucial role in determining future outcomes for students. Despite the improvements, 60% of treatment students are not able to do division at Grade 10, which requires advanced skills such as calculus. This highlights the importance of policy debates in developing countries to address the mismatch between student learning and grade-level requirements (Muralidharan et al., 2019). Furthermore, it reiterates the importance of the timing of implementing these programs at the primary school level itself. At the post-primary school level, even successful, intensive programs such as Avanti's might come "too late" to translate large gains in foundational skills into meaningful progress on school completion and advanced human capital accumulation.

The paper is organized as follows – Section 2 describes the empirical context and the experimental design. Section 3 describes the data collection and provides details on the final sample. Section 4 discusses the results and Section 5 concludes.

2 Empirical Context

2.1 Avanti's program

Avanti's approach combines trained facilitators, technology aids and parent counseling, along with an emphasis on peer learning and team work.[1] Avanti's pedagogy incorporates features from Eric Mazur's peer instruction and collaborative learning pedagogy (Fagen et al., 2009; Schell et al., 2013; Zhang et al., 2017) and other blended learning programs. Avanti has adapted this methodology for low-resource settings (common to educational settings in developing countries and in particular, India) for teaching Maths and Science in government secondary schools in India, using basic technology and infrastructure available or installed in schools, such as computers (without internet).

Conducted after-school for an hour for five days a week, Avanti classes run for 40 weeks in

[1] Peer learning for example, engages students during class through worksheets and multiple choice questions (ConcepTests), where they are organized into learning groups based on their academic achievement levels and collaboratively encouraged to solve the worksheets. Facilitators move from group to group, encouraging discussion and resolving queries when necessary.

an academic year, led by trained facilitators rather than teachers. Facilitators are usually fresh graduates who are required to have a Bachelor's degree, preferably (but not necessarily) in STEM, and do not having teaching certification or qualifications.They are trained on facilitating classroom sessions which rely on peer learning, team work and student collaboration. The facilitators rely on pre-recorded video lectures, presentations, worksheets and other academic materials developed by Avanti for Math and Science in the regional language and context. In the study schools, the program materials and discussions are both in Tamil (the local native language), although the schools' official medium of instruction may be English or Tamil. Finally, the program also helped provide occasional career counseling and guidance for students and parents. Table B1 provides a summary of the principal components of the Avanti program.

2.2 Experimental design

The study was conducted in a sample of 24 government schools under the management of the Chennai Municipal Corporation (CMC) in the city of Chennai, India. Of the total 281 schools under the CMC, data on grade 9 students for the academic years 2014-15 and 2015-16 was provided by the Education Department of the Corporation of Chennai. Schools were selected based on the following eligibility criteria: (a) minimum of 45 students in grade 9;[2] (b) maximum of 110 students;[3] (c) no previous experience or engagement with Avanti (to avoid contamination). The experimental design involved introducing the Avanti program in 12 randomly selected schools (henceforth, treatment schools) for the academic year 2016-17, leaving students from the remaining 12 schools to serve as a comparison group (control schools). The selection was done in an open paper lottery at the Chennai Municipal Corporation, under the observation of the research team to ensure that the selection was fair. Schools complied with the treatment selection and the program was introduced for all grade 9 students in the treatment schools. From each sample school, 45 students from each school were randomly sampled based on student registers at each school for a survey at the beginning of the academic year (baseline). However, out of 1080 students listed from student registers, 65 students could not be surveyed either due to school absence on the days the survey was taking place in their schools, or because they had changed or dropped out of that school between enrolment and the

[2]Based on power calculations at 80% power and statistical significance at 5% level, assuming a 10% sample attrition rate between the baseline and endline.

[3]For logistical reasons, Avanti could not handle schools with more than 110 students.

baseline survey. Our study sample therefore comprises of 1,015 students at baseline (94% of the listed students), who were followed up with at endline. During the year, some students dropped out or transferred to other schools resulting in a final (endline) sample of 880 students one year later (87% of the baseline sample).

3 Data

We implemented three sets of student surveys at baseline and endline. The first survey included questions on students' demographic and household information, academic performance, aspirations and expenditure on education (books, tuition etc.). Second, we conducted a battery of tests to measure gender attitudes and non-cognitive skills such as critical thinking, communication, goal-setting, problem solving, grit, self-esteem and team work (West et al., 2014; Duckworth and Quinn, 2009). The survey tools were administered in the local language (Tamil). Appendix C describes the survey questions as well as provides the construction of the composite scores used in the analysis. Third, we measured learning outcomes at the start and end of the program, using the standard ASER testing tool for Reading and Mathematics developed and implemented by ASER. The ASER instrument has been widely adopted in India (and other countries) to assess students' mastery over foundational skills at different stages of education. It is also frequently used in evaluations, such as those of Pratham's Read India program (Banerjee et al., 2010). The ASER test uses a five-point grading system and covers up to Grade 2 level reading skills and up to Grade 4 level mathematics ability . The test is administered orally for each student, by trained enumerators, and lasts up to ten minutes, following standard administration protocols laid out by ASER.[4] In addition to the ASER scores, we also collected exam scores for all subjects in the national standardized grade 10 examination in 2018 and matched them to students in the sample. Lastly, we conducted a short survey with students' parents to better understand parental involvement in, attitude towards, and support for the child's education, measured as cumulative category scores using a series of composite Likert scales.

[4] Information on the modalities of conducting the ASER test and tools is retrieved from aser-centre.org/Survey/Basic/Pack/Sampling/History/p/54.html

3.1 Sample description

Panel A of Table 1 describes the characteristics of the 1,015 students in our baseline sample in column (2). Aronud 40% of students are female, and students are on average 14-15 years of age (which is appropriate for grade 9 students). 77.4% of students come from socially disadvantaged backgrounds (Scheduled Castes, Scheduled Tribes and Other Backward Castes). From panel B, 27% of students have parents who are illiterate, and more than 80% of households earn a monthly income less than Rs. 10,000 (approximately 550 USD adjusted for purchasing power). Columns (3) and (4) of table 1 then report the averages separately for students in the control and treatment schools, while column (5) reports the difference and its statistical significance at conventional levels.

