

HARNESSING EDUCATION FOR CLIMATE ACTION – A RANDOMIZED EXPERIMENT IN BANGLADESH SECONDARY SCHOOLS

Pre-Analysis Plan

Teresa Molina, Shwetlena Sabarwal

September 25, 2023

1. Introduction

This study will explore the effectiveness of delivering climate change education in schools, using a randomized controlled trial (RCT) in Bangladesh. We will examine the impact of two information interventions on students' knowledge about and attitudes towards climate change. In addition, we hope to investigate whether enhanced climate change awareness leads to greater interest in and pursuit of STEM fields, particularly for girls. This plan outlines the study hypotheses and data analysis to be conducted.

2. Overview of the Study

2.1 Motivation and Intervention

Bangladesh, like many other countries, faces two pressing challenges—climate change and gender inequality. In recent decades, climate change has risen as a primary global issue, impacting ecosystems, economies, and societies worldwide. As a nation in a region susceptible to rising sea levels, extreme weather events, and changing rainfall patterns, Bangladesh is particularly vulnerable to the impacts of climate change.

Simultaneously, Bangladesh faces the persistence of gender inequality across various domains, including education, employment, and political representation. In the education and employment space, one important manifestation of this inequality is the underrepresentation of women in Science, Technology, Engineering, and Mathematics (STEM) fields.

The education sector could serve as a promising area for interventions that address both of these issues in tandem. An effective intervention that enhances students' knowledge about the climate change issue and fosters a drive to seek solutions could potentially motivate them to pursue STEM, a field with perhaps the greatest potential to contribute solutions to the climate crisis. Importantly, such an intervention could have a disproportionate effect on girls, whose education and career choices are often more socially motivated, and could therefore contribute to closing the gender gap.

2.2 Intervention and Timeline

In this study, we aim to test two information interventions, one light-touch intervention at baseline and one more intensive intervention at midline.

- Light-touch baseline intervention (March 2023): The first intervention involves showing students and teachers a video and handing out an informative pamphlet on climate change.
- Intensive midline intervention (September-October 2023): At midline, we will distribute a climate change information booklet to students. The booklet will provide information about the climate crisis, along with an emphasis on how science can be used to find creative solutions or mitigation strategies. In order to provide incentives for students to read and absorb the material, we will announce a test on the material to be conducted about 2 months later. Top performers on this test will receive a reward.

In addition to surveys conducted at baseline and midline, we aim for two additional survey rounds

(funding and schedule permitting).

- In-person endline (November 2023): This is when we will administer the test on the material from the climate change booklet. We will also conduct an in-person survey.
- Phone survey follow-up (January 2023) This follow-up survey will be conducted over the phone to check whether the (former 8th grade) students have chosen the science track in 9th grade.

2.3 Sample Selection and Treatment Assignment

Bangladesh is divided into 8 divisions and 64 districts. The evaluation sample, surveyed at baseline, consists of over 5,600 students and 560 teachers in 140 schools in two districts in Bangladesh.

After randomly selecting the Rajshahi division, we categorized its districts based on climate change risk exposure as either high risk or low risk. This categorization relied on the average rainfall in each district compared to the median rainfall of the entire division. Specifically, districts with average rainfall above the division's median were labeled as high risk, primarily due to the area's susceptibility to intense monsoon rains and flooding. Districts with average rainfall below the median were labeled as low risk. We then randomly selected one district from each risk category. Finally, 70 schools within each selected district were randomly chosen and then assigned to one of four following groups:

Table 1: Treatment Groups

	Intensive Control	Intensive Treatment
Light-touch Control	17 schools (Pure Control Group) Receive no intervention	18 schools (Intensive Only Group) Receive intensive intervention only
Light-touch Treatment	17 schools (Light Only Group) Receive light-touch intervention only	18 schools (Both Group) Receive both interventions

The evaluation sample comprises all 8th grade students in the selected schools. All schools will be administered the baseline, midline, endline, and phone follow-up.

3. Data

As described above, there will be a total of four surveys conducted: baseline, midline, endline, and phone follow-up. Our main outcomes of interest, described in more detail below, fall into three categories: climate change knowledge, climate change attitudes (skepticism, anxiety, and willingness to take action), and interest/effort in STEM fields.

