City Research Online

City, University of London Institutional Repository

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/31428/

Link to published version: https://doi.org/10.1016/j.econedurev.2016.08.001

Unintended Consequences of Rewards for Student Attendance: Results from a Field Experiment in Indian Classrooms

Sujata Visaria[a]
Rajeev Dehejia[b]
Melody M. Chao[c]
Anirban Mukhopadhyay[d]

July 29, 2016

Abstract

In an experiment in non-formal schools in Indian slums, a reward scheme for attending a target number of school days increased average attendance when the scheme was in place, but had heterogeneous effects after it was removed. Among students with high baseline attendance, the incentive had no effect on attendance after it was discontinued, and test scores were unaffected. Among students with low baseline attendance, the incentive lowered post-incentive attendance, and test scores decreased. For these students, the incentive was also associated with lower interest in school material and lower optimism and confidence about their ability. This suggests incentives might have unintended long-term consequences for the very students they are designed to help the most.

Keywords: educational economics, incentives, attendance, motivation, experiment

JEL Classifications: I21, I25, O15

[a]Department of Economics, Lee Shau Kee Business Building, Room 6081, Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong. Telephone: +852-23587608, E-mail: svisaria@ust.hk. Corresponding author.

[b]Wagner School of Public Policy, New York University, The Puck Building, 295 Lafayette Street, Room 3004, New York, NY 10012, USA. E-mail: dehejia@nyu.edu

[c]Department of Management, Lee Shau Kee Business Building, Room 5072, Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong. E-mail: mchao@ust.hk

[d]Department of Marketing, Lee Shau Kee Business Building, Room 4002, Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong. E-mail: anirban.mukhopadhyay@ust.hk

1 Introduction

A growing literature examines whether incentives can increase the effort and improve the school performance of students from underprivileged backgrounds (Angrist and Lavy 2009; Kremer, Miguel, and Thornton 2009; Fryer 2011; Bettinger 2012; Levitt et al. 2012). The underlying assumption is that target students have suboptimally low motivation to exert effort at school. This may be because they are unaware of the benefits of schooling, are too impatient to work for benefits that will accrue far in the future, or lack the self-control to trade off current costs against future benefits. A nearer-term incentive that rewards them for say, reading a book or attending school, can provide the "carrot" that will change their behavior.

Problems of impatience and self-control notwithstanding, some students do exert effort and achieve high test scores. The largest gains from incentives are not expected for these students: since they already exert high effort, any gains at the margin will presumably be small. Instead, researchers expect large treatment effects on children whose baseline academic outcomes and motivation are low. For such students, the promise of a large enough reward might create the motivation to do the task and in turn improve academic performance. If the student becomes habituated to the higher effort level, these effects can also sustain after the incentive is removed (Charness and Gneezy 2009).[1]

However it is also possible for incentives to backfire: for example, the extrinsic motivation provided by an incentive could crowd out students' intrinsic motivation to study and learn (Gneezy, Meier, and Rey-Biel 2011).[2] This is because attaching a price to a task that was initially enjoyable can make it less enjoyable (Deci and Ryan 1985). After the incentive is removed, the lack of extrinsic motivation coupled with the lower intrinsic motivation could lower student effort below what it would have been if no reward had been offered.

Two points emerge from this discussion. First, if incentives increase extrin-

---

[1] Charness and Gneezy (2009) find that when university students were given high-powered incentives to attend a gym, they became more likely to exercise even after the incentives were discontinued.

[2] A large literature in psychology also discusses the crowd-out of intrinsic motivation (see for example, Deci, Koestner, and Ryan (1999)).

sic motivation and do not change intrinsic motivation, then they should have the largest (positive) effects on students with low baseline motivation. If instead they do lower intrinsic motivation, then matters are less clear. Presumably students with high intrinsic motivation might have more of it to lose, but the decrease may still not be large enough to change effort or performance. Less motivated students, on the other hand, may be relatively disengaged to start with, and so the crowd-out might worsen their effort and performance. Since most studies have focused on the average effects across these two subgroups, it has been difficult to identify the channels at work.

Second, crowding-out is best detected by studying students' behavior after the reward has been discontinued. Although researchers have examined long-term effects of incentives to exercise, stop smoking, and engage in pro-social behavior (Gneezy, Meier, and Rey-Biel 2011), few papers in education have examined effects after the incentive period ended. A notable exception is Rodriguez-Planas (2012), who examines the effect of the high-school Quantum Opportunity Program in the US two years and five years after the program ended. Although she is unable to identify the mechanisms that caused the positive effects of the program to become smaller over time, she also finds that the fade-out differed by subgroups: long-term educational and employment outcomes were better for treated females, but not for treated males.[3]

In this paper we report on a field experiment where the attendance of students of non-formal schools in Indian slums was monitored and an incentive was offered for meeting an attendance target. To evaluate whether the effect of the incentive varies by students' baseline motivation levels, we examine separately students with low and with high prior attendance rates, both during and after the 39-day reward period. We find that both in the pooled sample as well as within the two subgroups, the incentive increased student attendance while it was in place.[4] However, the two subgroups were affected very differently after the incentive period ended. Students in the incentive group who had high

---

[3]We do not find evidence for such a gender difference. Unlike in Rodriguez-Planas (2012), our intervention did not provide students with additional mentoring or protection against sanctions. In any case, our students are significantly younger, and do not generally engage in risky behaviors where mentoring or (the lack of) sanctions might have differential impacts by gender.

[4]The effect on the low baseline attendance group is large in magnitude but imprecisely estimated.

baseline attendance attended school at the same rate as their counterparts in the control group. However, those with low baseline attendance were even less likely to attend school than they would have been if the incentive had not been offered.

Scores on a test administered three months after the incentive scheme were also affected in the same manner: the test scores of students with high baseline attendance were unaffected by the incentive scheme, but those of students with low baseline attendance became lower than if there had been no incentive at all. The reward also lowered these students' liking for school subjects, and lowered their expectations of themselves. Thus, in contrast to the existing literature, we find that although the incentive motivated students while it was in place, it had unintended negative consequences in the longer term for students with low baseline motivation.

Our results show that it is instructive to examine the effects of incentives for students with low and high initial motivation separately. However, the effects are not in line with the ideas that incentives primarily help students with low motivation, or that they hurt students with high motivation. The incentive appears to have had no long-term effects on students who were highly motivated to begin with. Instead, it had negative long-term impacts on students with low motivation, a group that arguably had the most to gain from improved performance.

The rest of this paper is organized as follows. Section 2 describes the empirical context. Section 3 describes the experimental intervention and data. Section 4 presents the empirical results. Section 5 discusses the implications of the study and concludes.