Panel C describes the cognitive, non-cognitive ability of students as well as the parental support they receive. As described before, we use the ASER tool to measure basic cognitive ability in reading and maths (details of scoring in Appendix C). The reading scores, for example, indicate that students are able to read words and short sentences on average, but cannot read paragraphs. Similarly, on basic maths, they can do simple operations of addition/subtraction, but not more complicated ones such as multiplication/division. These scores indicate that an average Grade 9 student has only mastered skills usually taught in Grades 2 and 3, and are lagging far behind in the grade-level skills needed. This finding resonates resonates with other studies in the Indian context that document the wide learning gap at the post-primary level (Pratham, 2016).

It is important to note from column (5), that students in treatment schools perform better than control students at baseline on reading and maths– on average, they are ahead by nearly one level in Reading and around half a level in Math, but receive less parental support and supervision. Since these characteristics are important outcome variables in our analysis, we follow Bruhn and McKenzie (2009) to control for them in all our specifications as discussed on in the next section. Lastly, indices that measure life skills and gender attitudes (see appendix C for their construction) appear to be balanced across students in the treatment and control schools.

4 Results

We now turn to examining the effect of the intervention for an individual i in school s. We estimate the following specification:

$$Y_{isE} = \alpha + \beta T_s + \gamma_1 Y_{isB} + \delta_1 X_i + \delta_2 X_s + \varepsilon_{is}$$

where Y_{isE} and Y_{isB} are the endline and baseline outcome variables of an individual i, T_s is an indicator that takes the value 1 for a treatment school, X_i and X_s are time-invariant characteristics of the individual and school respectively. We use age, gender, religion, dummy for caste (SC/ST, OBC and others), educational qualification of the guardian and household income for X_i and number of boys and girls, indicator of a co-education school and language of instruction for X_s. We estimate the above specification both without and with individual and school controls and report them in panels A and B in each table respectively. Lastly, since the randomization was done with 24 schools, we wild-bootstrap cluster our standard errors at the school level, as suggested by Cameron et al. (2008) for statistical inference with small clusters. This is reported in all tables as the p-val (OLS) value. Furthermore, we also report the p-value from a two-sided randomization inference test, denoted as p-val (RI) in all tables. This test, originally proposed by Fisher (1935) and developed by Young (2019); Heß (2017) allows statistical inference by comparing the realized treatment effect with multiple (500) placebo assignments. This procedure therefore has the advantage of providing inference with correct size, regardless of the sample and cluster size.

4.1 Impact on cognitive ability

We begin by examining the effect of the treatment on cognitive ability as measured by the ASER reading and math scores. Both the reading and math scores can take a value from 1 to 5 as described earlier. We examine the impact of our intervention in improving basic student cognitive ability in three ways as discussed below.

The first set of results are reported in table 2, where panel A reports the results without any individual and school characteristics, while Panel B controls for them. In columns (1) and (2), we standardize the ASER scores to have mean zero and standard deviation 1 in the control group, so that the treatment effect can be interpreted in standard deviations. We make two observations before discussing the results. First, baseline and endline scores are strongly positively correlated and second, controlling for individual and school characteristics does not substantially alter the magnitude and statistical significance of the intervention. Turning to the impact of the treatment, from Panel B (our preferred specification), after controlling for baseline ASER scores and characteristics, students in treatment schools have a 0.61 and 0.52 standard deviations higher ASER reading and math scores respectively as compared to students in the control school at the endline. This is equivalent to a 40.7% and 15.5% gain in the reading and math scores respectively vis-a-vis the control group students. As a benchmark, this impact compares favorably with similar programs evaluated in the Indian context using ASER, such as Pratham's Read India campaign (Banerjee et al., 2016) and Mindspark program (Muralidharan et al., 2019). Since the ASER scores are discrete and ordinal, we also use an alternate specification and report results from a rank-ordered logistic regression in columns (3) and (4). The results and their inference are qualitatively similar for both outcome variables.

Lastly, in table A1, we transparently report the difference in the endline cognitive level by comparing treatment and control students with the *same* cognitive level at baseline. For example in column (1), we restrict the sample of students (in both treatment and control schools) to only those who could read at least a word at baseline. We then find that students in treatment schools had a 0.88 higher ASER reading score at endline than their control school counterparts. Similarly in column (2), treatment school students who could read a paragraph at baseline had a 0.7 higher ASER score at endline as compared to their control school counterparts who could also read a paragraph at baseline. Lastly in column (3), we see no difference between treatment and control students who could read a story at baseline. Similarly, turning to Math scores in columns (4)-(6), we find that students in the treatment schools have a higher math score at endline, as compared to their control school counterparts who start at the same baseline math level. Put together, we conclude from the above analysis that the Avanti program intervention robustly resulted in substantial improvements in basic reading and maths cognition of students.

4.2 Impact on non-cognitive skills and parental support

Apart from examining the impact on cognitive ability, we also examine if the intervention affected non-cognitive skills, gender attitudes and parental support to the child. These might be particularly important in the context where remedial education intended for cognitive ability has any spillovers on non-cognitive ability as well as parental involvement in the child's education. We conduct a battery of standard tests to capture non-cognitive skills and gender attitudes along with survey data to understand the parental role in helping their child in school (see Appendix C for details). Given the richness of this data collection process, the primary empirical challenge is that a large set of outcomes could be potentially affected by the intervention. The richness of the data implies a danger of 'cherry-picking' outcomes that show large treatment effects. We deal with these systematically in three ways: first, we aggregate all components of a family of outcomes to form an aggregate index, such as for grit or for decision-making. Second, we undertake a principal component analysis to use the underlying variation that drives these family of outcomes. Third, we follow Banerjee et al. (2010); Kling et al. (2007) and calculate the average standardized effect over the family of outcomes. For example, for a family with N different outcomes, each denoted by n, the average effect of the treatment $\hat{\beta}$ is calculated as:

$$\hat{\beta} = \frac{1}{N} \sum_{n=1}^{N} \frac{\hat{\beta}_n}{\hat{\sigma}_n}$$

where $\hat{\sigma}_n$ is the standard deviation of the control group for outcome n. The system across all N outcomes is estimated using seemingly unrelated regression models to account for correlation among the coefficients for all outcomes in one family and the variance-covariance system is used to calculate the standard error of the estimate. We consider three sets of outcomes described below.