1. **Climate change knowledge:** share of correct answers to a set of climate change questions, standardized.
2. **Climate change attitudes:** indices constructed by standardizing responses to several questions and averaging across all standardized variables within the following sub-categories
 - a. *Skepticism*
 - b. *Anxiety*
 - c. *Willingness to take action*
3. **Interest/effort in STEM**
 - a. *STEM interest:* index constructed by standardizing responses to one or more (depending on the wave) questions related to interest in STEM, confidence in STEM ability, and plans to pursue STEM in the future – and averaging across all standardized variables.
 - b. *Intent to pursue science stream:* Indicator for whether the respondent plans to choose (or has chosen) the science stream in grade 9.
 - c. *Math aptitude:* share of correct answers to set of math questions, standardized.

For index variables, slightly different questions will be asked in each wave, though there will be a consistent set of questions asked in all waves. All questions used to create the indices will either have Likert scale, binary, or a 1-10 scale response options.

We will also collect information on student demographics (gender, age, parental education, household assets). At baseline, we will also survey four teachers per school to collect climate knowledge and attitudes.

Table 1 provides summary statistics for basic demographic controls, main outcome variables of interest, and teacher characteristics, by intensive treatment status (our main treatment of interest). Characteristics of intensive treatment and intensive control schools are largely similar, with one statistically significant exception: teachers were less likely to have an MA degree in intensive treatment schools.

Table 1: Summary Statistics, by “Intensive” Treatment Status

	(1) Treatment	(2) Control	(3) Difference
Age	13.72 (1.02)	13.77 (1.01)	-0.05 (0.06)
Female	0.57 (0.50)	0.65 (0.48)	-0.08 (0.05)
Mother Secondary School	0.49 (0.50)	0.49 (0.50)	0.00 (0.03)
Father Secondary School	0.43 (0.50)	0.43 (0.49)	0.00 (0.03)
Asset Index	0.00 (1.34)	0.00 (1.26)	0.00 (0.10)
Knowledge Index	-0.02 (1.02)	0.02 (0.98)	-0.03 (0.08)
Skepticism Index	-0.02 (0.64)	0.00 (0.63)	-0.02 (0.03)
Anxiety Index	-0.02 (0.65)	-0.06 (0.67)	0.04 (0.03)
Willingness to Contribute Index	0.02 (1.00)	-0.02 (0.99)	0.04 (0.06)
Interest in Science Index	-0.04 (1.01)	0.05 (0.98)	-0.09 (0.06)
Plans on Science Track	0.55 (0.50)	0.57 (0.50)	-0.01 (0.03)
Math Score	-0.05 (0.98)	0.04 (1.02)	-0.09 (0.11)
Teacher Age	44.32 (3.92)	43.22 (3.69)	1.10 (0.68)
Teacher Female	0.30 (0.24)	0.25 (0.21)	0.05 (0.04)
Teacher MA degree	0.47 (0.29)	0.65 (0.28)	-0.18*** (0.05)
Teacher Knowledge Index	0.05 (0.54)	0.02 (0.58)	0.03 (0.10)
Teacher Skepticism Index	-0.02 (0.36)	0.01 (0.38)	-0.03 (0.07)
Teacher Anxiety Index	-0.05 (0.45)	0.03 (0.51)	-0.08 (0.09)
Individuals	2713	2951	5664
Schools	72	68	140

Notes: In column 3, standard errors clustered at the school level are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 2 reports differences across the four separate treatment arms. Column 1 reports means for the pure control group, and the remaining columns represent differences relative to the pure control.