2   The Empirical Context

Our experiment was conducted in collaboration with Gyan Shala, a non-government organization that runs non-formal education centers (hereafter referred to as "classes") in the slums of Ahmedabad in the state of Gujarat in western India. In 2010, Gyan Shala had 343 classes operating across 5 areas in the city (CfBT Education Services 2010). Each Gyan Shala class caters to a single

grade, and is housed in a single room, usually rented from a local resident. Students pay no fees. The median class in our sample has 22 students, all of whom are from the same or from neighboring slums.[5] Each classroom has basic school supplies. Teaching is mainly lecture-based, but each student has a workbook with exercises to do in school. Three subjects are taught: language (Gujarati), mathematics and science.

Gyan Shala's mission is to provide children of low socioeconomic status a high quality education at a low cost. Operational costs are low because teachers do not have a formal teaching qualification, and therefore would not be hired by formal schools. Most teachers have only a high school diploma. To ensure quality, Gyan Shala trains these teachers intensively: the typical school year includes 30 training days. The teachers closely follow day-wise lesson plans that they receive from a "design team" made up of subject specialists who hold bachelor's or master's degrees. A supervisor visits each class once a week to observe and provides inputs as needed. When students in particular classes find individual topics difficult to understand, design team members visit the classroom to gauge the problem and to help the teacher. The information gathered is fed back into future lesson plans.

The parents of Gyan Shala students are for the most part self-employed or casual workers in the unorganized sector. They have low education levels and therefore limited ability to support their children's learning at home. Gyan Shala hopes to provide these parents with an attractive alternative to the local municipal school, while also demonstrating that a good education need not be expensive. An independent evaluation conducted by Educational Initiatives (EI) in 2010 found that Gyan Shala students outperformed their peers in municipal schools on language, mathematics and science by wide margins (Educational Initiatives Private Limited 2010). On average Gyan Shala students were also better able to answer the more difficult, "non-straightforward" questions on EI's tests. A short-lived experimental intervention where Gyan Shala's teaching techniques were adopted in municipal schools also generated significant impact, with treatment municipal schools outperforming control municipal

---

[5]An important consideration for Gyan Shala is that children be able to walk to school unescorted, since this lowers the time and transport costs of attending school and helps to lower absenteeism.

schools (Educational Initiatives Private Limited 2010).

Gyan Shala's main effort has been to run classes for grades 1, 2 and 3. Our experiment was conducted in grade 2 and grade 3 classes, but we report here only the results for grade 3 classes because those are the only students who took a test administered by Educational Initiatives (EI), that provides us with an independent assessment of their achievement.[6] The EI examination only tested mathematics and science.

The goal of this study was to examine the effect of an incentive for student effort, on student performance.[7] The administrators at Gyan Shala identified attendance as the appropriate task to target. We believe this choice is justified for a couple of reasons. First, research in higher-income countries has shown that student attendance is correlated with performance (Roby 2004; Paredes and Ugarte 2011), and it is likely that this relationship is even stronger in our context, where parents can provide limited support at home. Second, at an unannounced visit made by our investigators two months into the 2011-12 school year, 75 percent of students in sample classes were present.[8] While considerably lower than the standards set by school boards in some developed countries, this number is also not so low that it might be mainly caused by structural factors outside students' control.[9] Gyan Shala administrators believed that a significant factor behind the absence was truancy: students often missed school because they wanted to play instead, it was a festive season, or because their siblings had a day off at their school.

---

[6]Our results are qualitatively unchanged when we include Grade 2 students in the attendance analysis.

[7]This is part of a larger project aimed at understanding the impact of economic and psychological interventions on student achievement. For more detail, see Chao et al. (2016). The psychological intervention was implemented orthogonally to the reward intervention and we do not examine its effect in this paper.

[8]This number matches the 75 percent average attendance rate for Gujarat state reported by previous research (Educational Consultants India Limited 2007).

[9] For example, the California legislature defines as a chronic truant a student who is absent from school without a valid excuse for ten or more percent of school days in one school year (California Department of Education 2015).

3   The Data and the Experimental Intervention

Our study took place during the school year that ran from June 2011 to April 2012. Our sample consists of roughly 12 students randomly sampled from each of 68 grade 3 classes, that are spread evenly across all 5 city zones where Gyan Shala operates. Figure 1 summarizes the sequence of events in our study. Investigators made six unannounced visits to the classrooms; we label these visits Time 0 through Time 5. At all six visits, they took roll call of the sample students to check if they were present.[10] At three of these visits (Time 1, Time 3 and Time 5) they also conducted 10-minute surveys with the sample students. Survey questions were about the students' like and dislike for particular subjects, and their expectations and attitudes about learning and exerting effort on difficult tasks. At Time 1 students were also asked to provide demographic information about themselves and their family members.

An important feature of our interview visits is that we attempted to conduct interviews with all sample students, even if they were absent from school at the time of the visit. Investigators tried to find out when the student might be available, and then made up to 3 follow-up visits either to their homes or to the school, within a window of a few weeks after the original visit. As a result, we have interview data for 79 percent of the students who were absent on the day of the class visit.[11]

Table 1 presents summary statistics from the Time 1 interview, and checks whether there were significant differences between the control and treatment classes. About half of the 799 sample students were female. They were on average 9 years old. Since we did not interview their parents, we had to rely on the children's reports of household assets to infer socioeconomic status. We also measured their height and weight, on the assumption that their body mass in-

---

[10]All visits were scheduled to begin at least an hour after the school day began, so as to not miss latecomers. However, since the Gyan Shala classes are located within the students' own neighborhoods, a teacher could have sent word to summon absent students to class when the investigator arrived. To prevent this from contaminating our attendance measure, we instructed the investigators to code any child who entered the classroom after she had entered it as "E" (for "entered during visit"). In our analysis such students are considered absent.

[11]These students are coded as absent from school for that visit, but their interview data are non-missing.

dex may be correlated with their socioeconomic status. Note however that all children are residents of low-income neighborhoods and so variation in SES is likely to be small. The average child had a body mass index of 13.8, which places them between the 3rd and 5th percentiles of a normal international population (World Health Organization 2007).

Ninety-three percent of children reported that at least one person in their household owned a mobile phone. A quarter reported that their parents had a motorized vehicle. Three-quarters had a toilet in the house, and a little over a third had a VCR or DVD player. Computers were almost non-existent. There were no significant differences between the control and treatment groups on these dimensions.

At the Time 0 visit conducted about 6 weeks after the school year had begun, investigators found 75 percent of the sample students present in class. This is in line with the administrative attendance records, according to which these students were present for 78 percent of days during the first two months of the school year. We do not find significant differences across treatment (mean = 0.02) and control (mean = -0.03) classes in the z-score of the students' scores on the previous year's final exam (conducted by Gyan Shala).