First, we examine the impact of our intervention on parental support. This consists of components (evaluated on a scale of 1-5) on how involved parents were with school events, homework, participating in school activities and talking to the child about school. We also examine the expenditure of the household on school material (such as books etc.). Column (1) of table 3 show the effect on the index (aggregated across all components) while column (4) reports from using the principal component analysis. Like before, panels A and B provide results without and with the individual and school controls respectively. Column (1) in table 4 reports the average standardized effect. Lastly, table A2 in the appendix A reports on the impact of the intervention on each component of the parent support index. We see robustly across all methods that there is no effect of the treatment on parental support to the child. Even the estimated coefficients are very small in magnitude (as compared to for example, the mean in the control group). The statistical inference is also robust to using the randomization inference test, as reported by the p-val (RI) values in the table.

Second, we conduct a set of standard non-cognitive skill tests. These ‘life-skills’ include communication skills, critical thinking, decision making, goal setting, grit and problem solving. Each metric is measured on a scale of 1 to 5 on a number of questions. Similar to the previous case, we report the index (aggregated across all components) in column (2) of table 3, the principal component in column (5) of the same table and the average standardized effect in column (2) of table 4. Furthermore, we also report the impact of the intervention on each component separately (instead of the index) in table A3 in appendix A. Again, we find no robust impact of the intervention on any measure of non-cognitive skills.

Lastly, we turn to examining whether the intervention had an impact on student attitudes towards gender. We ask students on how strongly they agree or disagree with respect to gender roles, equality in opportunities, higher studies, opportunities to boys and education of a girl (see appendix C for detailed questions). We report the index (aggregated across all components) in column (3) or table 3, the principal component in column (6) of the same table and the average standardized effect in column (3) of table 4. Lastly, we also report the impact of the intervention on each component separately in table A4 in appendix A. We do not find a systematic impact of the intervention on gender attitudes.

To briefly summarize, we find that while the intervention had large and robust impacts in improving basic cognitive ability of students, it did not change their non-cognitive abilities, gender attitudes or parental support.

4.3 Impact on standardized national exam scores

It is mandatory for all students in India completing grade 10 to write a standardized national exam– the All India Secondary School Examination (AISSE). Students are tested in five subjects, namely: Math, English, Science, Social Sciences and the local language (Tamil, in our case). The AISSE exam is an important milestone in a student’s education because these scores are used for future admissions in higher education as well for job applications early in the career. Given the large improvement in basic cognitive ability we observe as a result of our intervention, we now turn to examining whether it translates into an improvement in the AISSE scores. Along with the list of individual characteristics, we now also control for the baseline cognitive ability. For the students in our sample, we obtain the list of their AISSE scores for all the five subjects in all our 24 schools. Each exam score is normalized to have mean 0 and standard deviation 1 in the control group. Similar to the format so far, results are reported without individual and school controls in panel A of table 5 and with these controls in panel B. Firstly, we see that baseline ASER reading and math scores are strongly positively correlated with the AISSE exam scores, but after controlling for them, the treatment does not have any significant impact on these exam scores. If anything, the estimated coefficients are negative (and statistically significant in the case of Tamil). Moreover, the statistical inference is robust to using randomizaiton inference tests, as reported by the p-val (RI) values. This implies that while the intervention does robustly improve basic cognitive skills, it fails to translate into significant gains in standardized national exam scores one year later. The mechanism behind why this is the case is a question beyond the scope of this paper, and left for future research endeavors.

4.4 Attrition

As mentioned previously, we had a sample of 1015 students at baseline, but were only able to measure outcomes for 880 students at the end of the academic year (an attrition of 13.3%). This attrition is caused both due to students dropping out of school as well as transfers to other schools, for example, due to family migration. In section D in the appendix, we check for differential attrition between control and treatment groups, and see if this affects our estimates of treatment effects. Specifically, we do so in two ways: first, we examine whether student and parent characteristics, cognitive and non-cognitive ability and parental support are differentially correlated with dropouts across treatment and control schools. We find no evidence of differential attrition. Second, to account for potential endogenous attrition from the sample, we estimate Lee bounds on the treatment effects Lee (2009). These bounded estimates are consistent with our main analysis. Table D2 shows that the attrition-adjusted lower bound on the point estimate is 0.27 for Reading and 0.63 for Maths. This increases our confidence in the robustness of our results after taking into account the attrition of students from our sample.

5 Conclusion

This paper reports on the results of an evaluation of a remedial class program in India, which was specifically targeted towards post-primary school students in grade 9. We find significant learning deficiencies at baseline, with students having reading and math levels at grade 2 or 3 on average. We find that the program, led by facilitators (rather than teachers) with technological inputs and academic materials, does succeed in improving basic learning outcomes of the students. The effect sizes are non-trivial, and show the potential of delivery of programs which rely on education technology combined with in-person facilitation and teamwork. However, these gains fail to translate into any improvement on a national standardized Grade 10 exam, indicating that such programs targeted at the post-primary level may not be sufficient to overcome the accumulated learning deficiencies at grade-level. These results are in line with other recent studies such as Beg et al. (2020), who also find that learning outcomes do not translate into gains in official exams in Pakistan, a context similar to India. As Pritchett and Beatty (2012) note, the regular grade-level curriculum may be excessively difficult for learners in these settings, who have continued to accumulate deficiencies over many years and still to need to build on basic skills. Directions for future research could include understanding and unpacking which elements of the program were most effective, including targeting of the program towards younger students. While this was not possible in our research design given the limited sample size and combination of elements in the Avanti program, there is value in disentangling the contribution of its individual elements as well as targeting. Nevertheless, our findings indicate the urgency to address learning levels more urgently at the primary level itself, and rethink post-primary educational initiatives in the context of high learning gaps. This also reinforces the importance of models of 'Teaching at the Right Level' (Banerjee et al., 2016), and ensuring that learning is driven based on the child's needs rather than the dictates of the curriculum. These will be important lessons to inform the implementation of India's newly drafted National Education Policy (2020), which highlights the importance of foundational numeracy and literacy, and the role of education technology for achieving educational goals.