Table 2: Summary Statistics by Treatment Arm

	(1) Pure Control (Mean/SD)	(2) Light-Touch Only (Diff/SE)	(3) Intensive Only (Diff/SE)	(4) Both (Diff/SE)
Age	13.80 (1.03)	-0.06 (0.08)	-0.06 (0.08)	-0.11 (0.07)
Female	0.67 (0.47)	-0.04 (0.07)	-0.13* (0.07)	-0.08 (0.07)
Mother Secondary School	0.48 (0.50)	0.03 (0.04)	0.03 (0.04)	-0.00 (0.04)
Father Secondary School	0.40 (0.49)	0.06* (0.03)	0.04 (0.03)	0.02 (0.04)
Asset Index	-0.11 (1.25)	0.23** (0.11)	0.09 (0.14)	0.14 (0.11)
Knowledge Index	-0.06 (0.97)	0.15 (0.10)	0.00 (0.10)	0.08 (0.12)
Skepticism Index	0.01 (0.67)	-0.02 (0.05)	-0.01 (0.04)	-0.05 (0.05)
Anxiety Index	-0.05 (0.70)	-0.01 (0.05)	0.06 (0.04)	0.01 (0.05)
Willingness to Contribute Index	0.03 (1.06)	-0.10 (0.07)	-0.00 (0.07)	-0.00 (0.08)
Interest in Science Index	-0.02 (1.04)	0.13 (0.09)	-0.02 (0.09)	-0.04 (0.08)
Plans on Science Track	0.55 (0.50)	0.03 (0.05)	0.04 (0.04)	-0.04 (0.04)
Math Score	-0.16 (1.00)	0.42*** (0.15)	0.01 (0.16)	0.22 (0.14)
Teacher Age	43.66 (3.47)	-0.92 (0.93)	0.27 (0.90)	1.05 (0.95)
Teacher Female	0.30 (0.22)	-0.11** (0.05)	-0.01 (0.06)	-0.00 (0.06)
Teacher MA degree	0.68 (0.25)	-0.06 (0.07)	-0.14* (0.07)	-0.29*** (0.06)
Teacher Knowledge Index	-0.04 (0.61)	0.13 (0.15)	-0.04 (0.16)	0.23 (0.14)
Teacher Skepticism Index	-0.03 (0.37)	0.08 (0.10)	-0.08 (0.10)	0.10 (0.09)
Teacher Anxiety Index	0.02 (0.37)	0.01 (0.14)	-0.16 (0.10)	0.01 (0.11)
Individuals	1530	1421	1383	1330
Schools	34	34	36	36

Notes: In columns 2-4, standard errors clustered at the school level are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

4. Hypotheses

Overall Impacts

We predict our information interventions will increase knowledge, decrease skepticism, and increase willingness to contribute to fighting the climate crisis. Effects on anxiety are ambiguous: on the one hand, more knowledge about the negative future consequences of climate change might increase anxiety, but on the other hand, a better understanding of the phenomenon might reduce anxiety. We also predict the

interventions will increase interest and effort in STEM classes.

We expect larger effects from the midline intervention compared to the (lighter touch) baseline intervention.

Heterogeneity

We expect the effects on STEM interest to be larger for girls than boys, given the possibility that girls' education and career choices are more heavily influenced by prosocial goals.

If the baseline has any statistically significant effect on our outcomes of interest, we expect the midline intervention to have larger effects on schools that also received the baseline intervention (that is, complementarity between the two).

Patterns of heterogeneity by district-level climate risk could go in either direction. On the one hand, those exposed to more climate risk might have more experience with and therefore awareness of climate change, which would make these information interventions less effective. On the other hand, those exposed to climate risk might be more responsive to the information interventions because their firsthand experience motivates them to absorb and retain the material better.

5. Estimation methodology

5.1 Effects of both interventions at endline

The main focus of the paper will be on outcomes measured during the endline ($t=2$) and phone follow-up ($t=3$) surveys, for which we will be able to estimate the effects of both interventions. We will use the following ANCOVA specification, estimating separate regressions for $t=2$ and $t=3$:

$$Y_{it} = \beta_{t1} \text{Intensive}_j + \beta_{t2} \text{Light}_j + \gamma_{t1} Y_{i0} + \gamma_{t2} X_{ij0} + \mu_d + \epsilon_{ijt}, \quad (1)$$

where Y_{it} represents the relevant outcome variable for individual i (in school j in district d) measured at time t . The main outcomes (described above) will be: climate change knowledge, skepticism, anxiety, willingness to contribute, interest in science, intent to pursue the science track, and math scores.

The main coefficient of interest is β_{t1} , the coefficient on ***Intensive_j***, which is an indicator equal to 1 for schools that received the intensive midline treatment (groups “Intensive Only” and “Both” in Table 1). This provides us with the average effect of the intensive intervention on outcomes at time t . We are also interested in β_{t2} , the coefficient on an indicator for the light-touch baseline treatment (***Light_j***), which is equal to 1 for groups “Light Only” and “Both.” This provides us with the average effect of the light-touch intervention on outcomes at time t .