Students told us how much they liked each of the three subjects they were taught, on a 7-point scale. For these questions, we showed them drawings of faces, and first asked them to choose either a smiling, neutral or sad face to indicate how they felt about the subject. If they chose the smiling face, we asked them to choose one of three happy faces where the faces and smiles were small, medium or large, to indicate how intensely they liked it. If instead they chose the sad face, we showed them three unhappy faces to choose from, where the faces, frowns and tears became incrementally larger.

As can be seen, mathematics was very popular among students, with an average rating of 2.5 on a scale ranging from -3 to +3. The difference between control and treatment schools was not significant. Science was relatively less popular, with an average rating of 2. To elicit students' opinions about their ability to pick up new skills, we asked them if they thought they could learn to solve a crossword puzzle. (They knew what crossword puzzles were because they had been introduced to them shortly before the Time 1 interview.) Ninety-six

percent of students answered in the affirmative. We also tried to elicit students' optimism about their ability to rise to an academic challenge. To do this, we told them about a hypothetical child attempting a difficult sum, and asked them to predict the child's performance, on a scale of 1 (low) to 5 (high).[12] If a student predicted the child would perform well, our interpretation is that the student is optimistic that one can succeed at a challenging academic task. If they predicted the child would perform poorly, we say that the student is pessimistic that one can overcome challenges in academic work. The average prediction was 2.2. The difference between treatment and reward schools was not statistically significant. We therefore conclude that the control and treatment groups were balanced on observables.

After the Time 0 (August), Time 1 (September-October) and Time 2 (November) visits had taken place, in December the supervisors introduced the incentive scheme in randomly selected classes. In each city zone, classes were first stratified by neighborhood and then randomized so that classes with and without the incentive scheme were in different neighborhoods. This was to prevent students in control classes from hearing about the incentive scheme. The scheme promised a reward to all students in the class who attended more than 85% of school days during the 39-day period between December 14th and January 31st.[13] To inform students about the scheme, the supervisors put up on the wall a chart with each student's name and each school date during the incentive period. Next, following a script that the research team had prepared, they told the students that when they skipped school, it became harder for them to understand the material that was taught, and this also affected their ability to learn subsequent material. The school had decided that any student who attended school regularly would receive a reward. Their attendance would be marked on the chart every day during the specified period. At the end of this period, all students who had attended more than 33 days would be eligible for a reward. The students were then shown samples of the reward (each reward was two pencils and a brightly colored eraser shaped like an animal), and were told that

---

[12]We made it clear that this child found the sum difficult, so as to prevent the student from assuming that their hypothetical child was bright and so would not find the sum difficult.

[13]This implies that both sample and non-sample students in a class were exposed to the same treatment condition.

the supervisor would give them one of these as a reward.[14] On each day during the reward period, the teacher was asked to fill in the chart, but not to mention it directly to any student. In the classes that were assigned to the control group, the supervisors gave each teacher a similar chart to fill in every day. The chart was not made public, and the supervisor did not make any announcements in class.[15]

The Time 3 visits took place during the incentive period, thus allowing us to examine how students responded to the scheme while it was in place. At the end of the incentive period, our project coordinator collected all the charts and identified the students who had met the threshold, all of whom received their rewards from the supervisors at small ceremonies in the classroom. All rewards were distributed within two weeks of the end of the incentive period. No further announcements about attendance were made.

Two further visits took place at Time 4 and Time 5, roughly one month and two months after the incentive period ended. Finally, in March, all grade 3 students took a test in mathematics and science, administered by Educational Initiatives (EI).[16] Their tests were aimed at uncovering student ability, and so did not directly test the material covered in the classroom. Questions were designed to test a variety of types of knowledge, ranging from fact and concept recognition to complex problem-solving and analysis skills. Thus rote learning was unlikely to guarantee a high test score. Note also that since Gyan Shala teachers strictly follow a daily lesson plan, they were unlikely to be able to teach to the test.

---

[14]Although these rewards had small monetary value, we had found in a pilot the previous year that they were appealing to the students.

[15] Thus our incentive scheme involved a speech by the supervisor that explained that regular attendance was important, promised a reward to students who met the attendance threshold, and publicly monitored each student's attendance. It can be argued that this represents a bundle of behavioral nudges, and we are unable to disentangle the pure effect of a reward scheme absent these other elements. It is also true, however, that to make the reward scheme salient to the students, the school would have to explain the rationale behind it. Also, to implement the scheme transparently, it would be necessary to ensure common knowledge between the student and the teacher/incentivizer about the student's attendance and eligibility for the reward.

[16]Educational Initatives provides an independent testing service. The scores on tests administered by EI have been used to evaluate student performance in previous research on education in India (Muralidharan and Sundararaman 2011).

All test questions were multiple choice. Students were given question papers, the exam administrator read an exam question aloud, asked students to circle the correct alternative, and then moved on to the next question. Test administrators unaffiliated with Gyan Shala then took these question papers and filled in an optimal mark recognition (OMR) sheet for the student. Due to a budget constraint, Gyan Shala opted to have a random subsample of exam scripts graded. These were then processed, and the test scores were delivered both to EI and to Gyan Shala. EI then prepared a summary report of the students' performance in each class. This report also classifies each question in the test according to the type of knowledge it was testing. Using this information, we classify the questions as "simple," "intermediate" and "complex" and analyze not just the total scores in the math and science tests, but also the scores in each category.

We have test score data for 584 students. These 584 students are not a perfect subset of our sample of 799 students described above. We observe test scores for only 308 of the 799 sample students. For 276 students we have test score data, but since they were not in our sample, we do not have attendance and interview data.[17] In linear probability regressions, neither assignment to treatment nor baseline attendance at Time 0 predict the probability that we observe a test score for a student. When we evaluate the effect of the intervention on only the 308 students for whom we have all data, our results are qualitatively unchanged.

4   Empirical Specification and Results

4.1   Attendance

In this section we examine the effect of the incentive scheme on attendance. We examine separately the effect when the incentive was in place and after it had been removed.

We start by depicting the key patterns as seen in the raw means from the data. Next we run regressions with additional controls and student fixed effects.

---

[17]However, since we know which classes they belonged to, we know whether they were in the treatment or control condition.

As noted above, we have 799 students in the sample. For each of these students we have data on whether they were present in school at six different points in time (Time 0 through Time 5).

As column 2 in Table 2 shows, average attendance rates vary from a low of 72 percent to a high of 86 percent over the 6 visits. Columns 5-7 and columns 8-10 show how the attendance rates varied between the control classes and the incentive classes, and how the subsequent attendance rates differed between baseline (Time 0) attenders and non-attenders. In each of these two subgroups of students, in the control classes, attendance dipped from Time 1 to Time 2 and then increased at Time 3.[18] Recall that the intervention was the promise of a reward for attending 85 percent or more of school days during a 39-day period in December-January. Since the Time 3 visit took place during this 39-day period, the difference in attendance between the incentive and control classes at Time 3 reflects the effect of the incentive on attendance. At Time 3, 90 percent of incentive class students were present, compared to 81.5 percent of control class students.