References

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Delivering Education: A Pragmatic Framework for Improving Education in Low-Income Countries," *World Bank Policy Research*, 2015.

Angrist, Noam, Simeon Djankov, Pinelopi K. Goldberg, and Harry A. Patrinos, "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications," *Educating Global Practice & Development Economics*, 2019, *8742.*

Banerjee, Abhijit V. and Esther Duflo, *Poor economics: A radical rethinking of the way to fight global poverty*, Public Affairs, 2011.

_ , Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani, "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India," *American Economic Journal: Economic Policy*, 2010, *2* (1), 1–30.

_ , _ , James Berry, Esther Duflo, Harini Kannan, Shobhini Mukherji, Marc Shotland, and Michael Walton, "Mainstreaming an effective intervention: Evidence from randomized evaluations of Teaching at the Right Level in India," Technical Report, National Bureau of Economic Research 2016.

_ , Shawn Cole, Esther Duflo, and Leigh Linden, "Remedying education: Evidence from two randomized experiments in India," *The Quarterly Journal of Economics*, 2007, *122* (3), 1235–1264.

Bank, World, "World Development Report," 2018.

Barkman, Susan and Krisanna Machtmes, "Solving Problems Survey," 2002.

Beg, Sabrin, Adrienne M. Lucas, Waqas Halim, and Umar Saif, "Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not," 2020.

Blimpo, Moussa P., "Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin," *American Economic Journal: Applied Economics*, 2014, *6*, 90–109.

Bruhn, Miriam and David McKenzie, "In pursuit of balance: Randomization in practice in development field experiments," *American economic journal: applied economics*, 2009, *1* (4), 200–232.

Cameron, Colin A., Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Cilliers, Jacobus, Isaac Mbiti, and Andrew Zeitlin, "Can Public Rankings Improve School Performance? Evidence from a Nationwide Reform in Tanzania," *IZA Institute of Labor Economics*, 2019, *12172.*

Dhar, Diva, Tarun Jain, and Seema Jayachandran, "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India," Technical Report, National Bureau of Economic Research 2018.

Duckworth, Angela and Patrick Quinn, "Development and validation of the Short Grit Scale (GritS)," *Journal of Personality Assessment*, 2009, *91*, 166–174.

Evans, David K. and Anna Popova, "What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews," *The World Bank Research Observer*, 2016, pp. 242–270.

Fagen, Adam P., Catherine H. Crouch, and Eric Mazur, "Peer Instruction: Results from a Range of Classrooms," *The Review of Economics and Statistics*, 2009, *91* (3), 437–456.

Fisher, Ronald A, "The Design of Experiments," 1935.

Glewwe, Paul and Karthik Muralidharan, "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications," *Handbook of the Economics of Education*, 2016, *5*, 653–743.

Heß, Simon, "Randomization inference with Stata: A guide and software," *The Stata Journal*, 2017, *17* (3), 630–651.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Kremer, Michael, Conner Brannen, and Rachel Glennerster, "The Challenge of Education and Learning in the Developing World," *Science*, 2013, *340* (6130), 297–300.

_ , Edward Miguel, and Rebecca Thornton, "Incentives to learn," *The Physics Teacher*, 2009, *40*, 206–209.

Kumar, Naveen, "Public School Quality and Student Outcomes: Evidence from Model Schools in India," 2020.

Lee, David, "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *Review of Economic Studies*, 2009, *3* (76), 1071–1102.

Mbiti, Isaac and Dominic Rodrigue-Segura, "Back to basics: Curriculum Reform and Student Learning in Tanzania," 2020.

_ , Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani, "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania," 2019, *24876.*

Mincemoyer, Claudia and Daniel Perkins, "Youth skills evaluation project at Penn state," 2001.

Muralidharan, Karthik, "Priorities for primary education policy in Indias 12th five-year plan," *India Policy Forum*, 2013, *9* (1), 1–61.

_ , Abhijeet Singh, and Alejandro J Ganimian, "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 2019, *109* (4), 1426–60.

_ and Nishith Prakash, "Cycling to school: Increasing secondary school enrollment for girls in India," *American Economic Journal: Applied Economics*, 2017, *9* (3), 321–50.

_ and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 2011.

Pratham, "Annual Status of Education Report," Technical Report, Pratham 2016.

Pritchett, Lant and Amanda Beatty, "The Negative Consequences of Overambitious Curricula in Developing Countries," *American Economic Journal: Applied Economics*, 2012.

Rogers, Maria, Clarisa Markel, Jonathan D. Midgett, Bruce A. Ryan, and Rosemary Tannock, "Measuring Children's Perceptions of Parental Involvement in Conjoint Behavioral Consultation," *Assessment for Effective Intervention*, 2013.

Schell, Julie, Brian Lukoff, and Eric Mazur, "Catalyzing Learner Engagement Using Cutting-Edge Response Systems in Higher Education," 2013.

UNESCO, "More Than One-Half of Children and Adolescents Are Not Learning Worldwide," 2017.

West, Martin, Matthew Kraft, Amy Finn, and Angela Duckworth, "Promise and Paradox: Measuring Students Non-cognitive Skills and the Impact of Schooling," 2014.

Young, Alwyn, "Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results," *The Quarterly Journal of Economics*, 2019, *134* (2), 557–598.

Zhang, Ping, Lin Ding, and Eric Mazur, "Peer Instruction in introductory physics: A method to bring about positive changes in students attitudes and beliefs," *Physical Review Physics Education Research*, 2017.