To control for sampling strata, we include district fixed effects. A vector of controls, X_{ij0} , will be chosen using the post double selection lasso method proposed by Belloni et al (2014) from the following set of possible controls:

- female dummy
- age dummies (under 12 or missing, 13, 14, and over 14)
- mother's education dummies (below secondary, secondary or higher, or missing)
- father's education (below secondary, secondary or higher, or missing)
- household wealth index (generated using principal components analysis),
- teacher climate knowledge (averaged across all surveyed teachers in school)
- teacher climate skepticism (averaged across all surveyed teachers in school)
- teacher climate anxiety (averaged across all surveyed teachers in school)
- teacher gender (female share of surveyed teachers in school)
- teacher age (average age of surveyed teachers in school)
- teacher education (average share of surveyed teachers in school with a Masters' degree)
- baseline values of other outcome variables (knowledge, skepticism, anxiety, willingness to contribute, interest in science, intent to pursue the science track, and math scores)

To address potential missing values in Y_{i0} (the outcome variable measured at baseline) and continuous X_{ij0} variables, we will replace all missing values with the sample mean and control for a dummy indicating missing values. Standard errors will be clustered by school, the unit of randomization.

In order to examine whether there are any complementarities between the baseline and midline interventions, we will estimate the following specification:

$$Y_{it} = \beta_{t1} \text{LightOnly}_j + \beta_{t2} \text{IntensiveOnly}_j + \beta_{t3} \text{Both}_j + \gamma_{t1} Y_{i0} + \gamma_{t2} X_{ij0} + \mu_d + \epsilon_{ijt}, \quad (2)$$

where LightOnly_j , IntensiveOnly_j , and Both_j represent dummies for the respective groups. β_{t1} provides the effect of the light-touch intervention alone, β_{t2} provides the effect of the intensive intervention alone, and β_{t3} provides the effect of the combination of interventions. We will test for the significance of $\beta_{t3} - \beta_{t2} - \beta_{t1}$, which will provide evidence on how the two interventions interact. If this term is significantly positive, this indicates the two interventions are complementary.

5.2 Short-run effects of baseline interventions

Using data collected at midline ($t=1$), we will also be able to estimate the short-term effect of the light-touch intervention (administered at $t=0$). We will use the following ANCOVA specification:

$$Y_{it} = \beta_{t2} \text{Light}_j + \gamma_{t1} Y_{i0} + \gamma_{t2} X_{ij0} + \mu_d + \epsilon_{ijt}, \quad (3)$$

where the variables are defined as above.

5.3 Heterogeneity

To examine heterogeneity by gender, we will split the sample into male and female students. To examine heterogeneity by climate risk, we will split the sample into low-risk and high-risk districts. For both analyses, we will estimate regressions (1), (2), and (3) separately for each sub-group, using the same vector of controls selected in the pooled regression for both sub-groups. We will report tests of equality of the β coefficients across subgroups.

5.4 Attrition and Missing Data

Missing outcome variables in follow-up waves could result from respondents leaving the sample as well as from non-response to certain questions. We test for selective non-response by estimating the following regression for each outcome variable Y_{it} , where N_{it}^Y is an indicator equal to 1 if outcome Y_{it} is non-missing.

$$N_{it}^Y = \alpha_{t1} \text{LightOnly}_j + \alpha_{t2} \text{IntensiveOnly}_j + \alpha_{t3} \text{Both}_j + \alpha_{t4} Y_{i0} + \alpha_{t5} X_{ij0} + \mu_d + \epsilon_{ijt}. \quad (4)$$

X_{ij0} represents the same vector of controls used in equation (2) above. We will cluster our standard errors at the school level. If any of the α coefficients are individually significant or if they are jointly significant at the 10% level, we will use inverse probability weighting to deal with selective non-response. To construct (outcome-specific) weights, we will estimate an expanded version of equation (4) as a probit regression, adding interactions between each treatment dummy and Y_{i0} and X_{ij0} . Weights will be calculated as $\frac{1}{N_{it}^Y}$, the inverse of the predicted probability of a non-missing outcome variable.

References

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2), 608-650.