Time 4 and Time 5 visits occurred after the reward period had ended, and therefore allow us to see if the incentive had a persistent effect even after it had been discontinued. As we can see, both at Time 4 and Time 5, attendance was lower than at Time 3 for all students. However at Time 4, incentive students still remained more likely to attend than control students. This effect wore off with time, so that in both subgroups (high and low baseline attenders), incentive students were less likely to be present at Time 5 than control students.

In Table 3 we run linear probability regressions according to the specification below.

$$
\begin{aligned}
y_{ict} = \ & \alpha_i + \beta_1 \text{Time } 3_t + \beta_2 \text{Time } 4_t + \beta_3 \text{Time } 5_t + \beta_4 (\text{Reward}_c \times \text{Time } 3_t) \\
+ \ & \beta_5 (\text{Reward}_c \times \text{Time } 4_t) + \beta_6 (\text{Reward}_c \times \text{Time } 5_t) + \epsilon_{ict} \quad (1)
\end{aligned}
$$

---

[18]It is common for parents to take their families back to their hometown during the Diwali holidays (that were just before Time 2) and not return in time for school reopening. We verify that the explanation "student is out of town" was much more common for absence at Time 2 than at the other visits (63% v. 44%, p-value = 0.013).

Here y_{ict} is a binary variable that takes value 1 if student i in class c was present at the investigator visit at time t, and is zero otherwise. The α_i represent student fixed effects that capture all time-invariant personal and location-specific characteristics that may influence attendance. Thus if there there are fixed personal characteristics correlated with low socioeconomic status preventing a student from attending school regularly, these do not affect our results. The inclusion of student fixed effects also helps to address the potential concern that student-specific fixed differences affected the response to the intervention.[19] The Time 0 observations are removed from the sample because as we shall see below they are used to classify students by baseline attendance levels. Standard errors are clustered at the class level to control for intra-class correlation in attendance.[20]

The coefficient β_1 captures the Time 3 effect on the attendance rate of control students. The coefficient β_4 indicates how different this Time 3 effect was for incentivized students, and thus tells us the effect the incentive had while it was in place. As we see in Column 1, while the reward scheme was on, it increased the likelihood that the average student attended school. At the Time 3 visit, the probability that the investigators found a sample student present in the control classes was the same as before ($\hat{\beta}_1$ = 0.035, s.e. = 0.026, not statistically significant), but in the reward classes, the likelihood was 10.9 percentage points higher ($\hat{\beta}_1 + \hat{\beta}_4$ = 0.109, s.e. = 0.022, p-value = 0.000). The coefficient $\hat{\beta}_4$ can thus be interpreted to imply that at Time 3, the incentive increased the average student's attendance by a statistically significant 7.4 percentage points (or 9.9 percent). Since the reward scheme lasted 39 days, this translates to an additional 3.9 days of attendance by the average child.

Thus we find that while the incentive was in place, it caused attendance to increase. This is in line with expectations: if the incentive is attractive, it can increase student effort (Gneezy, Meier, and Rey-Biel 2011). However if the incen-

---

[19]Student fixed effects also absorb the dummy for Reward$_c$. However in unreported robustness checks we verify that our results are qualitatively unchanged even if we do not control for student fixed effects, and include an explicit control for being in the reward condition instead. The results are also qualitatively unchanged if we include a Time 2 dummy and its interaction with the Reward condition.

[20]These results are robust to fixed-effects logit specifications instead of linear probability models.

tive reduced intrinsic motivation, then after it was discontinued, student motivation should have become even lower: not only would students no longer have the extrinsic motivation to attend, they would also have lower intrinsic motivation. This could make the incentivized students even less likely to attend than the control students. Accordingly, we examine the effect of the reward 1 month after (Time 4) and 2 months after (Time 5) the reward period ended. We find once again that at Time 4 and at Time 5, control students were no more likely to attend school than before ($\hat{\beta}_2$ = -0.018, s.e. = -0.024, p-value = 0.464, and $\hat{\beta}_3$ = -0.028, s.e. = -0.025, p-value = 0.279). The incentive did not change this non-effect either ($\hat{\beta}_4$ = 0.014, s.e. = 0.039, p-value = 0.720; $\hat{\beta}_5$ = -0.034, s.e. = 0.039, p-value = 0.389), suggesting that the positive effect of the incentive scheme did not persist after the incentive was removed.

However, as discussed earlier, there is reason to believe that the incentive might have had different long-term effects on students with low and high baseline motivation to attend school. Accordingly, in columns 2 and 3 we divide the sample into two subgroups, using as a proxy for baseline motivation their attendance during the Time 0 unannounced visit. In column 2, we focus on baseline attenders, and find that the incentive increased their likelihood of attending school by a statistically significant 6.7 percentage points. After the incentive was removed, their attendance rate was no different from the control group (at either Time 4 or Time 5). This is consistent with either no reduction in their intrinsic motivation, or a very small reduction that did not change their attendance.

In column 3 we focus on baseline non-attenders (absent at Time 0). Although the magnitude of the incentive effect is large at 9.3 percentage points, it is imprecisely estimated. Strikingly however, at Time 5, these students were 13.9 percentage points less likely to attend school than similar baseline non-attenders in control classes. If this decline in attendance was uniform over the last two months of school, it suggests the average incentivized baseline non-attender attended 7.8 fewer days after the reward period ended. Thus, in contrast to the previous literature, we do find a negative long-term effect of the incentive, but only among students who had low baseline attendance. The incentive lowered their attendance rate in the post-incentive period below what it would have been if no incentive had been offered.

4.2 Test scores

We see the same pattern in student performance. In Table 4, we run regressions with the specification

$$y_{ic} = \beta_0 + \beta_1 \text{Reward}_c + \beta_2 X_{ic} + \epsilon_{ic} \tag{2}$$

where the dependent variable is student i's standardized score on the Educational Initiatives test administered at the end of the school year.[21] Controls include the student's z-score on the final exam (administered by Gyan Shala) in the previous year, the student's gender, the city zone where the class is located, and a dummy variable for the psychological intervention that was conducted in an orthogonal design to the reward intervention. Standard errors are clustered at the class level.

As column 1 in Table 4 shows, although the average treatment effect on the aggregate test score is negative, it is not statistically different from zero. This is also true when we analyze the mathematics (column 4) and science (column 7) scores separately.

However, as we see in columns 2, 5 and 8, this null effect was driven by baseline attenders (present at Time 0). As we noted above, the incentive increased these students' attendance during the incentive period, but had no effect on it afterwards. It is perhaps not surprising that the very small increase in days attended had no direct mechanical effect on their test scores. However the incentive also does not appear to have increased test scores through other means, such as for example, by increasing students' interest in school.