6 Tables

Table 1: Sample characteristics

	N	Whole sample	Control	Treatment	(4)-(5)
	(1)	(2)	(3)	(4)	(5)
Panel A: Student characteristics					
Female	1015	0.403	0.400	0.406	0.005
		[0.015]	[0.021]	[0.022]	
Age	1015	14.824	14.789	14.860	0.071
		[0.026]	[0.038]	[0.037]	
Hindu	1015	0.730	0.707	0.755	0.048*
		[0.014]	[0.020]	[0.019]	
SC/ST caste	1015	0.367	0.349	0.385	0.037
		[0.015]	[0.021]	[0.022]	
OBC	1015	0.407	0.408	0.406	-0.002
		[0.015]	[0.022]	[0.022]	
Panel B: Guardian characteristics					
Illiterate	1015	0.268	0.266	0.270	0.003
		[0.014]	[0.019]	[0.020]	
Any schooling	1015	0.708	0.713	0.704	-0.009
		[0.014]	[0.020]	[0.021]	
Income category	692	2.168	2.189	2.148	-0.041
		[0.025]	[0.040]	[0.030]	
Income $<$ Rs. 10k	692	0.825	0.805	0.844	0.038
		[0.014]	[0.022]	[0.019]	
Panel C: Student cognitive, non-cognitive and parental support					
ASER Reading	619	3.864	3.426	4.391	0.965***
		[0.046]	[0.059]	[0.058]	
ASER Maths	619	4.347	4.172	4.559	0.387***
		[0.033]	[0.050]	[0.039]	
Parental support index	852	0.114	0.209	0.010	-0.199**
		[0.044]	[0.063]	[0.062]	
Life skills index	852	0.031	0.032	0.031	0.001
		[0.061]	[0.082]	[0.093]	
Student attitude index	852	0.018	0.052	-0.020	-0.072
		[0.047]	[0.066]	[0.066]	

Notes: See Appendix C for measurement of student ability variables. Column (5) reports the difference between treatment and control and the asteriks report a t-test of whether this difference is statistically significant. * denotes significance at 0.1 level, ** at 0.05 and *** at 0.01 level.

Table 2: Impact on cognitive ability

	Standardized ASER Score		ASER Score	
	Reading	Maths	Reading	Maths
	(1)	(2)	(3)	(4)
	OLS	OLS	O. Logit	O. Logit
PANEL A: Without individual and school controls				
Treatment	0.607***	0.501***	1.412***	1.723***
	(0.135)	(0.110)	(0.315)	(0.344)
Baseline	0.620***	0.594***	1.237***	1.740***
	(0.0645)	(0.0481)	(0.137)	(0.183)
p-val (OLS)	0.00	0.00	0.00	0.00
p-val (RI)	0.00	0.00	0.00	0.00
R^2	0.46	0.45		
PANEL B: With individual and school controls				
Treatment	0.608***	0.520***	1.502***	1.922***
	(0.143)	(0.125)	(0.392)	(0.567)
Baseline	0.606***	0.561***	1.233***	1.700***
	(0.0600)	(0.0511)	(0.142)	(0.195)
p-val (OLS)	0.00	0.00	0.00	0.00
p-val (RI)	0.00	0.00	0.00	0.00
R^2	0.484	0.484		
N	674	674	674	674

Notes: The outcome variable in columns (1) and (2) are the ASER scores for Reading and Maths that have been standardized to have mean 0 and standard deviation 1 for the control group. Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income categories. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

Table 3: Impact on non-cognitive ability and parental support

	Parent Support	Life Skills	Student Attitudes	Parent Support	Life Skills	Student Attitudes
	Index	Index	Index	PCA	PCA	PCA
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Without individual and school controls						
Treat	0.295	0.160	-0.0514	0.0975	0.107	0.176
	(0.367)	(0.178)	(0.282)	(0.115)	(0.118)	(0.133)
Baseline	0.277***	0.205***	0.194***	0.280***	0.213***	0.324***
	(0.0320)	(0.0378)	(0.0409)	(0.0313)	(0.0385)	(0.0190)
p-val (OLS)	0.422	0.369	0.855	0.397	0.367	0.186
p-val (RI)	0.470	0.376	0.858	0.398	0.382	0.190
R^2	0.069	0.043	0.025	0.071	0.046	0.097
PANEL B: With individual and school controls						
Treat	0.0989	0.153	-0.137	0.0397	0.0965	0.170
	(0.411)	(0.224)	(0.352)	(0.138)	(0.167)	(0.142)
Baseline	0.282***	0.198***	0.144***	0.284***	0.206***	0.269***
	(0.0290)	(0.0365)	(0.0461)	(0.0293)	(0.0351)	(0.0233)
p-val (OLS)	0.810	0.494	0.697	0.774	0.562	0.230
p-val (RI)	0.822	0.456	0.686	0.742	0.492	0.222
R^2	0.116	0.063	0.122	0.116	0.067	0.161
N	852	852	852	852	852	852
Control mean	14.08	22.94	14.97	-0.135	-0.0254	-0.187

Notes: For each of the three composites, we report the impact on an aggregate index in columns (1) to (3), while principal component of each composite is reported in columns (4) to (6). Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

Table 4: Average standardized effects on non-cognitive ability

	Parent Support	Life Skills	Student Attitudes
	(1)	(2)	(3)
Mean	0.053	0.051	0.041
Std. Dev.	0.039	0.038	0.03
N	852	852	852