In column 3, we see the opposite result: the reward lowered test performance of baseline non-attenders. Their average score was 0.59 standard deviations lower than their counterparts in the control classes. We see a similar effect on the mathematics score (-0.48σ, column 6) and the science score (-0.59σ, column 9). Thus after the incentive was removed, these students both attended school less, and performed worse than if they had not faced the incentive.

---

[21]The score is standardized with respect to the mean score across all students in the 68 classes in the sample.

4.3 Possible Mechanisms

4.3.1 Lower Scores on Difficult Questions

In order to further understand the correlates of the decreased test scores, we examine separately the students' scores on questions of different difficulty levels.[22] As Table 5 shows, baseline non-attenders' scores on simple questions were unaffected by the incentive (columns 4 and 10). Scores were lower on the more difficult questions: intermediate and complex questions in mathematics (column 5, $\hat{\beta}_1$ = -0.475, s.e. = 0.210, p = 0.028 and column 6, $\hat{\beta}_1$ = -0.567, s.e. = 0.231, p = 0.017) and intermediate questions in science (column 11, $\hat{\beta}_1$ = -0.715, s.e. = 0.276, p = 0.012). (The coefficient for complex mathematics questions in column 3 is negative, but not significantly different from zero.) Thus the incentive appears to have lowered these students' ability or willingness to answer difficult test questions. The incentive did not have a significant effect on any of the test scores for baseline attenders.

4.3.2 Lower Liking for School Subjects

After the incentive was removed, baseline non-attenders in the incentive classes rated their liking for school subjects lower than they would have had they not been incentivized. This is apparent in Table 6, where we use data from the student interviews at Times 3 and 5 to run student fixed-effects regressions according to the specification:

$$y_{ict} = \alpha_i + \beta_1 \text{Time } 5_t + \beta_2 (\text{Reward}_c \times \text{Time } 5_t) + \epsilon_{ict} \tag{3}$$

The dependent variable is student i's rating at time t of mathematics, or of science (on a 7-point scale). Student fixed effects control for time-invariant observable and unobservable characteristics of the students. The coefficient β_2 estimates whether ratings by students in the incentive and control classes changed differentially after the incentive was discontinued. As we see in column 1, the average student in control classes rated mathematics 0.15 points higher (on a

---

[22]For a list of the knowledge categories that were tested and our classification of test questions into the "simple," "intermediate" and "complex" categories, see Table 8.

mean of 2.46) at Time 5 than at Time 3. On average, the reward had no differential effect ($\hat{\beta}_2 = 0.007$, s.e. = 0.110, p = 0.949). However, when we split the sample by students' baseline attendance, we see in column 3 that among baseline non-attenders, the coefficient $\hat{\beta}_2$ is negative, although imprecisely estimated. This is suggestive evidence that among these students, the increase in ratings was smaller than among the non-incentivized students.

A potential concern with column 3 is that since the investigators conducted the interviews when they visited the classrooms, students who were absent at the time of the visit were less likely to be interviewed. If, as we have shown above, the incentive lowered attendance at the Time 5 visit, then in column 3 we might be disproportionately estimating the effect of the incentive not on representative baseline non-attenders, but on those who chose to attend at Time 5, perhaps because they enjoyed school. To avoid this sample selection bias, at each interview visit (Times 1, 3 and 5), our investigators were required to make up to three attempts to find these students and interview them. This involved asking around to find out where and when the student would be available, and making follow-up visits accordingly. Note that since the Gyan Shala classes are in the same neighborhoods as the students' homes, it is relatively easy to locate homes and interview the students there if they are available. As a result, 84.5% of students who were absent on the day of the Time 5 visit, were nevertheless interviewed within a few weeks of the Time 5 visit. Although this is lower than the 95% interview rate of those who were present in school when the visit took place, it gives us a large enough sample to measure these children's liking for school subjects.

Therefore, in column 4 we restrict the subsample to baseline non-attenders who were also absent at the Time 5 visit. If repeated absence is indicative of disinterest, then both the incentive and the control students in this subsample should have low ratings for school subjects. Within this sample we find that although for control students the Time 5 ratings were higher than the Time 3 ratings, for incentive students they were actually lower ($\hat{\beta}_1 + \hat{\beta}_2$ = 0.650-0.923 = -0.273).[23]

---

[23]It could be suggested that the incentive actually increased ratings for these students at Time 3, and so the subsequent decline represents a reversion to baseline levels. However when we run these regressions without student fixed effects we do not find that ratings were higher for

When we consider the effect on students' rating for science on this subsample in column 8, the sign on $\hat{\beta}_2$ is negative but not statistically different from zero. We conclude that the incentive reduced baseline non-attenders' enjoyment of mathematics. This is consistent with the insight from psychology that intrinsic motivation is a key determinant of liking: as a student's intrinsic motivation to study a particular subject dwindles, they correspondingly like that subject less.

4.3.3 Lower Optimism about Ability to Perform and Learn

Finally, in Table 7 we analyze two other interview questions that measure students' opinion about their performance at challenges, and their ability to learn. Students were told about a hypothetical student attempting a challenging sum and asked to predict how he or she would perform on a scale of 1 to 5. As we have seen in Table 1, at baseline, the average student predicted the child would receive 2.2 stars from the teacher, and there was no significant difference between control and incentive classes. However as column 3 shows, the incentive caused baseline non-attenders to predict that the child would receive 0.3 fewer stars. This negative effect becomes even larger when we restrict the sample to students who were absent at both Time 0 and Time 5 (column 4, $\hat{\beta}_1$ = -0.469, s.e. = 0.174, p = 0.011).

We also tried to elicit students' confidence about their ability to learn something new. Since teachers had introduced students to crossword puzzles, we asked them if they thought they could learn to solve one.[24] Once again, among baseline non-attenders, the incentive lowered the belief they could learn this new skill (columns 7 and 8, although the coefficient in column 8 is imprecisely estimated). Thus, the reduction in attendance and test scores caused by the intervention appears to be correlated with lower self-reported enjoyment of school subjects, less optimism about ability to perform a challenging task, and less optimism about learning a new skill.

---

incentivized students than for control students at Time 3.

[24]Crossword puzzles were part of a worksheet exercise that students saw a few weeks before the Time 1 interviews. We asked students this question at all three interviews.

5 Conclusion

We have identified two issues that have received relatively little attention in the experimental incentive literature in education. First, even if incentives have positive effects on motivation while they are in place, they might have negative effects after they are removed. This makes it important to examine not just their impacts not just in the immediate term but also in the longer term. Second, if incentives lower intrinsic motivation, they might have more substantial behavioral impacts on students who had low motivation to start with. This could happen if a decrease in motivation lowers student effort and outcomes by more among students who were less motivated to start with.

In our study, students with high baseline attendance (and presumably high baseline motivation) were influenced positively by the incentive while it was in place, but were unaffected by it after it had been discontinued. This could be interpreted to mean that the incentive did not create a "habit" for these students to attend school more than their non-incentivized peers. However students with low baseline attendance were negatively affected. Not only did the incentive lower their attendance in the post-incentive period, it also lowered their test performance three months after the incentive scheme ended. In the long run they also enjoyed the material taught in school less, and were less optimistic and less confident about their ability to perform and learn.

In any incentive scheme, it is likely that some students will fail to earn the reward because they do not meet the target. When an attendance target is absolute (as it was in our case), students with high attendance levels meet it more easily, and the losers are disproportionately those with low attendance levels to start with. This paper shows that the incentive scheme can have unintended negative consequences for this very set of students, which is the group that incentive schemes typically intend to help.

A few caveats are in order. First, it could be argued that if students were unable to attend school due to circumstances beyond their control, then the reward scheme might have imposed an extremely challenging standard that made their constraints more salient and discouraged them further. We took care to choose a reasonable attendance target. As Table 2 shows, the average control student at-

tended 78 percent of school days during the incentive period, so that 85 percent represented only a 9% increase. According to school administrators, much of the absence could be explained by students' choices not to attend school rather than systemic problems at home or elsewhere.[25]

It is certainly possible that some students in the sample were discouraged by failing to meet the attendance target, and that since baseline non-attenders were more likely to miss the target, this discouragement effect was disproportionately strong among them. Since we have daily data from the incentive period for all classes in both the treatment and control groups, in unreported results we examine separately baseline non-attenders who met the incentive target of 85% of school days, and who did not. Among those who met the target, longer-term attendance (as measured by the Time 5 visit) did not decline significantly. Among those who failed to meet the target, the incentive lowered the attendance rate by 16.8 percentage points (p = 0.059). It is possible that the incentive scheme made these students' poor attendance salient to them and thereby de-motivated them even further. This underscores a central message of this paper, that rewards can have negative consequences on the students that educators intend to help the most.

Second, the attendance target could have been designed to be relative, so that students were rewarded for increasing their attendance by a certain proportion above their own baseline. Then students with low baseline attendance could have earned rewards with relatively small absolute increases in attendance and would have been less likely to be discouraged. This would have required catering the target to each student individually, and since student attendance varies within each classroom, would have required within-class variation in attendance targets.[26] Not only would this have been difficult to administer, it would have been difficult to ensure that each student understood what their

---

[25]Note also that the reward period was deliberately chosen during a period when there are no long-drawn festivals that often cause students to miss school. However it is true that we are unable to definitively rule out the possibility that students' absence was caused by circumstances outside their control.

[26]However this might have discouraged students with high baseline attendance, since some of them might have missed their own target even if they increased absolute attendance by more than their low baseline peers did.

own target was.[27] Although pedagogical best practices prescribe that each student be set an achievement target that is appropriate for them individually, it is rare, especially in developing country contexts where teaching resources are scarce, that different standards of achievement are applied to different students. Thus our experiment tests an incentive scheme that closely approximates one that might be implemented in such a setting. It cautions educators and policymakers that such a scheme could end up hurting students whose effort and motivation need the greatest boost, without generating significant benefits for those who are already performing at a high level.

Acknowledgements

---

[27]In Bettinger (2012)'s study, for eighth and ninth graders the eligibility to receive cash rewards was randomized at the student level. However, once they were selected into the incentive group, all students were assigned the same target. In Berry (2014)'s experiment, all students were offered rewards of the same value for meeting the same targets, but the type of reward was randomized at the student level.

References

Angrist, Joshua, and Victor Lavy. 2009. "The Effects of High Stake High School Achievement Awards: Evidence from a Randomized Trial." *American Economic Review* 99 (4): 1384–1414.

Berry, James. 2014. "Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India." Mimeograph.

Bettinger, Eric P. 2012. "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores." *The Review of Economics and Statistics* 94 (3): 686–698.

California Department of Education. 2015. "Truancy." Technical Report, Department of Education. Accessed on 19th November 2015.

CfBT Education Services. 2010. "The Gyan Shala Programme: An Assessment." Technical Report.

Chao, Melody M., Sujata Visaria, Anirban Mukhopadhyay, and Rajeev Dehejia. 2016. "When Mindsets and Identity Diverge: The Effects of Implicit Theories and Incentive Schemes in a Field Intervention." Mimeograph.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica* 77 (3): 909–931.

Deci, Edward L., Richard Koestner, and Richard M. Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." *Psychological Bulletin* 25 (6): 627–668.

Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-determination in Human Behavior*. New York, NY: Plenum Press.

Educational Consultants India Limited. 2007. "Study of Students Attendance in Primary & Upper Primary Schools: Abridged Report." Technical Report.

Educational Initiatives Private Limited. 2010. "Test of Student Learning for Gyanshala: Assessment Report." Technical Report.

Fryer, Roland. 2011. "Financial Incentives and Student Achievement: Evidence from Randomized Trials." *Quarterly Journal of Economics* 126:1755–1798.

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25 (4): 191–210.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *The Review of Economics and Statistics* 91 (3): 437456.

Levitt, Steven D., John A. List, Susanne Neckerman, and Sally Sadoff. 2012. "The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance." Technical Report. NBER Working Paper 18165.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy* 119 (1): 3977.

Paredes, Ricardo D., and Gabriel A. Ugarte. 2011. "Should Students Be Allowed to Miss?" *The Journal of Educational Research* 104:194201.

Roby, Douglas E. 2004. "Research on School Attendance and Student Achievement: A Study of Ohio Schools." *Educational Research Quarterly* 28 (1): 314.

Rodriguez-Planas, Nùria. 2012. "Longer-term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the United States." *American Economic Journal: Applied Economics* 4 (4): 121–139.