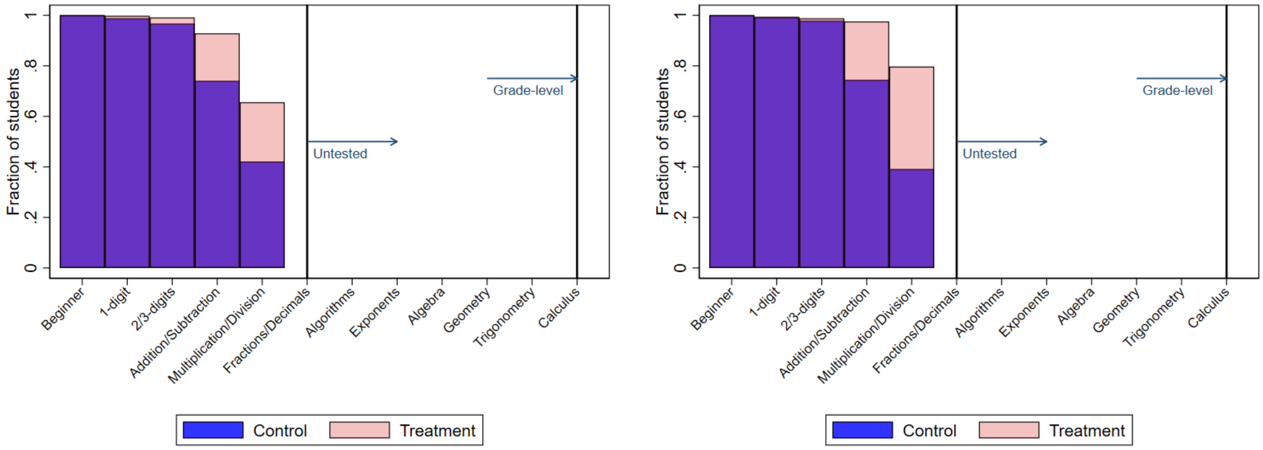
Notes: All outcomes presented in this table are aggregate measures composed of relevant survey questions. Details are presented in Appendix 2.

Table 5: Impact on standardized national exam scores

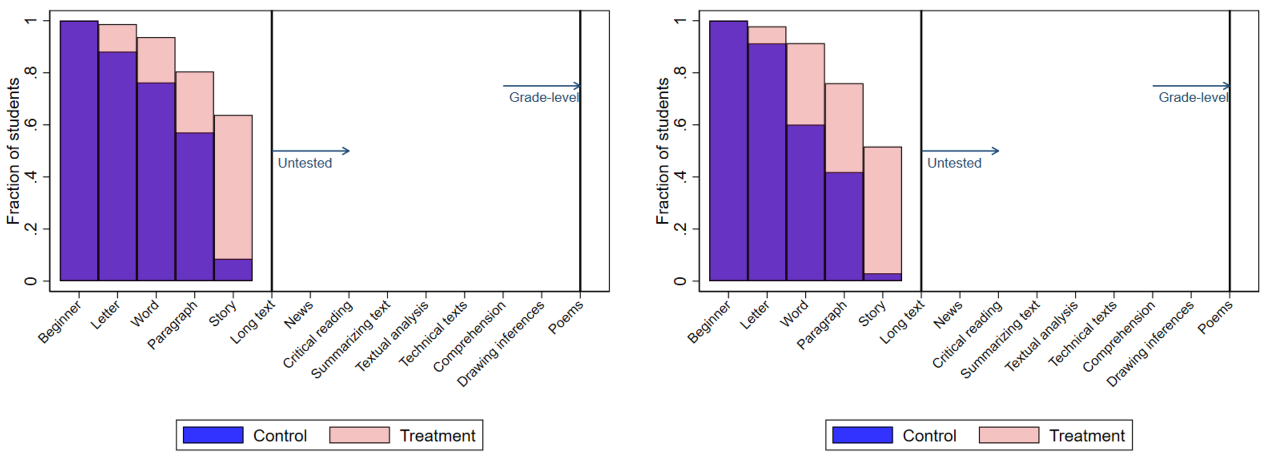
Z-scores:	Tamil	English	Maths	Science	Social Science	Total
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.382**	-0.253	-0.0355	-0.222	-0.226	-0.278
	(0.172)	(0.209)	(0.252)	(0.213)	(0.166)	(0.195)
Base Read	0.301***	0.204***	-0.0358	0.223***	0.196***	0.229***
	(0.0638)	(0.0527)	(0.0943)	(0.0530)	(0.0546)	(0.0644)
Base Math	0.300***	0.312***	0.264***	0.260***	0.248***	0.335***
	(0.0618)	(0.0541)	(0.0574)	(0.0678)	(0.0419)	(0.0592)
p-val (OLS)	0.026	0.23	0.89	0.296	0.17	0.154
p-val (RI)	0.012	0.31	0.91	0.22	0.13	0.09
R^2	0.318	0.251	0.137	0.232	0.205	0.268
N	668	668	668	668	668	668

Notes: All outcome variables have been normalized to have mean 0 and standard deviation 1 in the control group. 'Base Read' and 'Base Math' are the students' baseline scores on ASER Reading and Maths. Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

A Appendix Figures and Tables



(a) Maths scores



(b) Reading scores

Figure A1: Reading and math scores in control and treatment schools

Table A1: Impact on different baseline levels

	Reading level			Math Score		
Baseline level:	Word	Paragraph	Story	Digits	Subtraction	Division
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.884***	0.694***	0.299	0.404	0.699***	0.315***
	(0.328)	(0.239)	(0.246)	(0.312)	(0.108)	(0.122)
p-val (OLS)	0.00	0.00	0.22	0.195	0.00	0.01
p-val (RI)	0.00	0.00	0.040	0.042	0.00	0.00
R^2	0.125	0.087	0.018	0.035	0.246	0.091
N	189	210	220	98	188	333
Control mean	2.26	3.21	4.15	3.23	4.00	4.58

Notes: Each column restricts the sample of students to that level of reading/maths at baseline and examines the impact of the treatment on endline scores. Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

Table A2: Impact on Parental Support

	Events	Homework	School	Talk
	(1)	(2)	(3)	(4)
PANEL A: Without individual and school controls				
Treatment	0.0847	0.0405	0.00833	0.113
	(0.123)	(0.121)	(0.154)	(0.137)
Baseline	0.223***	0.182***	0.0597**	0.222***
	(0.0304)	(0.0276)	(0.0271)	(0.0287)
p-val (OLS)	0.490	0.739	0.957	0.411
p-val (RI)	0.494	0.748	0.960	0.414
R^2	0.044	0.029	0.004	0.045
PANEL B: With individual and school controls				
Treatment	0.0362	0.0851	-0.0868	0.0421
	(0.126)	(0.148)	(0.183)	(0.178)
Baseline	0.223***	0.191***	0.0633**	0.216***
	(0.0313)	(0.0295)	(0.0278)	(0.0322)
p-val (OLS)	0.774	0.565	0.635	0.813
p-val (RI)	0.802	0.540	0.654	0.810
R^2	0.075	0.058	0.042	0.075
N	852	852	852	852
Control mean	3.494	3.982	3.313	3.286