World Health Organization. 2007. "Growth Reference data for 5-19 years." Technical Report. Accessed on 21st March 2015.
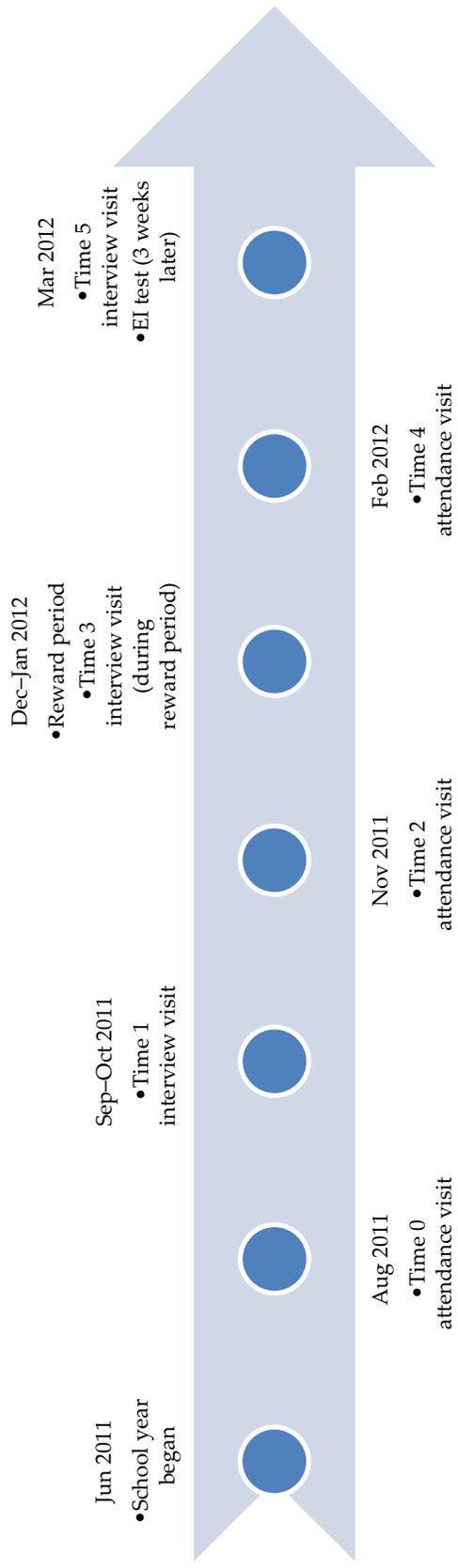
Figure 1: Sequence of Events

<p align="center">Table 1: Sample Characteristics</p>

	N (1)	All (2)	No reward (3)	Reward (4)	T-test of differences (5)
Student characteristics					
Female	799	0.51 (0.02)	0.49 (0.03)	0.54 (0.03)	*0.257*
Year of birth	769	2002.8 (0.06)	2002.8 (0.08)	2002.8 (0.08)	*0.785*
Body Mass Index (kg/m2)	768	13.83 (0.11)	13.85 (0.15)	13.81 (0.16)	*0.842*
Household assets					
Mobile phone	768	0.93 (0.01)	0.92 (0.01)	0.93 (0.01)	*0.810*
VCR/DVD	768	0.36 (0.03)	0.37 (0.04)	0.35 (0.04)	*0.791*
Computer	768	0.01 (0.00)	0.01 (0.01)	0.01 (0.01)	*0.659*
Autorickshaw/motorbike/car	799	0.24 (0.02)	0.22 (0.03)	0.26 (0.03)	*0.268*
Toilet in the house	768	0.73 (0.03)	0.69 (0.04)	0.78 (0.05)	*0.148*
School-related variables					
Present at Time 0	799	0.75 (0.02)	0.74 (0.03)	0.75 (0.03)	*0.817*
Administrative attendance record	797	0.78 (0.01)	0.79 (0.01)	0.78 (0.01)	*0.585*
z-score on previous year's exam	783	0.00 (0.06)	0.02 (0.08)	-0.03 (0.09)	*0.687*
Likes Math (range = [-3, 3])	621	2.46 (0.06)	2.51 (0.08)	2.41 (0.08)	*0.367*
Likes Science (range = [-3, 3])	621	1.99 (0.08)	2.09 (0.10)	1.87 (0.11)	*0.158*
Score on a difficult sum (range = [1, 5])	768	2.24 (0.09)	2.30 (0.12)	2.17 (0.13)	*0.481*
Able to solve a crossword puzzle (range = 0,1)	759	0.96 (0.01)	0.96 (0.01)	0.96 (0.01)	*0.723*

Means are computed from the baseline student survey data. t-tests account for correlation at the class level. Standard errors are in parentheses. Column 4 reports p-values for t-tests of differences between columns (3) and (4).

Table 2: Attendance Rates

| | N | All students | | | Present at Time 0 | | | Absent at Time 0 | | |
		Both conditions	No reward	Reward	Both conditions	No reward	Reward	Both conditions	No reward	Reward
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Random visits					
Time 0	799	74.59	74.10	75.13	100.0	100.00	100.00	0.00	0.00	0.00
Time 1	799	85.36	85.90	84.82	87.58	89.64	85.37	78.82	75.00	83.16
Time 2	799	71.84	70.26	73.56	74.66	74.76	74.56	63.55	57.41	70.53
Time 3	799	85.61	81.53	90.05	88.26	86.08	90.59	77.83	68.52	88.42
Time 4	799	77.47	76.26	78.80	80.37	80.91	79.79	68.97	62.96	75.79
Time 5	799	74.22	75.30	73.04	76.68	77.67	75.61	67.00	68.52	65.26
					Reward period					
Average attendance	798	80.38	78.49	82.45	83.54	82.24	84.94	71.13	67.78	74.93
Above the 85% threshold	798	52.51	48.32	57.07	57.82	54.87	60.98	36.95	29.63	45.26

The top panel shows mean attendance rates of sample students, as recorded during each of the unannounced random visits. The bottom panel shows the average percentage of school days that students attended, and the mean number of students who attended at least 85% of school days during the 39-day reward period.

26

Table 3: Effect of reward scheme on attendance at unannounced visits

	All students (1)	Present at Time 0 (2)	Absent at Time 0 (3)
Time 3	0.035	0.039	0.023
	(0.026)	(0.027)	(0.046)
Time 4	-0.018	-0.013	-0.032
	(0.024)	(0.024)	(0.055)
Time 5	-0.028	-0.045	0.023
	(0.025)	(0.029)	(0.049)
Reward × Time 3	0.074**	0.067*	0.093
	(0.034)	(0.034)	(0.069)
Reward × Time 4	0.014	0.011	0.022
	(0.039)	(0.039)	(0.078)
Reward × Time 5	-0.034	0.002	-0.139**
	(0.039)	(0.045)	(0.068)
Sample mean	0.789	0.815	0.712
	(0.006)	(0.007)	(0.142)
Observations	*3,995*	*2,980*	*1,015*
R-squared	*0.015*	*0.015*	*0.022*
Number of students	*799*	*596*	*203*

All columns report student fixed-effects linear probability regressions, where the dependent variable takes value 1 if the student was present at the unannounced visit, and 0 otherwise. Standard errors in parentheses are clustered at the class level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Effect of reward scheme on test scores

	Aggregate			Mathematics			Science		
	All students (1)	Present at Time 0 (2)	Absent at Time 0 (3)	All students (4)	Present at Time 0 (5)	Absent at Time 0 (6)	All students (7)	Present at Time 0 (8)	Absent at Time 0 (9)
Reward	-0.062 (0.202)	0.055 (0.216)	-0.586** (0.235)	-0.055 (0.207)	0.036 (0.233)	-0.483** (0.202)	-0.052 (0.182)	0.069 (0.179)	-0.594** (0.278)
Sample mean	0.450 (0.043)	0.070 (0.051)	-0.014 (0.083)	0.050 (0.043)	0.083 (0.052)	-0.030 (0.078)	0.030 (0.042)	0.044 (0.050)	0.000 (0.084)
Observations	584	419	152	584	419	152	583	418	152
R-squared	0.076	0.101	0.151	0.059	0.070	0.126	0.077	0.107	0.141

All columns report OLS regressions. The dependent variable is the student's z-score on the test administered by Educational Initiatives. We control for the student's z-score on the previous year's final exam. A female dummy, zone dummies and a dummy for the orthogonal psychological intervention are included. Standard errors in parentheses are clustered at the class level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Effect of reward scheme on test scores, broken by difficulty level

| | Mathematics | | | | | | Science | | | | | |
| | Present at Time 0 | | | Absent at Time 0 | | | Present at Time 0 | | | Absent at Time 0 | | |
	Simple (1)	Intmdt (2)	Complex (3)	Simple (4)	Intmdt (5)	Complex (6)	Simple (7)	Intmdt (8)	Complex (9)	Simple (10)	Intmdt (11)	Complex (12)
Reward	0.142	0.046	-0.141	-0.075	-0.475**	-0.567**	0.072	-0.033	0.186	-0.234	-0.715**	-0.308
	(0.183)	(0.225)	(0.220)	(0.197)	(0.210)	(0.231)	(0.190)	(0.168)	(0.145)	(0.201)	(0.276)	(0.269)
Sample mean	0.056	0.066	0.990	-0.520	-0.029	0.015	0.049	0.001	0.089	0.030	-0.040	0.056
	(0.051)	(0.051)	(0.052)	(0.083)	(0.079)	(0.078)	(0.051)	(0.050)	(0.049)	(0.073)	(0.085)	(0.085)
Observations	*419*	*419*	*419*	*152*	*152*	*152*	*419*	*418*	*419*	*152*	*152*	*152*
R-squared	*0.059*	*0.058*	*0.066*	*0.089*	*0.111*	*0.155*	*0.057*	*0.111*	*0.074*	*0.147*	*0.119*	*0.108*

All columns report OLS regressions. The dependent variable is the student's *z*-score on the test administered by Educational Initiatives. We control for the student's *z*-score on the previous year's final exam. A female dummy, zone dummies and a dummy for the orthogonal psychological intervention are included. Standard errors in parentheses are clustered at the class level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6: Effect of reward scheme on change in students' liking for maths and science

	Mathematics				Science			
	All students	Present at Time 0	Absent at Time 0	Absent at both Time 0 & Time 5	All students	Present at Time 0	Absent at Time 0	Absent at both Time 0 & Time 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Time 5	0.149**	0.131*	0.159	0.650**	0.120	0.100	0.098	0.300*
	(0.060)	(0.071)	(0.133)	(0.292)	(0.103)	(0.109)	(0.185)	(0.152)
Reward × Time 5	0.007	0.053	-0.159	-0.923**	0.178	0.180	0.204	-0.073
	(0.110)	(0.115)	(0.209)	(0.372)	(0.144)	(0.156)	(0.237)	(0.353)
Sample mean	2.661	2.705	2.57	2.451	2.293	2.288	2.301	2.451
	(0.024)	(0.026)	(0.058)	(0.123)	(0.032)	(0.037)	(0.066)	(0.115)
Observations	*1,437*	*1,068*	*349*	*102*	*1,437*	*1,068*	*349*	*102*
Number of students	*785*	*581*	*194*	*60*	*785*	*581*	*194*	*60*
R-squared	*0.017*	*0.021*	*0.009*	*0.125*	*0.019*	*0.016*	*0.022*	*0.035*

All columns report student fixed-effects regressions, where the dependent variable is the student's rating of her liking for the subject, in interviews at Time 3 and Time 5. A female dummy, zone dummies and an indicator for the orthogonal psychological intervention are included. Standard errors in parentheses are clustered at the class level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Effect of reward scheme on student optimism and confidence

	All students	Present at Time 0	Absent at Time 0	Absent at both Time 0 & Time 5	All students	Present at Time 0	Absent at Time 0	Absent at both Time 0 & Time 5
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Performance on a difficult sum				Ability to solve a crossword puzzle			
Reward	-0.081	-0.015	-0.309**	-0.469**	-0.012*	-0.007	-0.029*	-0.024
	(0.095)	(0.109)	(0.144)	(0.174)	(0.007)	(0.005)	(0.014)	(0.025)
Sample mean	1.601	1.585	1.634	1.564	0.994	0.997	0.984	0.981
	(0.034)	(0.039)	(0.068)	(0.106)	(0.003)	(0.002)	(0.009)	(0.018)
Observations	*777*	*576*	*191*	*55*	*776*	*575*	*191*	*55*
R-squared	*0.075*	*0.062*	*0.142*	*0.237*	*0.029*	*0.014*	*0.094*	*0.129*

All columns report OLS regressions using student interview data from Time 3. In columns 1, 2 and 3 the dependent variable is the number of stars (ranging from 1 to 5) the student expects a child will receive for a difficult maths sum. In columns 4, 5 and 6 the dependent variable indicates whether the student expects he/she can learn how to solve a crossword puzzle. A female dummy, zone dummies and a dummy for the orthogonal psychological intervention are included. Standard errors in parentheses are clustered at the class level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Classification of Educational Initiatives test questions by difficulty level

Difficulty	Knowledge Tested	Number of questions
	Mathematics	*30*
Simple	Number sense, related concepts and basic number competency	5
	Arithmetic operations: Addition & Subtraction	3
	Arithmetic operations: Multiplication	3
Intermediate	Word problems & visual based problems	6
	Basic shapes & geometry	4
	Applications in daily life: money, time, calendar, length, etc.	5
Complex	Problem solving (advanced or challenging problems)	4
	Science	*30*
Simple	Recollection or recognition of science facts & concepts	3
	Definition or description of scientific terms, organisms or materials	4
	Knowledge of use of scientific instruments, tools and procedures	3
Intermediate	Classification/comparison of organisms/processes: giving examples	5
	Representing or explaining processes or observed phenomena	5
	Extraction, translation and application of knowledge or information	3
Complex	Complex analysis, data interpretation, integrating different concepts	4
	Hypothesis formulation or prediction of outcome	3

The tests that Educational Initiatives administered had 30 questions each for mathematics and science. Each question was meant to test particular learning objectives. We classify these learning objectives and the corresponding questions into three categories: "simple," "intermediate" and "complex".