Notes: Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

Table A3: Impact on Non-Cognitive Life-skills

	Communication	Critical Thinking	Decision	Goal Setting	Grit	Problem Solving
	(1)	(2)	(3)	(4)	(5)	(6)
PANEL A: Without individual and school controls						
Treatment	0.0139	0.0371	-0.00247	-0.0326	0.0627	0.0801**
	(0.0388)	(0.0513)	(0.0469)	(0.0457)	(0.0452)	(0.0343)
Baseline	0.160***	0.212***	0.163***	0.119***	0.0861**	0.117***
	(0.0317)	(0.0256)	(0.0323)	(0.0364)	(0.0345)	(0.0427)
p-val (OLS)	0.719	0.470	0.958	0.476	0.166	0.0190
p-val (RI)	0.748	0.464	0.972	0.488	0.140	0.0640
R^2	0.022	0.044	0.028	0.016	0.013	0.020
PANEL B: With individual and school controls						
Treatment	0.0171	0.0242	0.0126	-0.0573	0.0737*	0.0858**
	(0.0466)	(0.0698)	(0.0580)	(0.0650)	(0.0424)	(0.0431)
Baseline	0.150***	0.221***	0.147***	0.111***	0.0777**	0.115***
	(0.0345)	(0.0226)	(0.0344)	(0.0362)	(0.0355)	(0.0410)
p-val (OLS)	0.714	0.729	0.828	0.378	0.0820	0.0460
p-val (RI)	0.722	0.696	0.820	0.294	0.156	0.0320
R^2	0.041	0.076	0.046	0.048	0.046	0.039
N	852	852	852	852	852	852
Control mean	3.779	3.876	3.993	3.987	3.484	3.824

Notes: Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

Table A4: Impact on Gender Attitudes

	Roles	Equal Opp.	Study	Boys Opp.	Wife Education
	(1)	(2)	(3)	(4)	(5)
PANEL A: Without individual and school controls					
Treatment	0.00237	-0.0605	-0.00768	0.264***	0.0891
	(0.127)	(0.0758)	(0.0671)	(0.0904)	(0.0974)
Baseline	0.140***	0.0766**	0.189***	0.159***	0.204***
	(0.0330)	(0.0335)	(0.0472)	(0.0353)	(0.0279)
p-val (OLS)	0.985	0.424	0.909	0.00300	0.360
p-val (RI)	0.982	0.448	0.920	0.00800	0.400
R^2	0.021	0.007	0.034	0.036	0.047
PANEL B: With individual and school controls					
Treatment	0.0228	-0.0806	-0.00798	0.214**	0.0903
	(0.131)	(0.0861)	(0.0759)	(0.101)	(0.113)
Baseline	0.126***	0.0800**	0.181***	0.0839**	0.165***
	(0.0331)	(0.0316)	(0.0458)	(0.0374)	(0.0287)
p-val (OLS)	0.862	0.350	0.916	0.0340	0.423
p-val (RI)	0.860	0.344	0.932	0.0320	0.364
R^2	0.077	0.034	0.061	0.111	0.101
N	852	852	852	852	852
Control mean	2.204	4.123	4.268	2.506	2.036

Notes: Individual controls include gender, age, religion, caste and dummy variables for guardians' education and income. School controls include language of instruction, size of class and an indicator variable for whether the school is co-educational. 'Baseline' captures the students' responses to the survey question at baseline. Wild-bootstrapped standard errors are clustered at the school-level and reported in parentheses. p-val (OLS) reports the p-value for the treatment coefficient estimated by the wild-bootstrapped clustered standard errors, while p-val (RI) reports the p-value using randomized inference method. * denotes significance at 0.1, ** at 0.05 and *** at 0.01 respectively.

B Program Description

Table B1: Avanti Program Components Description

Technology: 15 short five-minute videos in the local language that explain concepts through the use of real world examples.	1 hour 15 minutes (15 videos) per week
Peer-learning: Students organized in groups of 6 and tasked with collaboratively solving worksheets. Facilitators intervene to encourage discussion and resolve questions only when necessary.	1 hours 30 minutes (6 worksheets) per week
Trained Facilitators: Facilitators (from the local area) led lectures to summarize key concepts based on facts emerging from an in-class discussion.	1 hour 40 minutes (20 min per day) per week

C Survey Questions and Indicators

C.1 ASER

ASER Levels	**ASER Reading**	**Grade level**	**ASER Maths**	**Grade level**
Level 1	Beginner		Beginner	
Level 2	Identifying letters	$<$ Grade 1	Number recognition 0 - 9	$<$ Grade 1
Level 3	Reading words	$<$ Grade 1	Number recognition 10 - 99	$<$ Grade 1
Level 4	Reading a paragraph	Grade 1	Subtraction with borrowing	Grades 2 or 3
Level 5	Reading a story	Grade 2	Division with remainder	Grades 3 or 4

Table C1: Curriculum Mapping for ASER Tools

C.2 Non-cognitive

Data on students' non-cognitive skills were collected using several composite indicators captured in our survey. The dimensions along which we capture non-cognitive indicators include parents' support (Rogers et al. (2013)), student attitudes toward gender (Dhar et al. (2018)), and various life skills like decision making, critical thinking (Mincemoyer and Perkins (2001)), communication, goal setting, problem solving (Barkman and Machtmes (2002)), and grit (Duckworth and Quinn (2009)). The reponses are measured on a Likert Scale ranging from Never (1) to Always (5). For the analysis, we study the treatment's impact on the individual metrics, aggregate scores as well as the principal component that retains most of the sample's information. The tables below detail the variable label associated with each composite question in the analysis and the associated question from the survey.

Table C2: Parent Support and Student Attitude Questions

Variable	**Survey Question**
Panel A: Parental Support	
School	How often do your parents help you with your school work?
Talk	How often do your parents talk to you about what you are doing in school?
Homework	How often do your parents ask you about homework?
Events	How often do your parents go to meetings or events at school?
Panel B: Student attitudes towards gender	
Roles	A woman's most important role is to take care of her home, feed kids and cook for her family.
Equal Opp.	Men and women should get equal opportunities in all spheres of life - education, healthcare, food, decision making.
Study	Girls should be allowed to study as far as they want.
Boys Opp.	Boys should be allowed to get more opportunities and resources for education than girls.
Education	Wives should be less educated than their husbands.

Table C3: Life Skills Composite Questions

Variable	**Survey Question**
Decision Making	I look for information to help me understand the problem I consider the risk of a choice before making a decision. I think about all the information I have about the different choices I think of past choices when making new decisions
Critical Thinking	I can easily express my thoughts on a problem. I usually have more than one source of information before making a decision. I compare ideas when thinking about a topic. I keep my mind open to different ideas when planning to make a decision. I am able to tell the best way of handling a problem.
Communication	I try to keep eye contact. I recognize when two people are trying to say the same thing, but in different ways. I try to see the other person's point of view I change the way I talk to someone based on my relationship with them. I organize thoughts in my head before speaking. I make sure I understand what another person is saying before I respond.
Goal Setting	I look at the steps needed to achieve the goal. I think about how and when I want to achieve a goal. After setting a goal, I break goals down into steps so I can check my progress. Both positive and negative feedback help me work toward my goal
Problem Solving	I first figure out exactly what the problem is. I try to determine what caused the problem. I do what I have done in the past to solve the problem I compare each possible solution with the others to find the best one. After selecting a solution, I think about it for a while before putting it into action
Grit	New ideas sometimes distract me from previous ones Setbacks don't discourage me. I bounce back from disappointments faster than most people. I have been obsessed with a certain idea or project for a short time but later lost interest. I am a hard worker. I often set a goal but later choose to pursue (follow) a different one. I have difficulty maintaining (keeping) my focus on projects that take more than a few months to complete. I finish whatever I begin. I am diligent (hard working and careful).

C.3 Controls

Table C4: Controls Description

	Variable Name	**Variable Description**
Individual Controls	Student Gender	Indicator; Male/Female
	Student Age	Continuous; Years
	Religion	Indicator; Hindu/Non-Hindu
	Caste	Indicators for SCST, General, and OBC castes
	Guardian Education	Categorical; Illiterate, Schooling
	Guardian Income	Categorical; <10,000, =>10,000
School Controls	Language of Instruction	Indicator; Tamil/English
	Size of class	Continuous, by gender
	Coeducational School	Indicator; Coed/Single Sex

D Attrition

Of the 1015 students in our baseline sample we only have 880 students in a the endline (an attrition of 13.3%). We investigate the differential attrition across treatment and control schools on observable characteristics as well as its impact on our treatment estimates in two ways. First, we report whether there is differential attrition between treatment and control. We estimate the following specification:

$$M_i = \alpha + \beta X_i + \gamma T_i + \delta T_i \times X_i + \varepsilon_i$$

where M_i is a dummy variable that takes the value 1 if the student stays in the sample and 0 if the student drops out between the baseline and endline. δ is the coefficient of interest since it tells us whether characteristics of students who stay in the sample is statistically different on average between control and treatment schools. We report the estimates for δ in table D1. As reported in the table, attrition is small in magnitude and statistically insignificant between treatment and control schools for most student characteristics.

Table D1: Differential attrition between control and treatment

Student characteristic (X_i)	$\hat{\delta}$	S.E.	p-value
	(1)	(2)	(3)
Female	-0.008	0.05	0.858
Age	0.053**	0.0271	0.05
Hindu	0.079	0.052	0.127
OBC	-0.01	0.045	0.827
Guardian Educ: Illiterate	-0.12	0.163	0.457
Guardian Educ: Schooling	-0.084	0.16	0.596
Guardian Income < Rs.10,000	-0.006	0.051	0.907
Guardian Income >= Rs.10,000	0.09	0.074	0.228
Baseline Reading ability	0.009	0.024	0.698
Baseline Maths ability	-0.025	0.019	0.17
Lifeskills (PC)	-0.0013	0.013	0.919
Gender Attitudes (PC)	-0.0231	0.018	0.204
Parent Support	-0.004	0.006	0.487
Education exptd. ('000)	0.0051	0.0032	0.109

Notes: * denotes significance at 0.1 level, ** at 0.05 and *** at 0.01 level.

In addition, to account for potential endogenous attrition from the sample, we also estimate Lee bounds on the treatment effects with bootstrapped standard errors (Lee (2009)) and report the results in table D2. Where possible, baseline values are used to tighten bounds, identified as Group 1 in table D2. In other cases (Group 2), bounds estimated using this method are tightened using gender, religion, and guardian education and income categories as covariates [5], with weighted averages of bounds defined for sub samples defined by these covariates. The table D2 shows that the attrition-adjusted upper bound on the point estimate is 0.56 for reading and 0.99 for Maths.

Table D2: Lee Bounds on Treament Effects

	Lower bound		Upper bound	
	Estimate	S.E.	Estimate	S.E.
	(1)	(2)	(3)	(4)
Group 1: Baseline Ability used to tighten bounds				
Panel A: Cognitive and Non-cognitive abilities				
Reading	0.27**	0.11	0.56***	0.11
Maths	0.63***	0.12	0.99***	0.13
Parent Support	0.073	0.34	0.30	0.34
Group 2: Individual Controls used to tighten bounds				
Lifeskills (PC)	-0.028	0.17	0.15	0.18
Gender Attitudes (PC)	0.15	0.075	0.297*	0.098
Panel B: Standardized National Exam Scores				
Tamil	-1.05	2.22	1.87	1.55
English	1.03*	1.47	3.67	1.23
Maths	1.64*	1.23	3.30	0.96
Science	-1.60	1.45	0.098	1.51
Social science	-2.23	1.77	0.28	2.01
Total	-1.18	6.82	8.98	6.65

Notes: * denotes significance at 0.1 level, ** at 0.05 and *** at 0.01 level.

[5] The variables used to tighten bounds are gender, religion, education level of the guardian and household income.